



**HAL**  
open science

# Full-band quantum transport simulation of advanced nanodevices

Sylvan Brocard

► **To cite this version:**

Sylvan Brocard. Full-band quantum transport simulation of advanced nanodevices. Micro and nanotechnologies/Microelectronics. Université de Grenoble, 2014. English. NNT : 2014GRENT117 . tel-01317452

**HAL Id: tel-01317452**

**<https://theses.hal.science/tel-01317452>**

Submitted on 19 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Nano électronique, nano technologies**

Arrêté ministériel : 7 août 2006

Présentée par

**Sylvan BROCARD**

Thèse dirigée par **Marco PALA** et  
codirigée par **Gérard GHIBAUDO**

préparée au sein du **Laboratoire IMEP-LAHC**  
dans l'**École Doctorale EEATS**

# Full-band quantum transport simulation of advanced nanodevices

Thèse soutenue publiquement le **20 octobre 2014**,  
devant le jury composé de :

**Dr. Francis BALESTRA**

Directeur de recherche au CNRS, IMEP-LAHC, Examineur

**Prof. Giuseppe IANNACCONE**

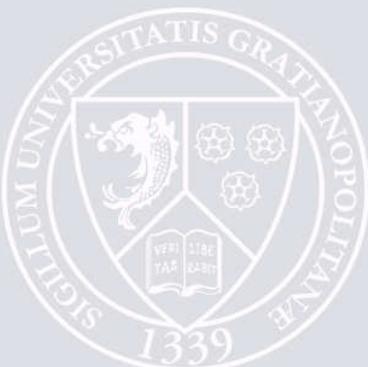
Professeur à l'Université de Pise, Rapporteur

**Dr. Marc BESCOND**

Chargé de recherche au CNRS, IM2NP, Rapporteur

**Dr. Marco PALA**

Chargé de recherche au CNRS, IMEP-LAHC, Directeur de thèse



# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	Trends in semiconductor technology . . . . .	7
1.2	Tunneling field-effect transistors . . . . .	8
1.3	Purpose of this work . . . . .	9
<b>2</b>	<b>Electronic band structure</b>	<b>13</b>
2.1	Quantum Mechanics . . . . .	14
2.1.1	Crystallographic structure . . . . .	15
2.1.2	Calculating the band structure . . . . .	16
2.2	Atomistic Empirical Pseudopotentials method . . . . .	17
2.2.1	Symmetrical and antisymmetrical form factors . . . . .	19
2.2.2	Bulk parametrization . . . . .	20
2.2.3	Atomistic approach . . . . .	21
2.2.4	Strain . . . . .	22
2.2.5	Low-field mobility calculations . . . . .	23
2.3	The $\mathbf{k} \cdot \mathbf{p}$ approximation method . . . . .	26
2.3.1	Equations . . . . .	26
2.3.2	8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian . . . . .	27
2.3.3	Quantization . . . . .	28
2.3.4	Strain . . . . .	29
<b>3</b>	<b>Full-quantum simulations</b>	<b>31</b>
3.1	Context . . . . .	32
3.2	Self-consistent Poisson-Schrödinger solution . . . . .	33
3.3	The Non Equilibrium Green's Function formalism . . . . .	34
3.3.1	The Dyson equation . . . . .	38

3.3.2	A simple example . . . . .	39
3.3.3	A dot attached to a lead . . . . .	40
3.4	The recursive Green's function algorithms . . . . .	41
3.4.1	Recursive scheme for the lesser-than Green's function . . . . .	42
3.5	The coupled mode-space approach . . . . .	43
3.6	The electron-phonon interaction . . . . .	45
3.6.1	The self-consistent Born approximation . . . . .	47
<b>4</b>	<b>Comparison between Si and InAs in FinFETs</b>	<b>51</b>
4.1	III-V materials for nanoscale electronics . . . . .	52
4.1.1	FinFET as a solution for transistors scaling . . . . .	53
4.2	Devices under investigation . . . . .	54
4.3	Simulation results . . . . .	55
4.3.1	Impact of the device electrostatics . . . . .	55
4.3.2	Supply-voltage scaling . . . . .	58
4.4	Conclusion . . . . .	59
<b>5</b>	<b>Design of nanowire TFETs</b>	<b>61</b>
5.1	Technological context . . . . .	62
5.2	Nanowire geometry . . . . .	64
5.2.1	Nanowire width comparison . . . . .	65
5.2.2	Nanowire length comparison . . . . .	66
5.3	Impact of the hetero-junction . . . . .	68
5.4	Impact of strain engineering . . . . .	69
5.4.1	Analysis results . . . . .	71
5.5	SS degradation induced by traps . . . . .	74
5.6	Conclusion . . . . .	76
<b>6</b>	<b>Source region engineering of III-V Tunnel-FETs</b>	<b>79</b>
6.1	Tunneling path modification . . . . .	80
6.2	Use of the molar grading . . . . .	81
6.2.1	Simulation results . . . . .	81
6.2.2	Physical interpretation . . . . .	86
6.3	Insertion of a quantum well . . . . .	90
6.3.1	Role of the quantum well . . . . .	93

6.4 Conclusion . . . . .	94
<b>7 Conclusions</b>	<b>97</b>



# Acknowledgements

I would like to thank my supervisor for his indestructible support, help, mentoring, and patience through this PhD program. His sharp understanding and kindness never fail to inspire me.

I would also like to thank Denis Rideau for his contribution since the beginning of this project, his tutoring, and the pleasure of his company in partnership.

I am grateful to David Esseni for the active interest he took in this work, and for his inception of fruitful ideas.

My thanks extend to the IMEP-LAHC staff and other students for their warm welcome.

Finally I will express my love to my parents and family for their support of my endeavors.



# Chapter 1

## Introduction

Following the invention of the point contact transistor in 1947 by Bardeen and Brattain, and two years before his demonstration of the first bipolar junction transistor, William Shockley published the supporting theory, based on the drift-diffusion (DD) model. While at first the compact models based on drift-diffusion allowed for approximate analytical solutions, the fact that an analytical solution to the DD equations only exists under specific conditions drove the theoretical effort toward numerical solutions and led to the rise of Technology Computer Aided Design (TCAD).

Coincidentally, experimental and manufacturing results got farther from the predictions of simple analytical models. Device simulation offers both a way to predict the result of novel device architectures, and an insight into device properties that cannot be directly measured, furthering our understanding of the transport process.

As the carrier confinement in modern devices increased, it became increasingly clear that the DD model was insufficient to describe devices operation properly. This was addressed first by adding quantum corrections to the DD model, then to microscopic transport models based off the Boltzmann Transport Equation, to today's quantum transport models.

### 1.1 Trends in semiconductor technology

Given that the semiconductor industry is definitively moving away from bulk architectures, and toward planar and wire architectures (as illustrated in Fig. 1.1),

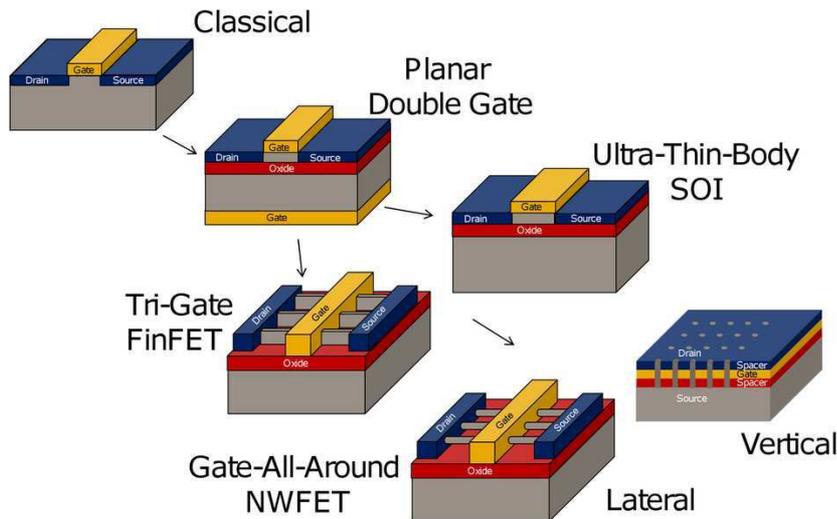


Figure 1.1: Evolution of the Field Effect Transistor (FET) Architecture. Source: [1]

mostly in an effort to improve electrostatic control of the channel, accurate nanoscale models become a necessity to investigate those new devices. Even though the existence of quantum effects in devices like dielectric tunneling was known since the 1950s, quantum models are a fairly recent endeavor, mainly because of the large computational effort involved. So far, those simulations have been focused on unstrained  $\langle 110 \rangle$ -oriented materials. However, whereas the scaling of the gate length used to be the prime factor of device improvement, alternative solutions have emerged to boost performances, such as strain engineering and crystal orientation. A suitable model should account for those technological boosters.

## 1.2 Tunneling field-effect transistors

More radically, since MOSFETs are approaching their ultimate performance limit, alternate device architectures are under investigation, like Tunnel FETs (TFET).

The principle of TFET technology is to build a p-i-n diode, with a band alignment such that the thermionic conduction is completely suppressed, as sketched in Fig. 1.2. Instead, the entirety of the current is driven by band-to-band tunneling

from the source valence states to the channel conduction states. By controlling the electrostatic state of the channel via a gate, it is possible to shorten the tunneling path the carriers have to take, reducing the transistor resistance and thus switching to the on state. There is no theoretical limit on how fast the on-off state transition can take place.

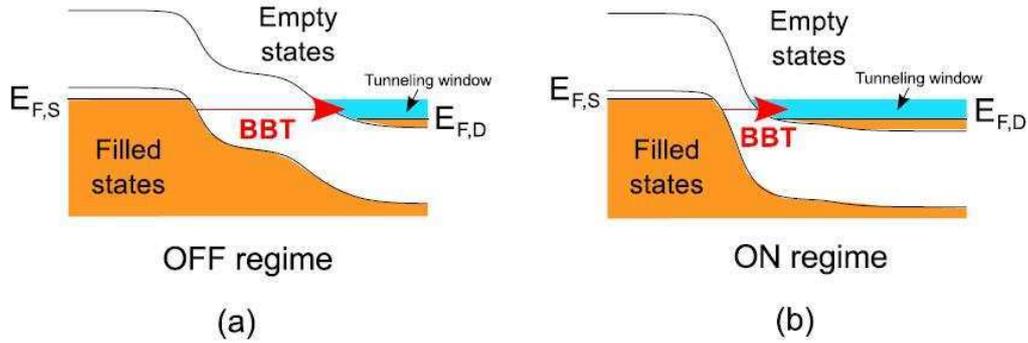


Figure 1.2: Sketch of the valence and conduction subband edges for a TFET in the on (a) and off (b) states. Charge carriers tunnel from the valence band in the source region to the conduction band in the drain (a) or channel (b). This makes the tunneling distance considerably shorter in the on state.

The main advantage of TFET over conventional transistors is that they can overcome the theoretical inverse subthreshold slope (SS) limit of 60mV/decade stifling MOSFET technology. However, as represented in Fig. 1.3, they typically display a lower on-current  $I_{on}$  than MOSFETs, since the current is due to band-to-band tunneling.

### 1.3 Purpose of this work

Since TFETs entirely rely on quantum phenomenon, quantum models are the only option to further learn about them. In particular, exact band-structure plays an essential role in the tunneling process.

Currently, if one is looking for a mature formalism for quantum transport, one will generally opt for the Non-Equilibrium Green's Function (NEGF) formalism.

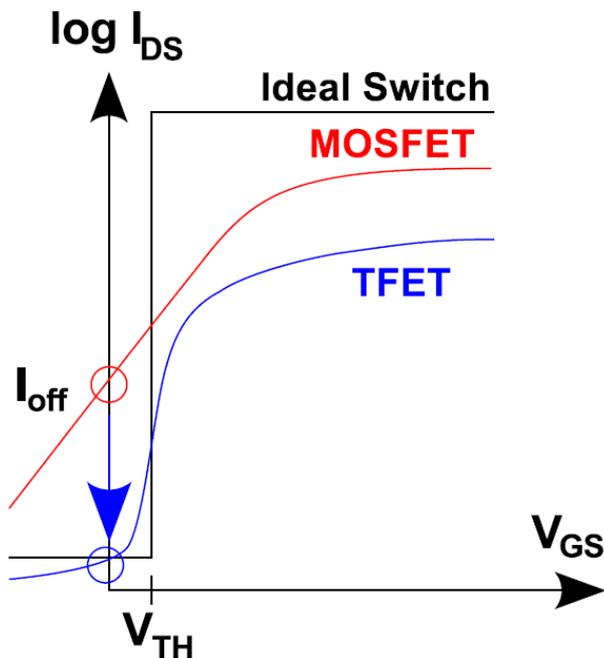


Figure 1.3: Sketch of an ideal switch, a MOSFET and a TFET archetypal  $I_{DS}$  versus  $V_{GS}$  curves. The MOSFET shows a better on-current than the tunnel-FET, but the TFET has a steeper slope, allowing faster on-off switching.

It earned its recognition by accounting for inelastic phonon scattering. One of its strength is that NEGF formalism doesn't depend on form of the Hamiltonian operator.

The most used Hamiltonian in quantum models for band-structure calculation is the effective mass approximation (EMA). However, the use of EMA supposes that the effect of boundary condition is small in regard to the lattice contribution, which is not the case in devices we're interested in. Other popular models include the k.P Hamiltonian and the tight-binding Hamiltonian. The k.p approach, while remaining computationally reasonable, has the flaw of needing a fair amount of parametrization, which is an issue when investigating novel devices, and the tight-binding approach computational burden keeps it from being attractive.

In this work we aim to use the Atomistic Empirical Pseudopotential Method (AEPM) within the NEGF framework to obtain a fully atomistic simulation of

planar and nanowire devices. The atomistic nature of that method allows for the use of any arbitrary strain and crystal orientation without the need for additional parametrization, as well as arbitrary molar fraction in ternary alloys.

In Chapter 2 of this work, we will present the theoretical framework used to compute nanowire bandstructures and different types of Hamiltonians. We will focus on the specificity of  $\mathbf{k} \cdot \mathbf{p}$  approximation and of the AEPM method and its implementation.

Chapter 3 covers the theory of quantum transport within the Non-Equilibrium Green's function method. It is focused on the application of that formalism to self-consistent simulations of nanodevices.

Chapter 4 compares the efficiency of Silicon and InAs in small scale FinFETs, and comments on the scalability of the devices based on these materials.

Chapter 5 is about the exploitation of the previous models to investigate the use of strain and material engineering to improve the performances of TFETs. In particular, it explores the viability of two technological boosters: the use of strain engineering and the use of a III-V materials heterojunction at the source-channel interface.

Chapter 6 introduces two original TFET architectures, in the form of a linear grading of the molar fraction in an AlGaSb ternary alloy in the source, or alternatively as a GaSb potential well in an AlSb source. The operating principle of these devices is described, and their performances are evaluated against conventional TFET design and MOSFET technology.



# Chapter 2

## Electronic band structure

This chapter is meant to expose the theory of full bandstructures calculations.

Nowadays, the effective mass approximation (EMA) is a widely used approach in quantum transport simulation, as its implementation is remarkably simple and its computation burden is low. However, the EMA is fundamentally a perturbative approach, and is meant for situations where the external perturbation varies on a much bigger scale than the lattice periodicity. Now that technology is approaching scales in the vicinity of 5nm, this condition starts to fail. EMA results diverge from both more accurate simulations and experiments, as the importance of band structure effects rises.

For the purpose of this work, we intend to go beyond the effective mass approximation. To that effect, we will use two approaches, namely the  $\mathbf{k} \cdot \mathbf{p}$  approximation method, and the empirical pseudopotential method (EPM). The  $\mathbf{k} \cdot \mathbf{p}$  method is already a tried and true approach, and combines reasonably accurate results with an acceptable computational load. Its main weaknesses reside in the need for extensive parametrization, and its loss of accuracy for large wave vectors. The atomistic EPM (AEPM) formalism is a greatly accurate method for band-structure calculation. It also benefits from a wide versatility but the considerable computational burden usually keeps it from being attractive for most simulators.

Section 2.1 briefly describes the fundamental concepts of quantum mechanics used in this chapter. Section 2.2 describes the AEPM method and how we developed it for this work, and Section 2.3 describes the  $\mathbf{k} \cdot \mathbf{p}$  method that we used for several simulations.

## 2.1 Quantum Mechanics

The quantum mechanics equation describing a semiconductor in equilibrium is the time-independent Schrödinger equation:

$$\mathbf{H}(\mathbf{r})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (2.1)$$

where  $\Psi(\mathbf{r})$  is the electron spatial wave function, and

$$\begin{aligned} \mathbf{H}(\mathbf{r}) = & \sum_i \frac{p_i^2}{2m_e} + \sum_j \frac{P_j^2}{2M_j} + \frac{1}{2} \sum_{j,j'} \frac{Z_j Z_{j'} e^2}{|\mathbf{R}_j - \mathbf{R}_{j'}|} \\ & - \sum_{j,i} \frac{Z_j e^2}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{1}{2} \sum_{i,i'} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_{i'}|}, \end{aligned} \quad (2.2)$$

where  $\mathbf{r}_i$  is the position of the  $i$ th electron,  $\mathbf{R}_j$  is the position of the  $j$ th nucleus,  $p_i$  and  $P_j$  are the momentum operator of the electrons and nuclei respectively, and  $Z_j$  is the atomic number of the  $j$ th nucleus.

It is not possible to fully solve this many-bodies Hamiltonian, so we use several approximations [2]. The first one consists of separating electrons into valence electrons and core electrons. Core electrons potentials are added together with nuclei potentials, to result in an effective ionic pseudopotential.

Then we invoke the Born-Oppenheimer approximation, which states that since ions are significantly heavier than electrons, their orbit is much slower. For all intent and purpose, they can be considered almost stationary when solving the Schrödinger equation for electrons. A perturbative Hamiltonian can thus be written as

$$\mathbf{H} = \mathbf{H}_{\text{ions}}(\mathbf{R}_j) + \mathbf{H}_e(\mathbf{r}_i, \mathbf{R}_{j0}) + \mathbf{H}_{e-\text{ions}}(\mathbf{r}_i, \delta\mathbf{R}_j), \quad (2.3)$$

where  $\mathbf{H}_{\text{ions}}$  is the Hamiltonian governing the ionic potential,  $\mathbf{H}_e(\mathbf{r}_i, \mathbf{R}_{j0})$  is the interaction between the electrons and ions at equilibrium, and  $\mathbf{H}_{e-\text{ions}}(\mathbf{r}_i, \delta\mathbf{R}_j)$  is the term governing electron-phonon interaction, which will be addressed in the next Chapter. The purpose of this Chapter is to compute the electronic Hamiltonian  $\mathbf{H}_e$ , given by

$$\mathbf{H}_e = \sum_i \frac{p_i^2}{2m_e} - \sum_{j,i} \frac{Z_j e^2}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{1}{2} \sum_{i,i'} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_{i'}|}. \quad (2.4)$$

To simplify further, we employ the mean-field approximation, by assuming that every electron experiences the same spatial potential  $V(\mathbf{r})$ . Hence, the Schrödinger equation for a single electron in an electron gas becomes

$$\mathbf{H}\Psi_n(\mathbf{r}) = \left( \frac{p^2}{2m_e} + V(\mathbf{r}) \right) \Psi_n(\mathbf{r}) = E_n \Psi_n(\mathbf{r}), \quad (2.5)$$

where  $n$  labels the different solutions of the equation.

### 2.1.1 Crystallographic structure

The semiconductors considered in this study are organized at the atomic level with either the diamond (for silicon or germanium) or zinc-blende structure (for III-V semiconductors). Both of these structures follow a face-centered cubic (fcc) lattice, with half the tetrahedral sites occupied by an extra atom. It is possible to construct this structure by starting from a fcc lattice with a single ion type, and adding another fcc lattice (identical for diamond structures, or with the other ion type for a zinc-blende structure), shifted by one quarter of the cubic cell diagonal in the  $[111]$  direction (see Fig. 2.1). Both these lattices elementary cells are defined by the Bravais vectors

$$\begin{aligned} \mathbf{a}_1 &= \frac{a}{2}(1, 1, 0) \\ \mathbf{a}_2 &= \frac{a}{2}(0, 1, 1) \\ \mathbf{a}_3 &= \frac{a}{2}(1, 0, 1), \end{aligned} \quad (2.6)$$

where  $a$  is the lattice parameter. This elementary cell contains two atoms. Its volume is  $\Omega = a^3/4$ .

From the basis vectors, we define the reciprocal lattice vectors as

$$\begin{aligned} \mathbf{b}_1 &= \frac{2\pi}{a}(1, -1, 1) \\ \mathbf{b}_2 &= \frac{2\pi}{a}(1, 1, -1) \\ \mathbf{b}_3 &= \frac{2\pi}{a}(-1, 1, 1). \end{aligned} \quad (2.7)$$

The first Brillouin zone of the fcc lattice is shown in Fig. 2.2, along with the usual notations for the remarkable symmetry points.

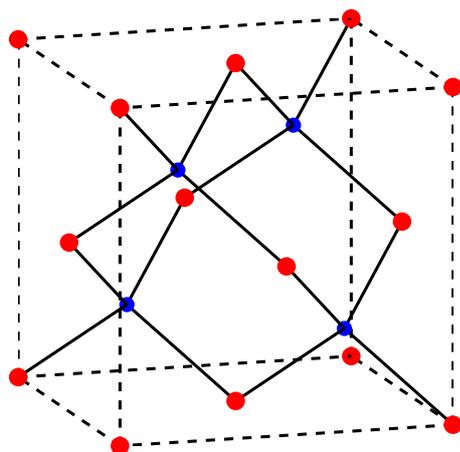


Figure 2.1: Zinc-Blende structure

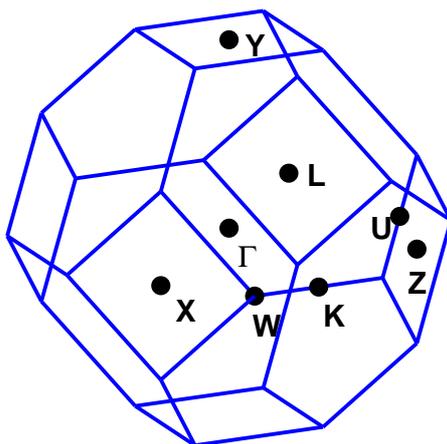


Figure 2.2: Bulk fcc first Brillouin zone and labelled high symmetry points

### 2.1.2 Calculating the band structure

To analyze the transport properties of a nano-scale device, we have to numerically solve the electron Hamiltonian as accurately as possible, while retaining a feasible

computational burden.

There exist so-called *ab initio* methods to compute band structures. They rely only on first principles to compute the Hamiltonian of a quantum system, like the density-functional theory, or the Hartree-Fock method. They are however too computationally expensive to be applied to anything but the smallest systems. To solve that issue, a number of semi-empirical approaches were created. These methods combine a certain amount of physical modeling with physical measurements used to fit the model parameters. They include the  $\mathbf{k} \cdot \mathbf{p}$  method, the empirical pseudopotentials method, and the tight-binding method, the first two of which are the topic of the rest of this chapter.

## 2.2 Atomistic Empirical Pseudopotentials method

The concept of pseudopotential was first introduced by Fermi in 1934 [3]. The essential idea is that the valence electrons in a crystal are submitted to an attractive potential from the nuclei, and to a repulsive potential from the core electrons that keep them outside of the core. When added, these two potentials result in a so-called pseudopotential. This pseudopotential is weaker than a Coulomb potential, as the nucleus and core potentials almost cancel each other. Fig. 2.3 is a real-space representation of the pseudopotential for a Si atom. This leaves us free to consider the valence electrons to be nearly-free in the crystal. When solving the electronic system, using the pseudopotentials instead of the true potentials remove the core electron states from the solutions, which is to our advantage since these states do not contribute to the conduction properties of a semiconductor. It does not affect the valence states in the solution, and is justified by the fact that the valence electrons contribution to the core potential is negligible.

In a periodic crystal, the electron wave functions are Bloch functions, as in

$$\Psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r}). \quad (2.8)$$

where  $u_{n,\mathbf{k}}(\mathbf{r})$  is periodic along all three axes of the crystal. Thus the wave function can be expressed in the reciprocal lattice as

$$\Psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} A_{\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (2.9)$$

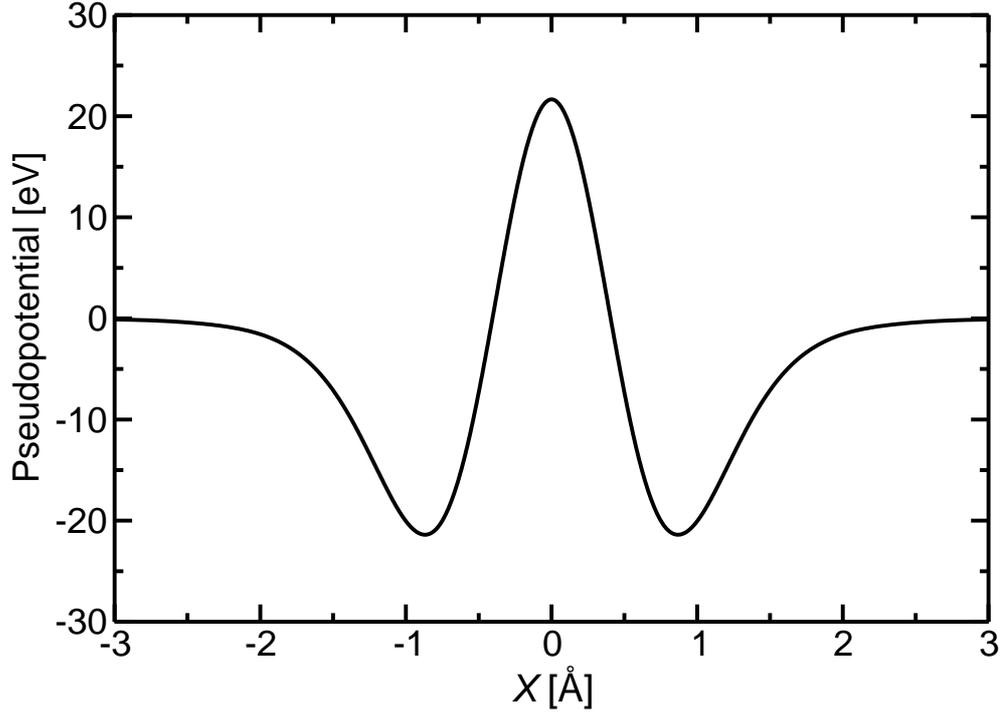


Figure 2.3: Real space representation of the Si atom pseudopotential used in this study.

where the  $\mathbf{G}$  are the reciprocal lattice vectors and the  $A_{\mathbf{k}}$  are the expansion coefficients of the periodic component of the Bloch function.

Similarly, the pseudopotential can be expanded in the reciprocal lattice

$$V(\mathbf{r}) = \sum_{\mathbf{G}} (V_{\mathbf{G}\alpha} S_{\alpha}(\mathbf{G}) + V_{\mathbf{G}\beta} S_{\beta}(\mathbf{G})) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (2.10)$$

where

$$S_{\alpha}(\mathbf{G}) = \frac{1}{N_{\alpha}} \sum_i e^{-i\mathbf{G}\cdot\mathbf{r}_{\alpha,i}}, \quad (2.11)$$

is referred to as the structure factor for the first ion type,  $S_{\beta}(\mathbf{G})$  is defined similarly for the second ion type, and  $V_{\mathbf{G}\alpha}$ ,  $V_{\mathbf{G}\beta}$  are the form factors of the two ions types, defined by

$$V_{\mathbf{G}\alpha} = \frac{1}{\Omega} \int_{\Omega} V_{\alpha}(\mathbf{r}) e^{-i\mathbf{g}\cdot\mathbf{r}} d\mathbf{r}, \quad (2.12)$$

where  $V_{\alpha}$  and  $V_{\beta}$  are the pseudopotentials for the two ion types.

Leading to the matrix representation of Eq. 2.5.

$$\frac{\mathbf{A}_k(\mathbf{G})}{2}(\mathbf{k} + \mathbf{G})^2 + \sum_{\mathbf{G}'} \mathbf{V}_{\mathbf{G}-\mathbf{G}'} \mathbf{A}_k(\mathbf{G}') = E_k \mathbf{A}_k(\mathbf{G}) \quad (2.13)$$

in which  $E_k$  are the energy solutions and  $\mathbf{A}_k(\mathbf{G})$  are the corresponding eigenvector wavefunctions.

### 2.2.1 Symmetrical and antisymmetrical form factors

In a zinc-blend structure, let's take the midpoint between the a and b ions as the origin in the unit cell. The ions coordinates can be written as  $\mathbf{r}_\alpha = (a/8)(1, 1, 1) = \tau$  and  $\mathbf{r}_\beta = (a/8)(-1, -1, -1) = -\tau$ . The general term in Eq. (2.10) can be expressed as

$$\begin{aligned} & (V_{\mathbf{G}\alpha} S_\alpha(\mathbf{G}) + V_{\mathbf{G}\beta} S_\beta(\mathbf{G})) e^{i\mathbf{G}\cdot\mathbf{r}} \\ &= \frac{1}{\Omega} \int_{\Omega} (V_{\mathbf{G}\alpha} e^{i\mathbf{G}\cdot\mathbf{r}_\alpha} + V_{\mathbf{G}\beta} e^{i\mathbf{G}\cdot\mathbf{r}_\beta}) e^{i\mathbf{G}\cdot\mathbf{r}} d\mathbf{r} \\ &= \frac{1}{\Omega} \int_{\Omega} ((V_{\mathbf{G}\alpha} + V_{\mathbf{G}\beta}) \cos(i\mathbf{G} \cdot \tau) - i(V_{\mathbf{G}\alpha} - V_{\mathbf{G}\beta}) \sin(i\mathbf{G} \cdot \tau)) e^{i\mathbf{G}\cdot\mathbf{r}} d\mathbf{r} \end{aligned} \quad (2.14)$$

The reciprocal space potential can thus be written as

$$V_{\mathbf{G}} = V_{\mathbf{g}}^s \cos(\mathbf{G} \cdot \tau) + iV_{\mathbf{g}}^a \sin(\mathbf{G} \cdot \tau), \quad (2.15)$$

where

$$V_{\mathbf{g}}^s = \frac{1}{\Omega} \int_{\Omega} (V_\alpha(\mathbf{r}) + V_\beta(\mathbf{r})) e^{-i\mathbf{g}\cdot\mathbf{r}} d\mathbf{r} \quad (2.16)$$

is the symmetric form factor of the total pseudopotential and

$$V_{\mathbf{g}}^a = \frac{1}{\Omega} \int_{\Omega} (V_\alpha(\mathbf{r}) - V_\beta(\mathbf{r})) e^{-i\mathbf{g}\cdot\mathbf{r}} d\mathbf{r} \quad (2.17)$$

is the antisymmetric form factor. These form factors can be adjusted based on measurements, hence the "empirical" in Empirical Pseudo Potential Method. In a diamond structure,  $V^a$  is always null. Fig. 2.4 shows the symmetric form factor for Si used in this study. Since the Si structure is symmetrical, it can also be seen as the Fourier transform of the Si pseudopotential from Fig. 2.3.

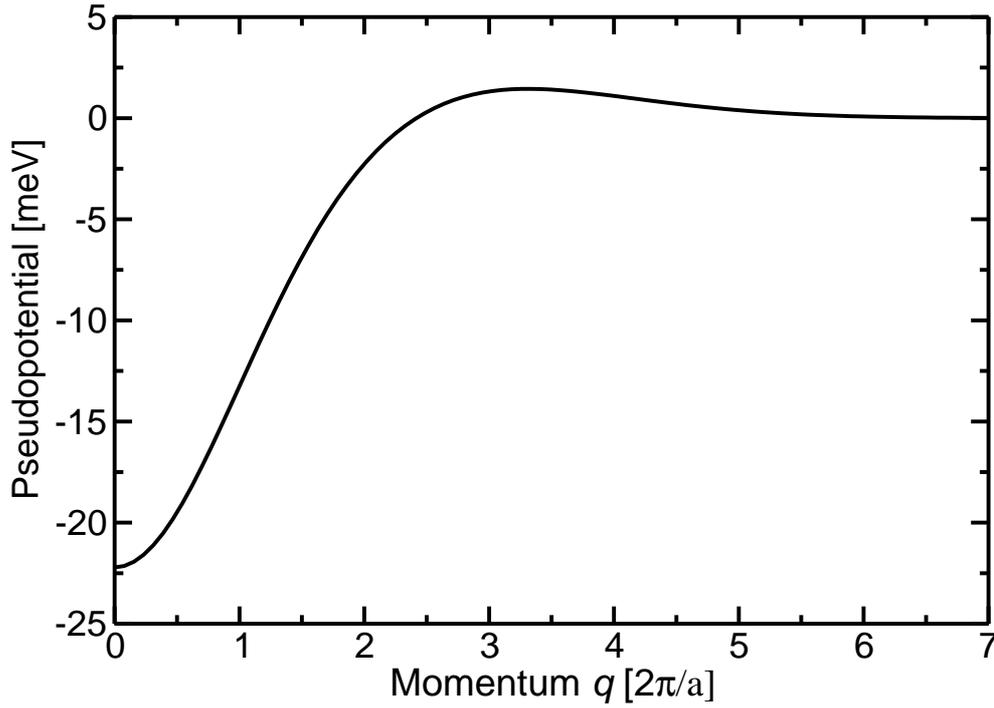


Figure 2.4: Plot of the Si form factor used in this study.

### 2.2.2 Bulk parametrization

In bulk, only the values of the form factors on reciprocal lattice vectors are relevant. Since the form factors are spherically symmetric their values are labeled according to the squared amplitude of the reciprocal lattice vector, in units of  $2\pi/a$ :

$$\begin{aligned}
 \mathbf{V}_0^s &= \mathbf{V}^s(0, 0, 0) \\
 \mathbf{V}_3^s &= \mathbf{V}^s(1, 1, 1) = \mathbf{V}^s(-1, 1, 1) = \mathbf{V}^s(1, -1, 1) = \dots \\
 \mathbf{V}_4^s &= \mathbf{V}^s(2, 0, 0) = \mathbf{V}^s(-2, 0, 0) = \mathbf{V}^s(0, 2, 0) = \dots \\
 &\dots
 \end{aligned} \tag{2.18}$$

Experimental results show that the form factors decrease drastically after a certain value therefore we can ignore  $V_{\mathbf{g}}$  for  $\mathbf{g}^2 > 11(2\pi/a)^2$ .  $V_0$  exact value is unimportant because it only contributes a constant term to the potential, and therefore an uniform energy shift on the energies scale.  $V_4^s$  can also be ignored because  $\cos(\mathbf{g} \cdot \mathbf{s})$  is null for  $\mathbf{g} = (2, 0, 0)(2\pi/a)$  and every other  $\mathbf{g}^2 = 4(2\pi/a)^2$

permutation. Similarly,  $\sin(\mathbf{g} \cdot \mathbf{s})$  is null whenever  $\mathbf{g}^2 = 8(2\pi/a)$  so  $V_8^a$  is an unimportant value. Thus only 3 form factor values are relevant for the computing of energy bands in bulk diamond structure materials:  $V_3^s$ ,  $V_8^s$  and  $V_{11}^s$ . For III-V semiconductors,  $V_3^a$ ,  $V_4^a$  and  $V_{11}^a$  also matter.

	$V_3^s$	$V_8^s$	$V_{11}^s$	$V_3^a$	$V_4^a$	$V_{11}^a$
Si	-0.224	0.055	0.072	0.000	0.000	0.000
InAs	-0.230	0.000	0.045	0.055	0.045	0.010
GaSb	-0.220	0.005	0.045	0.040	0.030	0.000

Table 2.1: Bulk pseudopotential form factors for some of the materials considered in this study. Source: [4].

Since pseudopotential band structure calculations rely on a small number of parameters, it is possible to fit them from experimental data. In the seminar paper [4], Chelikowsky and Cohen extracted pseudopotential parameters for multiple diamond and zinc-blende semiconductors that are still commonly in use today.

It should be noted that we have so far assumed a local potential. In reality there is a quasiparticle nonlocal self-energy term that enters into the diagonal elements of the pseudopotential matrix. Without taking it into account, pseudopotential calculations give inaccurate bandgap values. Cohen and Heine in [5] proved that this phenomenon can be approximated by adding a constant multiplier  $\beta$  to the kinetic energy term in the Schrödinger equation. Hence, Eq. (2.5) becomes

$$\mathbf{H}\Psi_n(\mathbf{r}) = \left( \frac{\beta p^2}{2m_e} + V(\mathbf{r}) \right) \Psi_n(\mathbf{r}) = E_n \Psi_n(\mathbf{r}). \quad (2.19)$$

This term itself is to be considered as another fitting parameter. We use the value of  $\beta = 1.205$  given in Ref. [6] for InAs/GaSb structures.

### 2.2.3 Atomistic approach

The previous section considered the case of a bulk material. In case of confined systems, heterojunctions, and superlattices, it is possible to use the pseudopotential method by constructing the pseudopotential matrix term in Eq. (2.13)

for the whole system instead of an elementary cell, and solving it. The atomistic approach consists of recovering the independent ionic form factors from the symmetric and antisymmetric components obtained in bulk measurements, converting them to real-space pseudopotentials, and summing them in real space according to the ionic coordinates to obtain a system that is then solved for energies. The nanowire surface is passivated with pseudo-hydrogen atoms in the empty tetrahedral sites.

In order to convert form factors from reciprocal space to real space, we have to evaluate their value over the whole range of  $\mathbf{q}$  amplitude, from  $\mathbf{q}^2 = 0$  to  $\mathbf{q}^2 = 11$ , and not just for specific reciprocal lattice values like in the bulk case. Since the values obtained for experimental results are always for bulk materials, it is common practice to interpolate the form factor coefficients. A satisfying interpolation function should vanish faster than  $\mathbf{q}^{-2}$  over the cutoff  $\mathbf{q}$ , as well as having a null or at least small derivative at  $\mathbf{q} = 0$ . Various interpolation functions have been proposed, like

$$V_q = a_1[\cos(a_2q) + a_3]e^{a_4q^4}/q^2 \quad (2.20)$$

in [7] and

$$V_q = a_1(q^2 - a_2)/(a_3e^{a_4q^2} - 1) \quad (2.21)$$

in [8]. Further, we will use the latter, mostly because it has the merit of producing a smoother pseudopotential function in the real space. Since we are working in real space, we are not concerned with limiting our  $q$  range, as it does not affect the computational burden.

For silicon, and for pseudo-hydrogen we use the parameters from Ref. [8]. For III-V compounds, we use the ones from [6]. They are listed in Table 2.2. Examples of band structures extracted from these parameters are shown in Fig. 2.5

### 2.2.4 Strain

There are two effects of strain on the calculated energy band structures. One is the displacement of the ions itself. Since the atomistic potential function already includes the ionic coordinates, this is taken into account naturally. The other comes from the fact that AEPM calculations are not self-consistent, thus they ignore the effect of local strain on the electron shell. Ref. [9] introduces an

	$a_1$	$a_2$	$a_3$	$a_4$
Si	0.2685	2.19	2.06	0.487
In	139.797	1.559	2.603	0.433
As	23.104	3.106	1.286	0.362
Ga	544918.497	1.752	7281.008	0.907
Sb	266.483	2.154	3.003	0.877

Table 2.2: Form factors fitting parameters for the materials considered in this study.

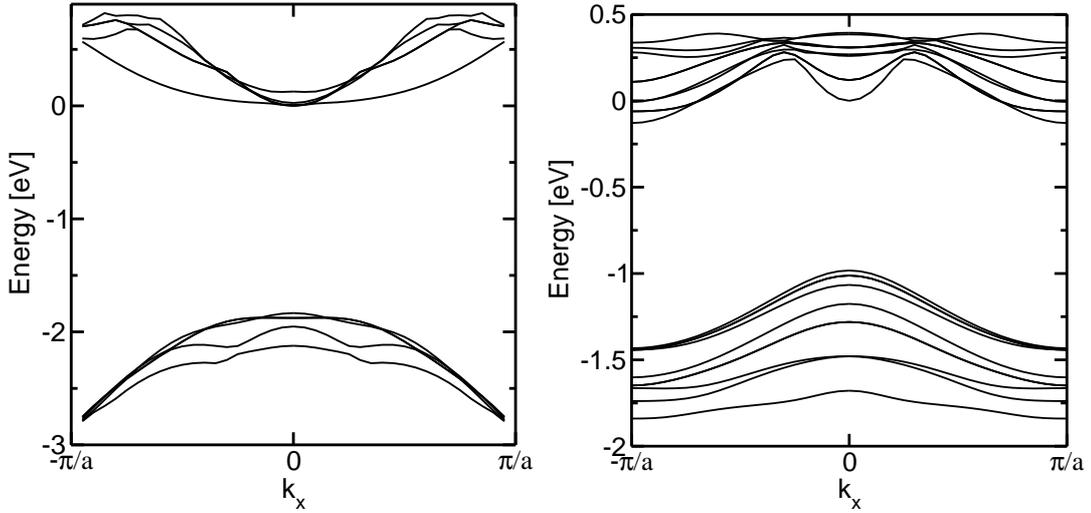


Figure 2.5: Example of conduction and valence bands obtained from the parameters listed in Table 2.2, in silicon (left) and InAs (right) square nanowires with a 2.2nm width.

additional fitting parameter  $\gamma_\alpha$  such that

$$V(q, \epsilon) = V(q, 0) [1 + \gamma_\alpha Tr(\epsilon)] , \quad (2.22)$$

where  $Tr(\epsilon)$  is the local hydrostatic strain at the atom position.

### 2.2.5 Low-field mobility calculations

When the AEPM Hamiltonian is discretized in the real space results in a matrix of rank  $N_x N_y N_z$ , where  $N_i$  is the number of discretized points along the  $i$ -th spatial direction. These numbers are usually very large even for narrow nanowires

because the mesh step size has to be small enough to properly describe the atomistic pseudopotentials as the one shown in Fig. 2.3. A typical value for the mesh step size is  $0.2\text{\AA}$ . Hence, the real-space AEPM Hamiltonian cannot be exploited to compute transport properties such as the transfer and output characteristics of nanowire transistors.

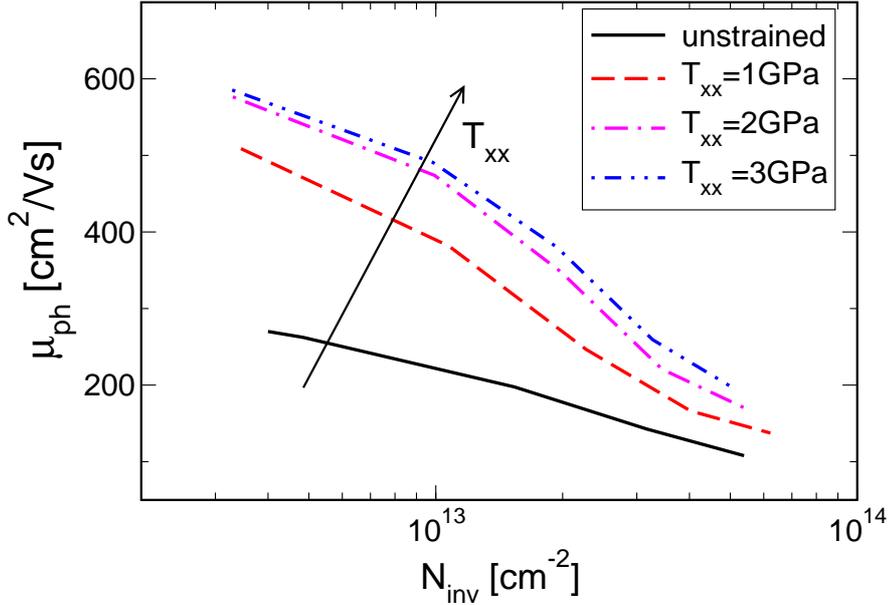


Figure 2.6: Example of phonon-limited mobility computed with the AEPM approach for a square Si nanowire with a  $5 \times 5\text{nm}^2$  cross section and length of 40nm. The uniaxial stress varies from 0 to 3 GPa. The transport direction is [100].

During the PhD work these Hamiltonians have been used mainly to perform bandstructure calculations and to validate simpler Hamiltonians such as the ones based on the  $\mathbf{k} \cdot \mathbf{p}$  approximation (see Sec. 2.3). However, we have also developed an approximated method to exploit the AEPM Hamiltonian to evaluate the low-field mobility of squared Si nanowires in the presence of phonon scattering and different levels of mechanical strength. We remark here that due to the large computational burden in terms of memory occupation and CPU time, these calculations can be performed only for simple structures with homogeneous interfaces and in the absence of defects. Since the real-space Hamiltonian  $H_{\text{RS}}$

cannot be handled due to memory constraints, we resolved to approximate it in a reduced basis. This was obtained by performing the unitary transformation  $H_{\text{SS}} = U_{\text{SS}}^\dagger H_{\text{RS}} U_{\text{SS}}$ , where  $H_{\text{SS}}$  is the Hamiltonian in a reduced subband space defined by the solution of the bandstructure problem for the nanowire unitary cell and  $U_{\text{SS}}$  is the unitary transformation matrix given by subbands at different wave vector points from 0 to  $\pi/a$ . This transformation is a generalization of the coupled-mode space approach that will be discussed in Chapter 3. Such an approximated representation of the Hamiltonian can work properly only if the energies important for the transport are close to the conduction band edge.

By using this reduced matrix, it is possible to largely soften the computational burden, which will allow us to solve a simple transport problem like the calculation of the low-field mobility of a uniform Si nanowire. The low-field mobility of such devices can be evaluated by following the prescription of [35, 28]. In order to extract the low-field mobility of such nanowires we have used the NEGF methods which will be detailed in Chapter 3.

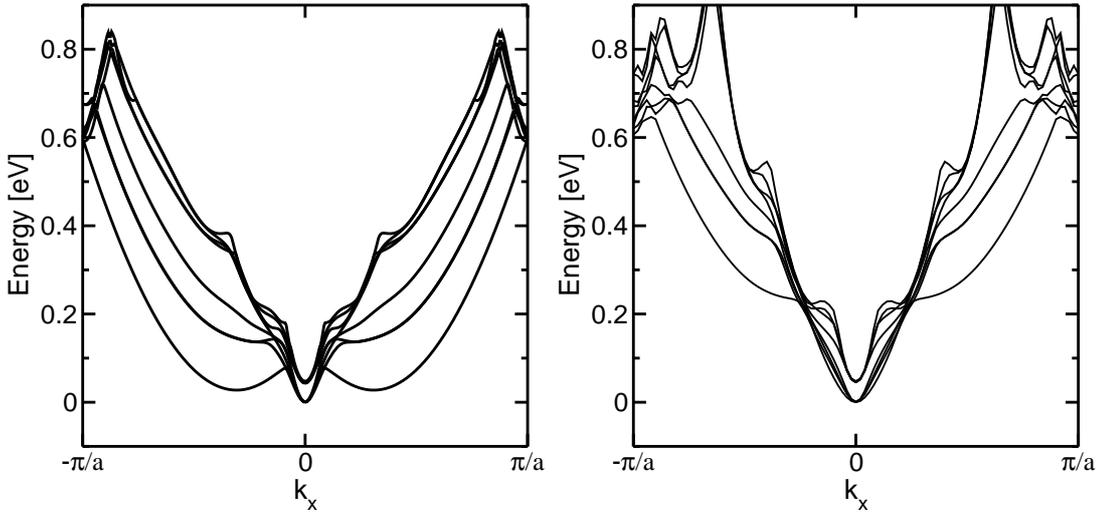


Figure 2.7: Electron bandstructures of a [100] Si nanowire with square section of  $5 \times 5 \text{ nm}^2$  with (left) no stress and with (right) tensile uniaxial stress of 3GPa.

The phonon-limited mobility of Si nanowires in the presence of elastic acoustic phonons and dispersion-less optical phonons is plotted in Fig. 2.6 for different

levels of uniaxial tensile stress for a [100] Si nanowire with square section of  $5 \times 5$  nm<sup>2</sup>. It is possible to appreciate the decrease of the mobility as a function of the inversion density in the device, as well as the mobility gain given by the effect of the uniaxial stress.

Such a mobility increase is mainly due to the modifications of the bandstructure properties induced by the mechanical stress, which are accurately taken into account by the atomistic EPM approach as shown in Fig. 2.7.

## 2.3 The $\mathbf{k} \cdot \mathbf{p}$ approximation method

The fundamental principle of the  $\mathbf{k} \cdot \mathbf{p}$  approximation is to calculate the energy levels at a band extreme, and use the eigenvectors and eigenvalues to expand the solution to a vicinity of the extreme in a perturbative manner. As such, it is difficult to come up with a set of parameters for a  $\mathbf{k} \cdot \mathbf{p}$  calculation that would produce satisfying result over the whole range of  $\mathbf{k}$  vectors.

### 2.3.1 Equations

Whereas in the EPM method we explicitly fit the potential functions to experimental results, the  $\mathbf{k} \cdot \mathbf{p}$  method involves directly fitting matrix elements. It should be noted that this leads to a higher amount of fitting parameters. Solutions of the Schrödinger equations are Bloch functions of the form

$$\Psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r}). \quad (2.23)$$

Expanding that form into the Schrödinger equation gives

$$e^{i\mathbf{k} \cdot \mathbf{r}} \left( \left( -\frac{1}{2} \nabla^2 + V(\mathbf{r}) \right) - i\mathbf{k} \cdot \nabla + \frac{1}{2} \mathbf{k}^2 \right) u_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} E_{n,\mathbf{k}} u_{n,\mathbf{k}}(\mathbf{r}). \quad (2.24)$$

Therefore the periodic component of the Bloch functions are solutions of

$$\left( -\frac{1}{2} \nabla^2 + V(\mathbf{r}) - i\mathbf{k} \cdot \nabla + \frac{1}{2} \mathbf{k}^2 \right) u_{n,\mathbf{k}}(\mathbf{r}) = E_{n,\mathbf{k}} u_{n,\mathbf{k}}(\mathbf{r}). \quad (2.25)$$

One should notice that the previous equation is identical to the Schrödinger equation for  $\mathbf{k} = 0$ . Therefore, its solutions at the  $\Gamma$  point are a complete set of the wavefunctions space, and the periodic components can be developed as

$$u_{n,\mathbf{k}}(\mathbf{r}) = \sum_i u_{i,\Gamma} c_i^n. \quad (2.26)$$

Injecting that development into (2.25) gives us the  $\mathbf{k} \cdot \mathbf{p}$  equation

$$\sum_i \left\{ \left( E_i^\Gamma + \frac{1}{2} \mathbf{k}^2 - E_{n,\mathbf{k}} \right) \cdot \delta_{i,i'} + \langle u_{i',\Gamma} | \mathbf{k} \cdot \mathbf{p} | u_{i,\Gamma} \rangle \right\} \cdot c_i^n = 0, \quad (2.27)$$

where  $\mathbf{p} = -i\nabla$  is the momentum operator and the  $E_i^\Gamma$  are the energies at the  $\Gamma$  point.

The quadratic approximation solutions to this equation are

$$E_{n,\mathbf{k}} \approx E_n^\Gamma + \frac{\mathbf{k}^2}{2} + \sum_{n' \neq n} \frac{|\langle u_{n,\Gamma} | \mathbf{k} \cdot \mathbf{p} | u_{n',\Gamma} \rangle|^2}{E_n^\Gamma - E_{n'}^\Gamma}. \quad (2.28)$$

We can therefore define an effective mass tensor  $m^*$  as

$$\frac{1}{m^*} = 1 + \frac{2}{\mathbf{k}^2} \sum_{n' \neq n} \frac{|\langle u_{n,\Gamma} | \mathbf{k} \cdot \mathbf{p} | u_{n',\Gamma} \rangle|^2}{E_n^\Gamma - E_{n'}^\Gamma}. \quad (2.29)$$

It should be noted that the specific case of a one-band  $\mathbf{k} \cdot \mathbf{p}$  results in the familiar effective mass approximation, as for small  $\mathbf{k}$ , Eq. (2.28) becomes

$$E_{n,\mathbf{k}} \approx E_n^\Gamma + \frac{\mathbf{k}^2}{2m^*}, \quad (2.30)$$

and in this case  $m^*$  is a scalar quantity.

### 2.3.2 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian

During the PhD program, in order to simulate III-V compounds, we have adopted the 8-band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian proposed in [10]. The explicit form is given by

$$\mathbf{H}_8 = \begin{pmatrix} C & 0 & P^+/\sqrt{6} & 0 & P^-/\sqrt{2} & -\sqrt{\frac{2}{3}}P_x & -P_x/\sqrt{3} & P^+/\sqrt{3} \\ C & -\sqrt{\frac{2}{3}}P_x & -P^+/\sqrt{2} & 0 & -P^-/\sqrt{6} & P^-/\sqrt{6} & P_z/\sqrt{3} & \\ & -P+Q & -S^* & R & 0 & \sqrt{3}2S & -\sqrt{2}Q & \\ & & -P-Q & 0 & R & -\sqrt{2}R & S/\sqrt{2} & \\ & & & -P-Q & S^* & S^*/\sqrt{2} & \sqrt{2}R^* & \\ & & & & -P+Q & \sqrt{2}Q & \sqrt{\frac{3}{2}}S^* & \\ & & & & & -P-\Delta_{SO} & 0 & \\ & & & & & & & -P-\Delta_{SO} \end{pmatrix}, \quad (2.31)$$

where

$$\begin{aligned}
C &= E_g + \frac{1}{2}\gamma_c(k_x^2 + k_y^2 + k_z^2) \\
P &= \frac{1}{2}\gamma_1(k_x^2 + k_y^2 + k_z^2) \\
Q &= \frac{1}{2}\gamma_2(k_x^2 + k_y^2 + k_z^2) \\
R &= \frac{1}{2}\sqrt{3}[\gamma_2(k_x^2 - k_y^2) - 2i\gamma_3k_xk_y] \\
S &= -\frac{1}{2}\gamma_32\sqrt{3}(k_x - ik_y)k_z \\
P^\pm &= P_0(k_x \pm ik_y) \\
P_z &= P_0k_z \\
P_0 &= \sqrt{E_p/2} \\
\gamma_c &= \frac{1}{m_c^*} - \frac{E_p}{3} \left( \frac{2}{E_g} + \frac{1}{E_g + \Delta_{SO}} \right).
\end{aligned} \tag{2.32}$$

Here, the modified Luttinger parameters  $\gamma_i$  are derived from the Luttinger parameters  $\gamma_i^L$  as

$$\begin{aligned}
\gamma_1 &= \gamma_1^L - \frac{E_p}{3E_g} \\
\gamma_2 &= \gamma_2^L - \frac{E_p}{6E_g} \\
\gamma_3 &= \gamma_3^L - \frac{E_p}{6E_g}.
\end{aligned} \tag{2.33}$$

It should be noticed that introducing the quantization in the  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian sometimes lead to the appearance of spurious solutions, which are caused by the non-ellipticity of the Hamiltonian. Spurious solutions can thus be eliminated by setting the value of parameter  $E_p$  such that the Hamiltonian operator is elliptic. We used the parameters from [11, 12].

### 2.3.3 Quantization

In order to simulate carrier transport in nanostructures, the effect of confinement must be introduced into the  $\mathbf{k} \cdot \mathbf{p}$  equation. This is done by replacing  $k_\alpha$  with  $-i\partial/\partial\alpha$ , which leads to an eigenvalue problem of a higher rank. In case of a nanowire in the  $x$  direction defined by a potential  $U(y, z)$  the 8-band  $\mathbf{k} \cdot \mathbf{p}$  equation becomes

$$\left[ \mathbf{H}_8 \left( k_x, -i\frac{\partial}{\partial y}, -i\frac{\partial}{\partial z} \right) + \mathbf{I}U(y, z) \right] \Psi_{n,k_x}(y, z) = E_{n,k_x} \Psi_{n,k_x}(y, z) \tag{2.34}$$

This equation is then discretized, and we end up with an operator of order  $8N_yN_z$ , that has to be diagonalized for each section of the wire. Since this is a problem too heavy to be solved in a short time, the problem can be projected on a sine functions basis to reduce its dimensionality [13].

### 2.3.4 Strain

Pikus and Bir [14] were the first to propose an analysis of strained semiconductors with the  $\mathbf{k} \cdot \mathbf{p}$  method. They gave a first order perturbation term to be added to Eq. (2.27) that is written as

$$W_{\mathbf{k},\mathbf{p}} = \sum_{i,j} -\mathbf{k}_i \epsilon_{i,j} \langle u_{n',\Gamma} | \mathbf{p}_j | u_{n,\Gamma} \rangle + \epsilon_{i,j} \langle u_{n',\Gamma} | \Xi_{i,j} | u_{n,\Gamma} \rangle, \quad (2.35)$$

where the  $i, j$  stand for  $x, y, z$  and  $\epsilon_{i,j}$  is the strain tensor. The first term in this equation covers the interaction between carrier momentum and strain. The second term is the deformation potential operator, and describes the change in potential and kinetic energy of the carriers due to the strain. The way to construct deformation potential matrix elements using group theory selection rules is described in [14]. According to [15], the resulting 8-band  $\mathbf{k}$ -dependent strain interaction operator can be explicitly written as

$$W_{\mathbf{k},\mathbf{p},8} = \begin{pmatrix} a_c e & 0 & -v^* & 0 & -\sqrt{3}v & \sqrt{2}u & u & -\sqrt{2}v^* \\ a_c e & \sqrt{2}u & -\sqrt{3}v^* & 0 & 0 & v & -\sqrt{2}v & -u \\ -p+q & -s^* & r & 0 & 0 & \sqrt{3/2}s & -\sqrt{2}q & \\ & -p-q & 0 & r & -\sqrt{2}r & s/\sqrt{2} & \sqrt{2}r^* & \\ & & -p-q & s^* & s^*/\sqrt{2} & \sqrt{2}r^* & \sqrt{3/2}s & \\ & & & -p+q & \sqrt{2}q & \sqrt{3/2}s & & \\ & & & & -a_v e & 0 & & \\ & & & & & & -a_v e & \end{pmatrix}, \quad (2.36)$$

where

$$\begin{aligned} e &= \epsilon_{c,xx} + \epsilon_{c,yy} + \epsilon_{c,zz} \\ p &= a_v e \\ q &= b[\epsilon_{c,zz} - (\epsilon_{c,xx} + \epsilon_{c,yy})/2] \\ u &= \frac{P_0}{\sqrt{3}}(\epsilon_{c,xz}k_x + \epsilon_{c,yz}k_y + \epsilon_{c,zz}k_z) \\ v &= \frac{P_0}{\sqrt{6}}[(\epsilon_{c,xx} - i\epsilon_{c,xy})k_x + (\epsilon_{c,xy} - i\epsilon_{c,yy})k_y + (\epsilon_{c,xz} - i\epsilon_{c,yz})k_z] \\ r &= \frac{\sqrt{3}}{2}b(\epsilon_{c,xx} - \epsilon_{c,yy}) - id\epsilon_{c,xy} \\ s &= -d(\epsilon_{c,xz} - i\epsilon_{c,yz}) \end{aligned}, \quad (2.37)$$

and the  $\epsilon_{c,ij}$  are the elements of the strain tensor in the crystal coordinate system. One should note that computing the strain effect does not increase the computational effort, as the order of the problem to be solved is unchanged.

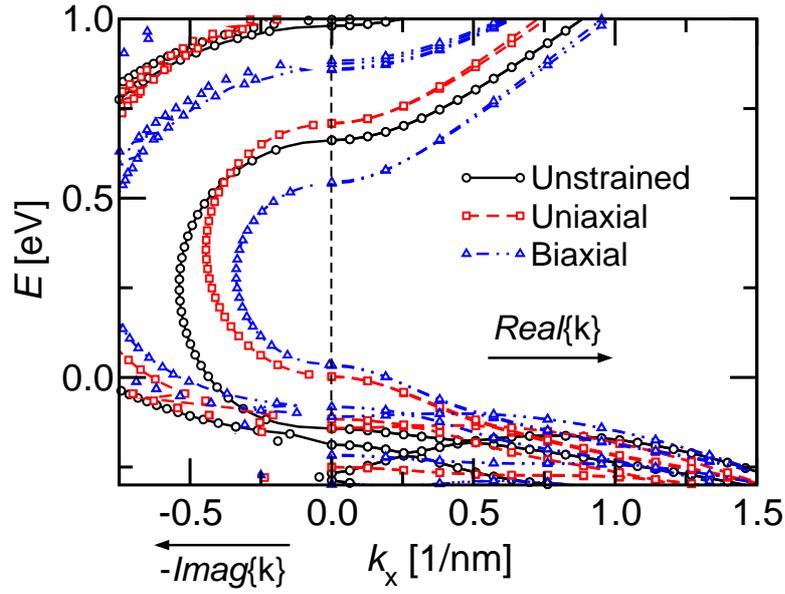


Figure 2.8: Energy dispersion in the gap for an square InAs NW with 5 nm width and for  $k_x$  along the [100] direction. Comparison between the unstrained, the compressive uniaxial stress ( $T_{xx} = 2$  GPa) and the tensile biaxial ( $T_{yy} = T_{zz} = 2$  GPa) stress conditions. The potential energy inside the wire is taken as zero. From Ref. [16].

The values used for deformation potentials are from [11]. Tab. 2.3 lists the potentials used in the simulations. Fig. 2.8 shows the effect of strain on the band structure in the  $\mathbf{k} \cdot \mathbf{p}$  method.

	$a_c$	$a_v$	$b$	$d$
InAs	-5.08	-1.0	-1.8	-3.6
GaSb	-7.5	-0.8	-2.0	-4.7

Table 2.3: Strain deformation potentials entering Eqs. 2.36 and 2.37 (units are eV).

# Chapter 3

## Full-quantum simulations

In this chapter we describe how to exploit the physical models and the Hamiltonians detailed in Chapter 2 to evaluate the transport properties of realistic nanodevices such as Si or III-V nanowire FETs within a full quantum approach.

The so-called non-equilibrium Green's functions (NEGF) approach is nowadays one of the most common methods to perform quantum transport calculations in nanodevices (see Sec. 3.1). It is based on the evaluation of the Keldysh-Green's functions, describing at once the dynamics and the kinetics of the carriers in the system. In our work it was used to compute electronic and transport properties of the devices under investigation and it was solved self-consistently with the 3-D Poisson equation as detailed in Sec. 3.2.

The main advantage of this approach is the capability to account for different physical problems at once. The equations of the NEGF method are presented in Sec. 3.3.

The main disadvantage of this approach is probably the numerical burden w. r. t. semi-classical methods arising from the CPU-time and memory requirements. In order to soften these drawbacks, we have implemented recursive numerical approaches able to speed up the calculations of the important quantities and we have computed approximated solutions in the so-called coupled mode space, allowing us to strongly reduce the computational effort. These methods are described in Sec. 3.4 and in Sec. 3.5, respectively. Finally, Sec. 3.6 presents the main formulas used to implement the phonon scattering in our calculations. The inclusion of this phenomenon is almost compulsory to realistically simulate electron devices at room temperature.

## 3.1 Context

The unending scaling of semiconductor device dimensions have accelerated the development of new simulation models and software tools which can correctly model quantum effects in extremely small devices. While quantum effects arising from lateral confinement can be accounted for also within a semiclassical approach based on the Boltzmann transport equation [17, 18], typically by introducing them only in the calculation of the subbands and the scattering rates, purely quantum-phase coherence phenomena such as electron tunneling, interferences, and localization effects demand a full-quantum approach.

Full-quantum simulation approaches [19, 20, 21, 22], on the other side, can overcome this limit, but are usually far more demanding from the computational standpoint with respect to the semi-classical ones. A quantum transport model can naturally include quantum mechanical effects along the transport direction like band-to-band-tunneling (BTBT) and interference effects, which are problematic to include in a semi-classical model. In the past few years, the increased performances of modern computers permitted to simulate realistic device structures by exploiting such methods. One possible way to tackle the full quantum transport problem in small devices is to directly solve the Schrödinger equation by using the finite difference or finite elements techniques and imposing open boundary conditions [23]. However, we will rather focus our attention on the Keldysh-Green's function formalism [24, 25, 26], because this method can handle also the problem of considering the inelastic electron-phonon interaction [25].

During the PhD program, we developed a 3-D Poisson-NEGF solver based on the atomistic empirical pseudo-potential and the 8-band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian (see Chapter 2) to simulate transport in direct-bandgap semiconductor materials. In particular, the next chapters will present the results of a simulation study which employs this model in order to simulate the electron mobility of Si nanowire FETs and the I-V curves of III-V tunneling-based nanowire transistors (Tunnel-FETs) and to systematically benchmark them against nanowire MOSFETs for low-power applications.

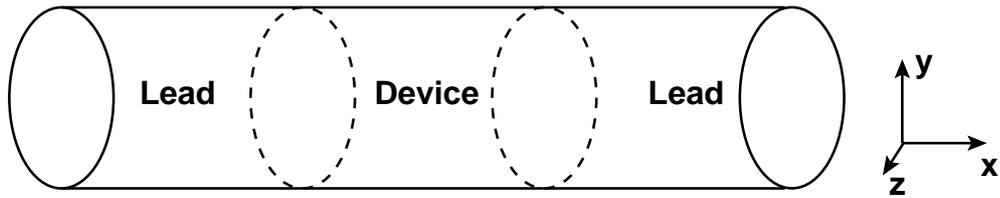


Figure 3.1: Schematic representation of a 3-D device connected to two external leads.

## 3.2 Self-consistent Poisson-Schrödinger solution

A sketch of the reference device structure for the numerical simulations is given in Fig. 3.1, representing a 3-D view and both longitudinal and transverse two-dimensional (2-D) sections. The complete domain is supposed to be formed by a device region with a finite extension and two contact regions infinite along the transport direction  $x$  (see Fig. 3.1). The domain has a finite width on the transverse plane. Due to the double confinement on the  $yz$ -plane shown in Fig. 3.1, only one direction of free propagation is left, thus defining the device as a quasi-1D system.

Although no restrictions are present on the geometry definition, in the following, a rectangular shaped domain and a rectangular discretization mesh will be supposed. The calculation of the fundamental physical observables of the system (i.e. charge and current distribution) is obtained, in steady-state conditions, by the solution of the time-independent Schrödinger equation, that for a single electron in a 3-D domain reads out

$$H\psi = (T + U)\psi = E\psi, \quad (3.1)$$

where  $\psi$  is the electron wave function,  $H = T + V$  is a generic Hamiltonian operator with a kinetic term  $T$ ,  $U$  the conduction(valence) band potential energy profile, and  $E$  the injection energy of the carrier. The presence of an external or induced magnetic field is neglected. Equation (3.1) has to be solved with proper boundary conditions depending on the chosen geometry of the device and depending on the model adopted for the source and drain contacts.

The electrostatic potential  $\phi$  is computed through the solution of the 3-D Poisson equation

$$-\nabla \cdot (\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})) = \rho(\mathbf{r}), \quad (3.2)$$

where  $\epsilon(\mathbf{r})$  is the position-dependent dielectric constant of the medium,  $\rho(\mathbf{r})$  is the charge density, and  $\mathbf{r} = (x, y, z)$  is the 3-D position vector. The charge density is given by the sum of the contributions due to free carriers, doping concentration and other fixed charge centers, as for example charges trapped in the oxide. The solution depends on the boundary condition (e.g. electrostatic potential at the electrodes). The electrostatic potential enters in (3.1) through the determination of the conduction band profile. On the other hand, the electron density depends on the solution of the Schrödinger equation through the square modulus of the electron wave function. Thus, a coupling between (3.1) and (3.2) is determined and an iterative self-consistent scheme has to be applied for the determination of the solution. Starting from an initial guess for the solution of the Poisson equation, the two equations are iteratively solved till a convergence criteria is verified. The calculation of all physical quantities of interest will be performed on the last obtained solution. The specific methods used for the calculation of the physical quantities of interest will be described in the following Sections with the introduction of the solution of the transport problem with open boundary condition in the space of transverse modes (or Coupled Mode Space approach) .

### 3.3 The Non Equilibrium Green's Function formalism

The Non-equilibrium Green's function formalism provides a microscopic theory for quantum transport. If in the case of ballistic transport the NEGF formalism is equivalent to the Landauer-Büttiker description of mesoscopic phenomena [19, 20], but the real power of NEGF is to provide a general approach for describing dissipative transport, combining quantum dynamics and a statistical description of the interactions [23, 27].

In the following, a steady state analysis is considered. In this framework, the Fourier transform of all the operators, and hence of all physical quantities, of interest allows for a description in terms of space and energy coordinates instead of time [27]. Further, a matrix notation with respect to the discrete spatial coordinates will be adopted.

The kinetic equations describing the non-equilibrium transport in the system,

according to the previous assumptions, can be expressed as [25]:

$$[(E + i\eta)I - H - \Sigma^r]G^r = I \quad (3.3)$$

$$[(E + i\eta)I - H - \Sigma^r]G^{<, >} = \Sigma^{<, >}G^{r\dagger}. \quad (3.4)$$

In Equations (3.3-3.4),  $G^r$  represents the retarded Green's function,  $H$  is the generic Hamiltonian of the isolated device (left hand side operator of (3.1)), while  $\Sigma^r$ , named retarded self-energy, represents the interactions of the device with external systems,  $E$  is the energy of the particle,  $\eta = 0^+$  is a positive infinitesimal quantity. Moreover, in Equation (3.4) the lesser-than and greater-than functions,  $G^<$  and  $G^>$ , describe the electron and hole density spectra, and the lesser-than and greater-than self-energy,  $\Sigma^<$  and  $\Sigma^>$ , are related to the in-scattering and out-scattering probabilities to the device from the interacting systems. This set of equations gives the information about the dynamical and statistical properties of the system, respectively. In presence of the electron-phonon interaction, the two equations are coupled through the definition of the retarded self-energy (see Sec. 3.6).

The retarded self-energy results from the sum of three terms

$$\Sigma^r = \Sigma_L^r + \Sigma_R^r + \Sigma_{\text{ph}}^r, \quad (3.5)$$

representing the interaction with the left contact (L), the right contact (R), and with the phonon bath (ph), respectively. The explicit dependence of each term on the energy has been omitted. In the following, the left and right contacts will be modeled as two reservoirs (source and drain) at thermal equilibrium with the chemical potential  $\mu_S$  and  $\mu_D$ , respectively, infinite along the transport direction, replicating uniformly the same confining potential  $U(y, z)$ , and reflectionless. This model is used to mimic the inclusion of ideally ohmic contacts. From a general point of view, the contact self-energy functions allow for the inclusion of the effect of inward and outward waves propagating through the device from an infinitely extended region maintaining a representation on a finite discrete space. Their definition requires the evaluation of the retarded Green's function associated to the semi-infinite portion of contacts outside the device and can be done analytically or numerically depending on the used Hamiltonian representation.

The inclusion of the electron-phonon interaction and additional details for the evaluation of the phonon self-energy will be presented in Sec. 3.6. The inclusion

of this collision mechanism will be approached within a low-order approximation (self-consistent Born approximation) [25, 28, 29], able, in any case, to capture all the fundamental properties of the electron-phonon interaction in a full-quantum description.

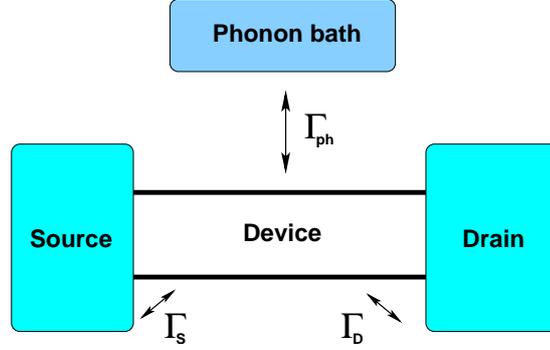


Figure 3.2: Sketch of the interaction model between the device and the external systems (contacts and phonon bath). The frequency of the interactions is described by the  $\Gamma$  functions.

The density spectral function  $A$  describing the available electronic states of the system (filled and not) is given by the relationship

$$i(G^> - G^<) = i(G^r - G^{r\dagger}) = A. \quad (3.6)$$

The imposition of an equilibrium condition for the carrier distribution in the contacts is fulfilled through the definition of the lesser-than and greater-than self-energy functions. They are expressed by the following relations

$$\Im[\Sigma_{L,R}^<(E)] = \Gamma_{L,R}(E)f(E - \mu_{L,R}) \quad (3.7)$$

$$\Im[\Sigma_{L,R}^>(E)] = -\Gamma_{L,R}(E)[1 - f(E - \mu_{L,R})], \quad (3.8)$$

where

$$\Gamma_{L,R}(E) = -\Im[\Sigma_{L,R}^>(E) - \Sigma_{L,R}^<(E)] = -\Im[\Sigma_{L,R}^>(E) - \Sigma_{L,R}^{r\dagger}(E)]. \quad (3.9)$$

The physical meaning of  $\Gamma$  permits a better understanding of the self-energy description of scattering mechanisms. The quantity  $1/\tau = -i\Gamma/\hbar$  represents a scattering probability describing the decay-rate of a carrier from an external

system to the device and vice-versa. In the case of an ideal ballistic conductor, the in- and out- scattering from the contacts remains the only interaction mechanism affecting the propagation of the carrier along the device. This physical description holds for both left and right contacts and for self-energy functions representing the electron-phonon interaction. In the latter case, as described in Fig. 3.2, the scattering rate gives the probability of interaction of the electron in the device with the outer system of phonons, here approximated as an unperturbed phonon bath. Interactions with the system are possible at every point of the device along the transport direction. In this general picture,  $1/\tau$  describes the total phase relaxation rate [30, 31].

The energy-resolved electron and hole density are determined as the expectation value of the respective field operator and they can be expressed as

$$n(\mathbf{r}; E) = \frac{1}{2\pi} \Im[G^<(\mathbf{r}, \mathbf{r}; E)] \quad (3.10)$$

$$p(\mathbf{r}; E) = -\frac{1}{2\pi} \Im[G^>(\mathbf{r}, \mathbf{r}; E)]. \quad (3.11)$$

Similarly, the current density distribution is defined as

$$\mathbf{J}(\mathbf{r}; E) = \frac{e}{\hbar} \lim_{\mathbf{r} \rightarrow \mathbf{r}'} \Re\{[H(\mathbf{r}) - H(\mathbf{r}')]G^<(\mathbf{r}, \mathbf{r}'; E)\}, \quad (3.12)$$

while the current density at the contact  $\alpha = S, D$  is given by the formula

$$J_\alpha(E) = \frac{e}{\hbar} \text{Tr}\{\Gamma_\alpha \Im[G^<(\mathbf{r}_\alpha, \mathbf{r}_\alpha; E) + 2G^r(\mathbf{r}_\alpha, \mathbf{r}_\alpha; E)f(E - \mu_\alpha)]\}, \quad (3.13)$$

where  $f(E - \mu_\alpha) = 1/(1 + \exp[(E - \mu_\alpha)/K_B T])$  is the Fermi-Dirac distribution function at the contact  $\alpha = S, D$ .

As a consequence of the presented relations, in the case of an ideal ballistic conductor it is more clear the division between the dynamical and the statistical description of the systems. Introducing the formulation of the contacts lesser-than Green's functions in and making use of (3.6), the electron density at a fixed energy and device position is given by

$$2\pi n(\mathbf{r}; E) = A_S(\mathbf{r}; E)f(E - \mu_S) + A_D(\mathbf{r}; E)f(E - \mu_D) \quad (3.14)$$

where  $A_S$  and  $A_D$  are the density spectral function due to the injection from the source and drain, respectively. The contact equilibrium statistics give the appropriate filling of the density of states. The knowledge of the retarded Green's function is required in the calculation of independently of the knowledge of its occupancy, thus uncoupling the kinetic equations.

### 3.3.1 The Dyson equation

This section presents the Dyson equation and illustrates simple 1-D examples based on effective mass approximation (EMA) Hamiltonians. Indeed, we developed a NEGF solver based on more complex Hamiltonians such as the 8-band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian introduced in Chapter 2, but, once we write the Hamiltonian as a tri-diagonal block matrix, all the transformations and the recursive algorithms are very general and remain the same also for different Hamiltonians.

Consider an unperturbed system whose Hamiltonian matrix is  $H_0$ . The corresponding (unperturbed) retarded GF  $g$  is defined as

$$g = (EI - H_0)^{-1}, \quad (3.15)$$

where we have implicitly assumed that the small imaginary term  $i0^+$  introduced in Eq.(3.3) is added to the energy  $E$ .

Assume that we perturb the system with a small Hamiltonian  $V$ . We can write the new retarded GF  $G$  as

$$G = (EI - H)^{-1}, \quad (3.16)$$

where  $H$  is the perturbed Hamiltonian defined as  $H = H_0 + V$ .

Using Equation (3.15) we can rewrite (3.16) as

$$G = (EI - H_0 - V)^{-1} = (g^{-1} - V)^{-1}, \quad (3.17)$$

which, with straight-forward matrix algebra, gives the Dyson equation for the retarded GF [25]

$$G = g + gVG = g + GVg. \quad (3.18)$$

Eq. (3.18) can be written explicitly in a matrix notation as

$$G(m, n) = g(m, n) + \sum_{p, q} g(m, p)V(p, q)G(q, n) = g(m, n) + \sum_{p, q} G(m, p)V(p, q)g(q, n), \quad (3.19)$$

which relates the perturbed GF elements  $G(m, n)$  to the unperturbed ones  $g(m, n)$  via the elements  $V(p, q)$  of the perturbation Hamiltonian. The indexes  $m, n, p$  and  $q$  run over the  $N_x$  discretization points.

### 3.3.2 A simple example

Consider now a semi-infinite chain of sites as in Fig. 3.3 and characterized by an unperturbed GF  $g$  and by a single energy level  $\epsilon_0$  for each site. The binding energy (hopping) between two adjacent sites in the EMA is equal to  $t = \hbar^2/2ma^2$ . Assume to perturb the system by adding a further site at position 0 as in Fig 3.3. This simplified example mimics the coupling of a semi-infinite contact with a real device [23]. Using Eq. (3.19) one can write the following system of equations

$$G_{0,0} = g_{0,0} + g_{0,0}V_{0,-1}G_{-1,0} \quad (3.20)$$

$$G_{-1,0} = g_{-1,-1}V_{-1,0}G_{0,0}, \quad (3.21)$$

which relates the unperturbed ( $g_{0,0}$  and  $g_{-1,-1}$ ) to the perturbed ( $G_{0,0}$  and  $G_{-1,0}$ ) GF components.

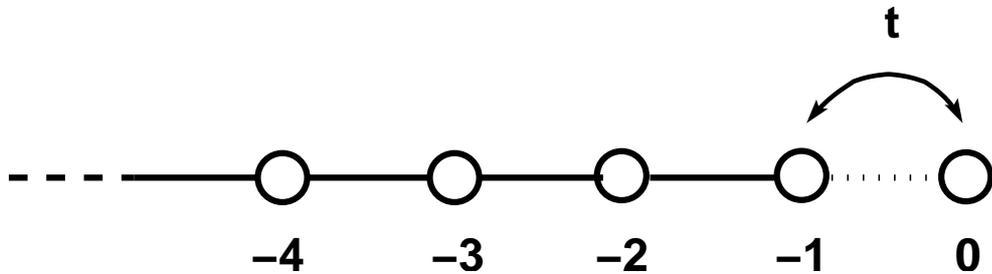


Figure 3.3: Sketch of the 1D semi-infinite atomic chain. Each atom is connected to the first neighbors with a coupling strength  $t$ .

All the other terms in the summations of Eqs. (3.21-3.21) vanish ( $g_{-1,0}$  is zero because in the unperturbed case we do not have the site labeled as 0). Now we can assume that the final system is identical to the initial one, because in both cases we deal with a semi-infinite lead of sites, and thus write

$$g_{-1,-1} = G_{0,0}. \quad (3.22)$$

In this case we are conjecturing that the last site in the unperturbed case is equal to the last one in the perturbed case. Considering a constant hopping term  $t$  between sites and writing the unperturbed GF as  $g_{0,0} = (E - \epsilon_0)^{-1}$  we can combine Eqs. (3.21) and (3.21) in order to obtain the following closed relationship for the unknown element  $G_{0,0}$  of the perturbed Hamiltonian

$$t^2 G_{0,0}^2 - (E - \epsilon_0)G_{0,0} + 1 = 0. \quad (3.23)$$

The term  $G_{0,0}$  is the surface GF which can be used to determine the contact retarded self energy. Solving Equation (3.23) and taking only the negative solution one can obtain

$$G_{0,0} = \frac{1}{t}(X - i\sqrt{1 - X^2}), \quad (3.24)$$

where  $X = (E - \epsilon_0)/2t$ .

### 3.3.3 A dot attached to a lead

Let's consider an isolated atom with energy  $\epsilon_0$  and coupling parameter  $t$  with an infinite reservoir on the left of the atom. In this case the Green's function of the isolated dot reads

$$g_{0,0}(E) = \frac{1}{E - \epsilon_0}. \quad (3.25)$$

By using the Dyson equation (3.19), we obtain

$$\begin{aligned} G_{0,0} &= g_{0,0} + g_{0,0}t_{0,L}G_{L,0} \\ G_{L,0} &= g_{L,L}t_{L,0}G_{0,0}, \end{aligned}$$

where  $g_{L,L}$  is the Green's function of the lead and is given by Eq. (3.24). This implies that the Green's function of the whole system reads

$$G_{0,0}(E) = \frac{1}{E - \epsilon_0 - t_{0,L}g_{L,L}t_{L,0}} = \frac{1}{E - \epsilon_0 - \Sigma_L}, \quad (3.26)$$

where

$$\Sigma_L = t_{0,L}g_{L,L}t_{L,0} \quad (3.27)$$

is the lead self-energy and accounts for the effect of the lead on the dot. In the wide band limit approximation valid in the linear regime we have that the retarded Green's function of the atom coupled to the external lead is

$$G_{0,0}(E) = \frac{1}{E - \epsilon_0 + i\Gamma_L/2}, \quad (3.28)$$

where  $\Gamma_L = -2\text{Im}\{\Sigma_L\}$  represents the coupling strength.

During the PhD work, in order to compute the contact self-energy of devices described by complex Hamiltonians such as the 8-band  $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian, we have adopted the numerical algorithm proposed by Sancho and Rubio in [32], which is similarly based on an iterative scheme derived from the Dyson equation (3.18).

### 3.4 The recursive Green's function algorithms

The direct inversion of the r.h.s of Eq. (3.3) and determine  $G$  can be almost impossible because this matrix has  $(N_x N_y N_z)^2$  elements and, moreover, it has to be done for a large number of energies  $E$  defined on a dense energy grid. There is an alternative way to determine exactly  $G$  or sub-matrices of  $G$  using the Dyson equation [31]. It is important to remember that, in order to calculate carrier concentration and current, we need only a small part of the diagonal elements of  $G$  (see Eqs. (3.10-3.11) and (3.13)), and so it is not necessary to know all the elements and to invert the entire matrix in the r.h.s of Eq. (3.3). Here we should point out that the presented recursive algorithms can be used only if the matrix  $H + \Sigma$  is a tri-diagonal block matrix. If this is not the case, the recursive algorithm cannot be used and Eq. (3.3) has to be inverted directly. Fig. 3.4 sketches the device sections along the transport direction  $x$ .

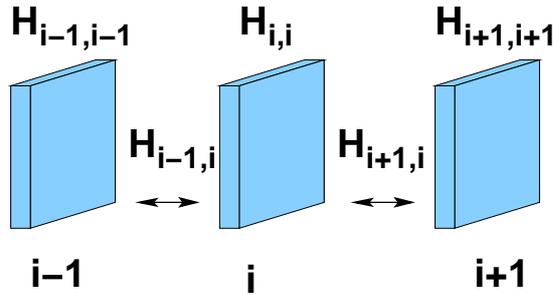


Figure 3.4: Sketch of the device sections along the transport direction. The Hamiltonian of the  $i$ -th section is  $H_{i,i}$ , whereas the hopping terms are  $H_{i,i\pm 1}$ .

The unperturbed system refers to an isolated slice, while the perturbation consists in adding one slice at a time on one side of the device, starting from one contact. In particular, using Eq. (3.18), we can write the relation between the unperturbed GF of the first slice  $g_{1,1}$  and the perturbed  $g_{1,1}^{\text{Left}}$ , which is not yet the final GF, because it is the GF of the slice connected only on the left side. In particular we can write

$$g_{1,1}^{\text{Left}} = [g_{1,1}^{-1} - \Sigma_{1,1}]^{-1} \quad (3.29)$$

where the superscript “Left” means coming from the left side, and the self-energy accounts for both the contacts and the phonon scattering. For a generic slice (i.e.

$i = 2, \dots, N_x$ ) the latter equation becomes

$$g_{i,i}^{\text{Left}} = [g_{i,i}^{-1} - \Sigma_{i,i} - H_{i,i-1} g_{i-1,i-1}^{\text{Left}} H_{i-1,i}]^{-1}. \quad (3.30)$$

Once we arrive at the right-end side of the device we know the self-energy of the drain  $\Sigma_D$ , hence we can calculate the perturbed GF  $G$ . Hence, the last equation of the chain will be

$$G_{N_x, N_x} = [g_{N_x, N_x}^{-1} - \Sigma_{N_x, N_x} - H_{N_x, N_x-1} g_{N_x-1, N_x-1}^{\text{Left}} H_{N_x-1, N_x}]^{-1}. \quad (3.31)$$

Eq. (3.31) gives us the final out-of-equilibrium GF at the  $N_x$ -th slice. We can identify in it the fluxes coming from the left and the right side.

Finally, starting from the  $G_{N_x, N_x}$  calculated in Eq. (3.31), one can calculate the GF corresponding to the slices from  $i = N_x - 1$  to  $i = 1$  by using

$$G_{i,i} = g_{i,i}^{\text{Left}} + g_{i,i}^{\text{Left}} H_{i,i+1} G_{i+1,i+1} H_{i+1,i}^\dagger g_{i,i}^{\text{Left}}. \quad (3.32)$$

We stress once again that the recursive algorithm to calculate the retarded GF is very general and has the same expressions for any form of Hamiltonian. In particular, we employed such algorithm to calculate the 3-D GF for the 8-band  $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian introduced in Chapter 2. A direct application of this procedure with the Hamiltonians discretized directly in real space is still too demanding from the computational point of view, thus we employed the mode-space transformation (see Sec. 3.5), in order to reduce the dimension of the matrices involved in the calculation.

### 3.4.1 Recursive scheme for the lesser-than Green's function

The same kind of numerical strategy used to compute specific elements of the retarded Green's function can be used to compute the same elements of the lesser-than Green's function  $G^<$  [31]. Such an algorithm is based on the Dyson equation for the lesser(grater)-than Green's function which is given by the formula

$$G^< = g^< + g \Sigma G^< + g^< \Sigma^\dagger G^\dagger + g \Sigma^< G^\dagger \quad (3.33)$$

or

$$G^< = g^< + G \Sigma g^< + G^< \Sigma^\dagger g^\dagger + G \Sigma^< g^\dagger, \quad (3.34)$$

where  $G^<$  and  $g^<$  are the perturbed and unperturbed lesser-than Green's functions.

Since the derivation of this recursive algorithm is quite lengthy, we give here the final set of formulae. The first step is, as before, to attach the left lead to the contact according to the equation

$$g_{1,1}^{<,Left} = g_{1,1}^{Left} \Sigma_{1,1}^{<} g_{1,1}^{\dagger,Left} \quad (3.35)$$

where  $g^{<,Left}$  is the lesser-than GF connected to just the left side.

The second step is to exploit the previous solution and to recursively compute  $g_{i,i}^{<,Left}$  for each slice  $i = 2, \dots, N_x - 1$  along the transport direction. This is accomplished by means of the formula

$$g_{i,i}^{<,Left} = g_{i,i}^{Left} [H_{i,i-1} g_{i-1,i-1}^{<,Left} H_{i-1,i}^{\dagger} + \Sigma_{i,i}^{<}] g_{i,i}^{\dagger,Left}. \quad (3.36)$$

Finally, the third step consists in attaching the right lead and in recursively compute the final lesser-than green's function  $G_{i,i}^{<}$  for all the slices from  $N_x$  to 1 according to

$$G_{i,i}^{<} = g_{i,i}^{<,Left} + g_{i,i}^{Left} H_{i,i+1} G_{i+1,i+1}^{<} H_{i+1,i}^{\dagger} g_{i,i}^{\dagger,Left} + G_{i,i+1} H_{i+1,i} g_{i,i}^{<,Left} + g_{i,i}^{<,Left} H_{i,i+1}^{\dagger} G_{i+1,i}^{\dagger}. \quad (3.37)$$

The formula to compute the element  $G_{i+1,i}^{<}$ , which is useful to compute the spatial distribution of the current density spectrum reads out

$$G_{i+1,i}^{<} = G_{i+1,i+1} H_{i+1,i} g_{i,i}^{<,Left} + G_{i+1,i+1}^{<} H_{i+1,i}^{\dagger} g_{i,i}^{\dagger,Left}. \quad (3.38)$$

### 3.5 The coupled mode-space approach

If we discretized a 3-D Hamiltonian such as the 8-band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian of Eq. (2.31) directly in real-space by using a finite difference method for each point  $\mathbf{r} = (x_i, y_j, z_k)$ , we obtain a block tri-diagonal matrix with sub-matrices  $H(x_i, x_j)$ , where  $i$  identifies the  $i$ -th section. Here,  $x$  is the transport direction and carriers are confined in the  $y - z$  plane. Indeed, the tri-diagonal form for the Hamiltonian is necessary in order to exploit the recursive GF method based on the Dyson equation to calculate the retarded and the lesser-than GF. The complete real-space Hamiltonian reads

$$H_{\text{RS}} = \begin{pmatrix} H(x_1, x_1) & H(x_1, x_2) & & & \\ H(x_2, x_1) & H(x_2, x_2) & H(x_2, x_3) & & \\ & H(x_3, x_2) & H(x_3, x_3) & & \\ & & & \ddots & \\ & & & & H(x_{N_x}, x_{N_x-1}) & H(x_{N_x}, x_{N_x-1}) \end{pmatrix} \quad (3.39)$$

Each  $H(x_i, x_j)$  has rank  $N_b N_y N_z$ , where  $N_b$  is the number of bands of the k.p Hamiltonian and  $N_y, N_z$  are the number of discretization points along the  $y$  and  $z$  directions, respectively. If we denote with  $N_x$  the number of sections in the transport direction, then the total Hamiltonian has rank  $N_b N_x N_y N_z$ .

The manipulation of matrices of such dimensions is an almost intractable problem from a computational standpoint. We thus resorted to the theoretically equivalent Coupled Mode- Space formulation [33, 34, 35, 13, 16]. By means of this approach, the rank of the diagonal and off-diagonal blocks of the Hamiltonian can be reduced by means of the unitary transformation

$$H_{\text{MS}}(x_i, x_j) = U(x_i)^\dagger H_{\text{RS}}(x_i, x_j) U(x_j), \quad (3.40)$$

where

$$[H(x_i, x_i) + H(x_i, x_{i+1}) + H(x_i, x_{i+1})^\dagger] \chi_m(x_i) = E_m(x_i) \chi_m(x_i) \quad (3.41)$$

is the rectangular unitary matrix composed by the  $M$  eigenvectors of the 2-D Schrödinger problem at the  $i$ -th device section expressed by the equation

$$U(x_i) = [\chi_1(x_i) \chi_2(x_i) \cdots \chi_M(x_i)]. \quad (3.42)$$

Eq. (3.40) stems from the transformation of the complete 3-D Hamiltonian of Eq. (3.39). Being  $H_{\text{RS}}$  a tri-diagonal block-matrix, also the mode-space Hamiltonian  $H_{\text{MS}}$  preserves the tri-diagonal form and we can express the blocks as in Eq. (3.39).

The eigenvalues  $E_m(x_i)$  and wave-functions  $\chi_m(x_i)$  are referred to as the modes of the section  $x_i$  and the number  $M$  of the modes retained in each section sets the rank of the submatrices  $H_{\text{MS}}(x_i, x_j)$  of the mode space Hamiltonian. In the mode-space, Eqs. (3.3-3.4) are re-written as

$$[(E + i\eta)I - H_{\text{MS}} - \Sigma_{\text{MS}}^r] G_{\text{MS}}^r = I \quad (3.43)$$

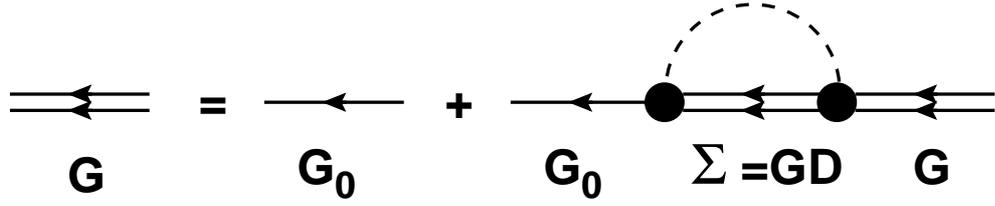


Figure 3.5: Feynman diagram representation of the Dyson equation of the electron-phonon interaction within the self-consistent Born approximation.

$$[(E + i\eta)I - H_{\text{MS}} - \sigma_{\text{MS}}^r]G_{\text{MS}}^{<, >} = \Sigma_{\text{MS}}^{<, >}G_{\text{MS}}^{r\dagger}, \quad (3.44)$$

where  $\Sigma_{\text{MS}}^{r, <, >}$  are the mode-space self-energies, which have been transformed as in Eq. (3.40).

Importantly, the real-space Green's function can be expressed in terms of the mode space Green's function elements  $G_{m,n}^{<}$  according to the following expansion

$$G_{\text{RS}}^{<}(x_i, x_j) = \sum_{m,n} G_{m,n}^{<}(x_i, x_j) \chi_m(x_i)^* \chi_n(x_j). \quad (3.45)$$

Eq. (3.45) will be used to re-build the electron concentration in the real space needed as the input of the Poisson equation in the PS self-consistent loop.

### 3.6 The electron-phonon interaction

The electron-phonon interaction was included within a perturbative model within the self-consistent Born approximation. The calculation of the phonon self-energy can be derived by a diagrammatic representation of the Dyson equation describing the second order contribution of the phonon interaction to the lesser-than Green's function [36]. Here,  $G$  represents the final perturbed Green's function (lesser-than Green's function in the following), being  $G_0$  the unperturbed one. The interaction between the electron and phonon system is represented by a proper self-energy function, confined between a starting and a final state (the circular vertex of the diagram).

Accordingly to the model, the phonon system is considered unperturbed by the interactions with the electron gas, hence the self-energy reported in Fig. 3.5 can be translated into

$$\Sigma_{\text{ph}}^{>, <} = D^{>, <}G^{>, <}, \quad (3.46)$$

where a real-space representation is supposed and the energy dependence has been omitted.  $D^{>,<}$  represent the lesser-than and greater-than Green's functions of the unperturbed phonon bath, respectively. Their analytical formulations are calculated from the free-phonons Hamiltonian in the interaction picture, that in the harmonic approximation for the lattice vibrations reads out

$$\phi_I(\mathbf{r}; t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} M_{\mathbf{q}} (a_{\mathbf{q}} e^{-i\omega_{\mathbf{q}} t} + a_{-\mathbf{q}}^{\dagger} e^{i\omega_{\mathbf{q}} t}) e^{i\mathbf{q} \cdot \mathbf{r}}, \quad (3.47)$$

where  $V$  is the volume of the sample,  $\mathbf{q}$  is the wave vector of the phonons,  $\omega_{\mathbf{q}}$  is the corresponding angular frequency,  $M_{\mathbf{q}}$  is the electron-phonon matrix element, and  $a_{-\mathbf{q}}^{\dagger}$  and  $a_{\mathbf{q}}$  are the phonon creation and annihilation operators. Fourier-transforming the time-dependent operators, the self-energy functions at a given energy  $E$  are thus given by

$$\Sigma_{\text{ph}}^{<}(\mathbf{r}, \mathbf{r}'; E) = \int \frac{d\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} |M_{\mathbf{q}}|^2 (N_{\mathbf{q}} + \frac{1}{2} \pm \frac{1}{2}) G^{<}(\mathbf{r}, \mathbf{r}'; E \pm \hbar\omega_{\mathbf{q}}) \quad (3.48)$$

and

$$\Sigma_{\text{ph}}^{>}(\mathbf{r}, \mathbf{r}'; E) = \int \frac{d\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} |M_{\mathbf{q}}|^2 (N_{\mathbf{q}} + \frac{1}{2} \pm \frac{1}{2}) G^{>}(\mathbf{r}, \mathbf{r}'; E \mp \hbar\omega_{\mathbf{q}}), \quad (3.49)$$

where  $N_{\mathbf{q}}$  is the average phonon density at the energy  $\hbar\omega_{\mathbf{q}}$  given by the Bose-Einstein distribution function

$$N_{\mathbf{q}} = [\exp(\hbar\omega_{\mathbf{q}}/K_B T) - 1]^{-1}. \quad (3.50)$$

The solution of the kinetic equations requires also the knowledge of the retarded self-energy, that can be calculated from the general relation

$$\Sigma^r(E) = P \int \frac{dE'}{2\pi} \frac{\Gamma(E')}{E - E'} - i \frac{\Gamma(E)}{2}, \quad (3.51)$$

where  $P$  is a principal value integral on the complex plane, and  $\Gamma(E)$  is defined as

$$\Gamma(E) = -\Im[\Sigma^{>}(E) - \Sigma^{<}(E)] \quad (3.52)$$

for every point inside the device.

The real part of  $\Sigma^r(E)$ , represented by the first term of the right side of (3.51), is a non-hermitian energy contribution giving a renormalization shift of

the particle energy levels. The second term is associated to the the scattering rate due to the electron-phonon interaction, as stated previously. The evaluation of Eq. (3.51) can be done analytically only for the case of elastic scattering while requiring a numerical evaluation of the principal value integral in the case of inelastic scattering process. This calculation can be computationally expensive due the necessity of the simultaneous knowledge of the Green's functions for every energy. Its contribution is usually omitted. The validity of this assumption has been numerically investigated in [37]. In the following only the information connected to the imaginary part of the retarded self-energy will be retained within the approximation

$$\Sigma^r(E) \approx -i \frac{\Gamma(E)}{2}. \quad (3.53)$$

It is worth highlighting that through (3.51) or, equivalently, (3.53), a dependence of  $\Sigma^r(E)$  on  $G < (E)$  is introduced. This implies, that in presence of the electron-phonon interaction the Equations (3.3) and (3.4) are coupled through a non linear relationship. As previously stated, in the ballistic case the two equations separately describe the dynamics and the statistical properties of the system. When the phonon scattering is included, hence, a self-consistent iterative solution of the two equations (3.3) and (3.4) with the phonon self-energy functions is required, in addition to the self-consistent loop with the Poisson equation. It is important to note that in the self-consistent Born approximation the conservation of the current flux is guaranteed along the entire device, verifying the fundamental fulfillment of the continuity equation.

### 3.6.1 The self-consistent Born approximation

The self-consistent Born approximation is surely the most widely used approach to include the electron-phonon interaction in the NEGF formalism. Approximations to the presented method or alternative approaches are in any case possible [38, 39, 40]. Comparative studies between different approaches are present in literature. In the following subsections the discrete form of Equations (3.48) and (3.49) will be presented in the inelastic and elastic case for both semiconducting nanowires deriving the mode-space formulation from the real-space representation.

In the case of acoustic intra-valley, phonon scattering is treated in the elastic

approximation, considering  $E \pm \hbar\omega_q \approx E$ . Within the scalar assumption for the acoustic deformation potential  $\Xi_{\text{ac}}$  the electron-phonon matrix element reads out

$$|M_{\mathbf{q}}|^2 = \frac{\hbar\Xi_{\text{ac}}^2 q}{2\rho u_l}, \quad (3.54)$$

where  $\rho$  is the material density,  $u_l$  is the sound velocity and a linear dispersion relation is assumed for the longitudinal acoustic mode. Due to the elastic approximation  $N_q \approx N_q + 1 \approx k_B T / (\hbar q u_l)$  and the lesser-than Green's functions involved in the integral become independent of the phonon wave vector. Consequently, the real-space self-energy functions are expressed as

$$\Sigma_{\text{ph}}^<(\mathbf{r}, \mathbf{r}'; E) = \frac{K_B T \Xi_{\text{ac}}^2}{2\rho u_l^2} G^<(\mathbf{r}, \mathbf{r}'; E) \delta(\mathbf{r} - \mathbf{r}'), \quad (3.55)$$

where the Dirac delta is derived from the remaining integrand

$$\delta(\mathbf{r} - \mathbf{r}') = \int \frac{d\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}. \quad (3.56)$$

In the case of optical phonons, approximating the momentum transfer for the  $j$ -th transition type as the momentum difference between the minimum energy states of the two valleys,  $\mathbf{K}_j$ , with a fixed energy transfer  $\hbar\omega_j$ , the deformation potential is independent of  $q$  and reads

$$|M_{\mathbf{q}}|^2 = \frac{\hbar D_t^2 \mathbf{K}_j^2}{2\rho\omega_j}. \quad (3.57)$$

Inserting (3.57) in (3.48) or (3.49), the  $j$ -th phonon self-energy reads

$$\Sigma_{\text{ph}}^<(\mathbf{r}, \mathbf{r}'; E) = \frac{\hbar D_t^2 \mathbf{K}_j^2}{2\rho\omega_j} G^<(\mathbf{r}, \mathbf{r}'; E \pm \hbar\omega_j) \left(N_q + \frac{1}{2} \pm \frac{1}{2}\right) \delta(\mathbf{r} - \mathbf{r}') \quad (3.58)$$

and

$$\Sigma_{\text{ph}}^>(\mathbf{r}, \mathbf{r}'; E) = \frac{\hbar D_t^2 \mathbf{K}_j^2}{2\rho\omega_j} G^>(\mathbf{r}, \mathbf{r}'; E \pm \hbar\omega_j) \left(N_q + \frac{1}{2} \mp \frac{1}{2}\right) \delta(\mathbf{r} - \mathbf{r}'). \quad (3.59)$$

The above relations have to be transformed from the real-space to the coupled mode-space representation. The same procedure presented in the previous section with the projection over the basis set of slice eigenfunctions is used. Dropping

out the superscripts and subscripts of the lesser-than and greater-than phonon self-energy, the  $n, m$  element reads out

$$\Sigma_{m,n}(x_i, x_j) = \int dydz \int dy'dz' \chi_n(x_i)^* \Sigma_{\text{RS}}(x_i, x_j) \chi_n(x_j). \quad (3.60)$$

For the case of intra-valley acoustic phonons, inserting (3.55) in the above equation and making use of the expansion (3.45), we obtain the mode-space self-energy [29, 16]

$$\Sigma_{n,n}^<(E) = \frac{K_B T \Xi_{\text{ac}}^2}{2\rho u_l^2} \sum_m G_{m,m}^<(E) I_{m,n}. \quad (3.61)$$

Here, all the contributions arising from the cross terms are neglected, reducing the self-energy formulation a diagonal matrix. This assumption has been validated numerically, noting that the overlap integral of multiple eigenfunctions associated to the off-diagonal terms of the Green's function give a negligible contribution compared with the terms accounted for. In Eq. (3.61) the term  $I_{m,n}$  represents the usual form factor

$$I_{m,n} = \int dydz |\chi_m|^2 |\chi_n|^2. \quad (3.62)$$

In the case of inter-valley scattering a similar result is obtained

$$\Sigma_{n,n}^<(E) = \frac{\hbar(D_t \mathbf{K}_j)^2}{2\rho\omega_j} \sum_m G_{m,m}^<(E \pm \hbar\omega_j) (N_q + \frac{1}{2} \pm \frac{1}{2}) I_{m,n} \quad (3.63)$$

and

$$\Sigma_{n,n}^>(E) = \frac{\hbar(D_t \mathbf{K}_j)^2}{2\rho\omega_j} \sum_m G_{m,m}^>(E \pm \hbar\omega_j) (N_q + \frac{1}{2} \mp \frac{1}{2}) I_{m,n}. \quad (3.64)$$

Acoustic and optical phonon parameters for sSi MOSFETs were taken from [41]. Moreover, a polar optical phonon (POP) scattering was also included in InAs device simulations due to its relevance in III-V compounds. However, although this is a non-local mechanism, it was accounted for by means of a local self-energy similar to Eqs. (3.63-3.64) with effective deformation potential  $D_{\text{pop}}$  and frequency  $\omega_{\text{pop}}$  reported in Tab. 3.1. This approximation was adopted to circumvent the numerical difficulties of implementing non local-self-energies, which would prevent the use of the recursive algorithms used to handle 3D quantum transport problems [42].

Phonons:	$D_{ac}$ [eV]	$D_{pop}$ [ $10^8$ eV/cm]	$\hbar\omega_{pop}$ [meV]
Electrons	5.8	20	30
Holes	1.0	20	30

Table 3.1: Phonon scattering parameters for InAs. Acoustic deformation potentials  $D_{ac}$  and phonon energy  $\hbar\omega_{pop}$  for polar phonons are from [11]. Mass density  $\rho=5.67$ [g/cm<sup>3</sup>] and sound velocity  $v_s=3\cdot 10^5$ [cm/s] are from [43]. The value of  $D_{pop}$  is discussed in the text.

Finally, it is important to recall that in the presented derivation of the phonon scattering self-energy functions a bulk model has been assumed for the phonon system. Effects of quantization on the phonon dispersion relation [44] due to the confinement on the transverse planes will be here neglected .

# Chapter 4

## Comparison between Si and InAs in FinFETs

This chapter is devoted to the application of the NEGF transport model presented in Chapter 3 to study the effect of the channel material on 3D FinFET device for the shortest technological node of the ITRS [45].

Sec. 4.1 addresses the motivation of this first study. This is given by the necessity of improving the efficiency of devices for low power applications. The first technological option to achieve such a goal is the adoption of a high-mobility semiconductor instead of Si as the channel material. Sec. 4.2 describes the device architectures that we have analyzed, while the results of our numerical simulations are presented in Sec. 4.3. Conclusions are given in Sec. 4.4.

## 4.1 III-V materials for nanoscale electronics

MOSFET technology scaling has encountered a soft technological wall in the last years and entered the so-called 'power-constrained scaling' phase. The surface power density in chips has reached the value of approximately  $100 \text{ W/cm}^2$  [46], and now remains at this value, given that increasing the power density further would force system architects to raise the cost and size of cooling systems, negating their usefulness in the most common applications. The natural response to this issue is to further reduce the supply voltage. However, this strategy has the cost of reducing switching speed, which is why  $V_{DD}$  has stagnated around 1V ever since the 130nm channel length node. Solving this issue is necessary to resume downscaling.

In an effort to improve the performances of nano-scale transistors, the industry has been looking toward materials that will outperform Si-based field effect transistors [47, 48, 49]. In the last years, it has been shown that, specifically, GaAs, InAs and InGaAs MOSFET have the ability to outperform Si based MOSFETs thanks to their intrinsic superior electron mobility [50, 51, 52, 53, 54, 55], as exposed in Fig. 4.1. This can lead to a reduction in supply voltage without losing performances. Thus there is already a sizable portion of the industry using those materials for a variety of applications ranging from radars to portable electronics [56].

However, before they replace silicon as a channel material for CMOS technology, III-V compounds must overcome difficult challenges. The first and most obvious one is they must provide a significant improvement over silicon, and they must also remain scalable for several technological nodes, and not just one.

A recurring issue for small-scale transistors is the difficulty of maintaining electrostatic integrity. That is to say the gate must provide a greater control of the electron concentration and energy level in the channel than in the drain. Single gate architectures have poor electrostatic control over the channel, the consequences of which are regrouped under the phrase "short channel effects" (SCE). To solve this issue an optimal device architecture has to be used.

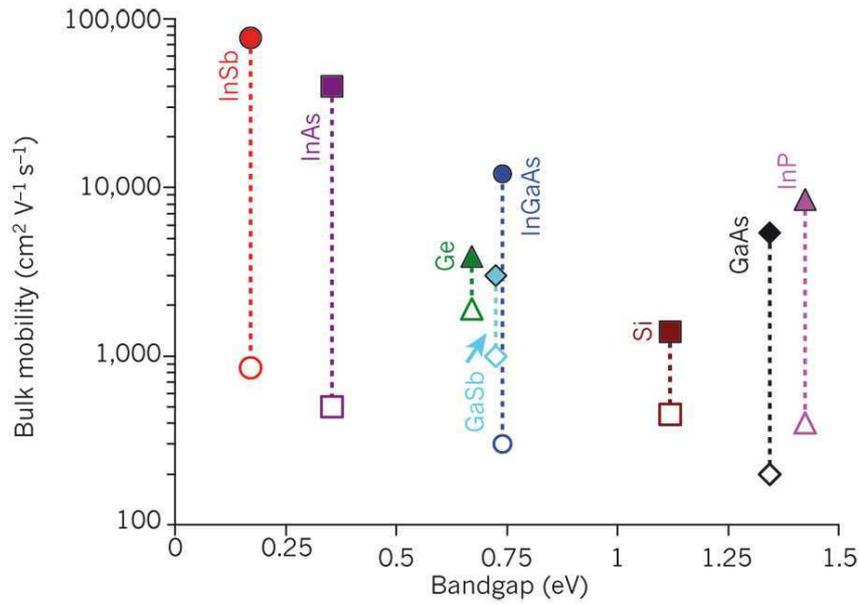


Figure 4.1: Electron and hole bulk mobility in silicon, germanium, and some III-V materials, versus their respective bandgap. Filled symbols show the electron mobility, and open symbols show hole mobility. III-V compounds offer a mobility increased by a factor of more than 10 over silicon. Source: [57]

#### 4.1.1 FinFET as a solution for transistors scaling

FinFETs (sometimes called tri-gate) are a new technology for the construction of nanoscale CMOS, recently introduced into the production cycle. A FinFET is in principle a rectangular MOSFET in which the electrostatic control of the channel is achieved by a gate on three sides. They have been shown to exhibit a better electrostatic control, and smaller short-channel effects than single-gate architectures [58, 59]. Previous simulations [60] have already shown that III-V FinFETs ( $\text{In}_{0.75}\text{Ga}_{0.25}\text{As}$  based) can perform similarly to Si FinFETs in short channel (12nm) devices with a supply voltage  $V_{DD}$  of 0.7V, while with a one-gate configuration, a Si FET outperforms a III-V FET significantly. This indicates that III-V based devices take better advantage of the improved electrostatic control than silicon ones.

This study focuses on exposing the conditions, if there are any, under which an InAs FinFET can outperform a Si FinFET, and providing an explanation for it.

## 4.2 Devices under investigation

For this study we have simulated two FinFETs with similar architecture, which is sketched in Fig. 4.2. The channel section is a  $W_{\text{fin}} = 5\text{nm}$  by  $H_{\text{fin}} = 20\text{nm}$  rectangle, with a 1nm thick gate on the top, left and right sides. Table 4.1 gives the device parameters used in those simulations. The InAs FinFET device uses low dopant concentration in the source and drain for  $V_{DS}$  values between 0.3V and 0.6V. For  $V_{DS} = 0.7\text{V}$  however, the strong ambipolar behavior of the FinFET forces us to adopt a strong dopant concentration, without which the device cannot reach the target  $I_{\text{off}} = 5\text{nA}/\mu\text{m}$  for any  $V_{GS}$  value.

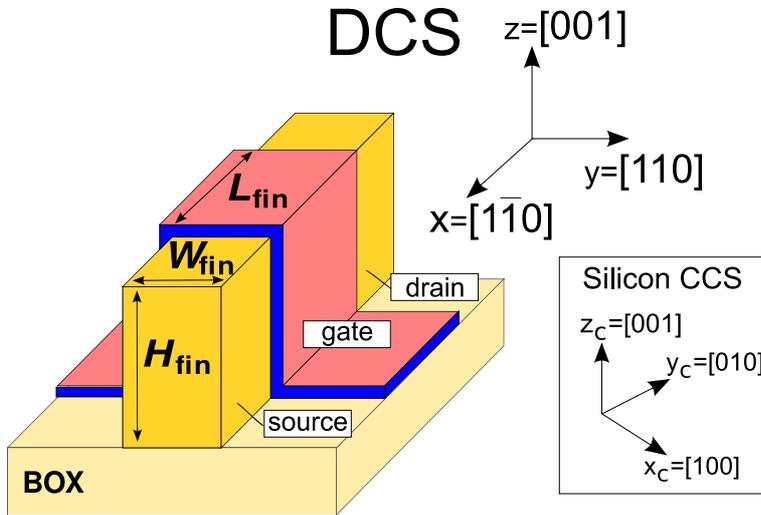


Figure 4.2: Schematic of a Si (110)/⟨110⟩ FinFET with a standard Manhattan layout in a (100) wafer. The device coordinate system (DCS) and the silicon crystal coordinate system (CCS) are shown.

We used the NEGF model described in 3, with an 8 band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian in the case of the InAs FinFET, and an effective mass one for the silicon FinFET [35], since the higher bandgap value in silicon makes band-to-band tunneling insignificant. Moreover, several simulations have previously shown that the validity of such an approximation in Si channels with thickness larger than 5 nm [61]. Phonon scattering was accounted for via the self-consistent Born approximation (see Sec. 3.6).

	$L_S/L_D$ [nm]	$T_{ox}$ [nm]	$N_S$ [cm <sup>-3</sup> ]	$N_D$ [cm <sup>-3</sup> ]
Si FinFET	10	1	10 <sup>20</sup>	10 <sup>20</sup>
InAs FinFET	20	1	6.5·10 <sup>18</sup>	6.5·10 <sup>18</sup>

Table 4.1: Device parameters used in the simulations. The gate dielectric is SiO<sub>2</sub>.  $L_S$  and  $L_D$  are respectively the length of source and drain regions,  $N_S$  and  $N_D$  are the doping concentrations in the source and drain (donors in both source and drain).  $L_D=20\text{nm}$  assures charge neutrality close to the drain contact for all the  $V_{GS}$  values in InAs. The source and drain do not need to be as long to reach charge neutrality in Si. †: higher doping value used for  $V_{DS}=0.7\text{V}$  to suppress ambipolar behavior.

### 4.3 Simulation results

The transfer characteristics of the Si and of the InAs based FinFET are shown in Fig. 4.3 at  $V_{DS}=0.4\text{V}$  (a) and at  $V_{DS}=0.6\text{V}$  (b). In both cases, the InAs device retains a better subthreshold slope value, very close to the ideal value of  $K_B T/e=60\text{mV/decade}$  value. Consequently, since the two MOSFETs as such small supply voltages work almost in the subthreshold voltage regime, the InAs FinFET presents the best on-current value at both  $V_{DS}=0.4\text{V}$  and  $V_{DS}=0.6\text{V}$ .

Similarly, Fig. 4.4 compares the surface charge density in the middle of the channel, far from the source and drain interfaces, in the same scenario as in Fig. 4.3. In the subthreshold regime, the InAs curve displays a higher slope, showing that the control of the carrier concentration in the channel is stronger than in the Si device. This is coherent with the smaller value of the inverse sub- $V_T$  slope of the InAs FinFET seen in the previous Figure.

These results show a promising scalability of III-V compounds for lower supply voltage applications. The InAs FinFET takes better advantage of the electrostatic control of the channel, and can switch over a smaller  $V_{GS}$  range than the Si FinFET.

#### 4.3.1 Impact of the device electrostatics

The improved subthreshold slope previously reported for the InAs MOSFET now needs to be clarified. A possible explanation could be a different impact

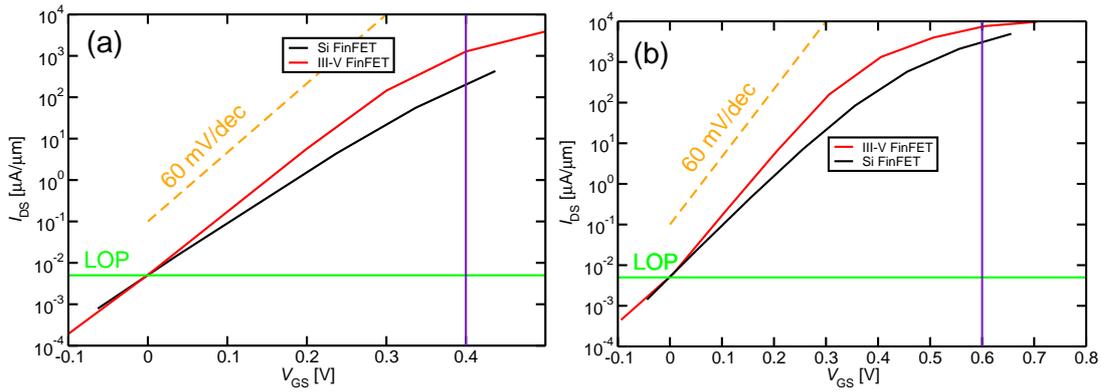


Figure 4.3:  $I_{DS}$  vs.  $V_{GS}$  characteristic for Si-based and InAs-based FinFETs with  $V_{DS}=0.4V$  (a) and  $V_{DS}=0.7V$  (b). The theoretical ideal SS is shown (dashed line), as well as the  $V_{GS,on}$  value (vertical solid line). The InAs FinFETs exhibit a near-ideal SS, albeit a lower saturation current. Thus, the InAs FinFET displays better  $I_{on}$  at  $V_{DS}=0.4V$ , and the Si FinFET displays better  $I_{on}$  at  $V_{DS}=0.7V$ .

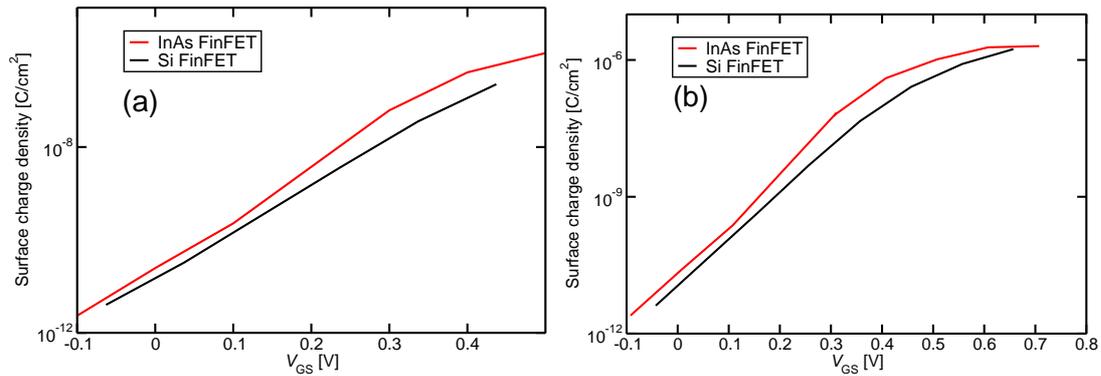


Figure 4.4: Surface charge density in the middle of the channel in Si and InAs FinFETs at  $V_{DS}=0.4V$  (a) and  $V_{DS}=0.7V$  (b). The charge density comparison follows the same trend as the  $I_{DS}$  vs.  $V_{GS}$  curves in fig.4.3, showing the on current difference has to do with the control of the carrier population in the channel.

of the tunneling current for the two devices. To verify this possibility we have studied in Fig. 4.5 the current spectra at  $V_{DS}=0.4V$  for the Si FinFET and the InAs FinFET, for  $V_{GS}=0V$  and  $V_{GS}=0.3V$ . This Figure shows current spectra with similar qualitative peculiarities. In fact, there is even a visible amount of tunneling transmission happening in the off state in InAs, but this is not enough to make the SS value worse than in silicon. The lower SS value in InAs cannot

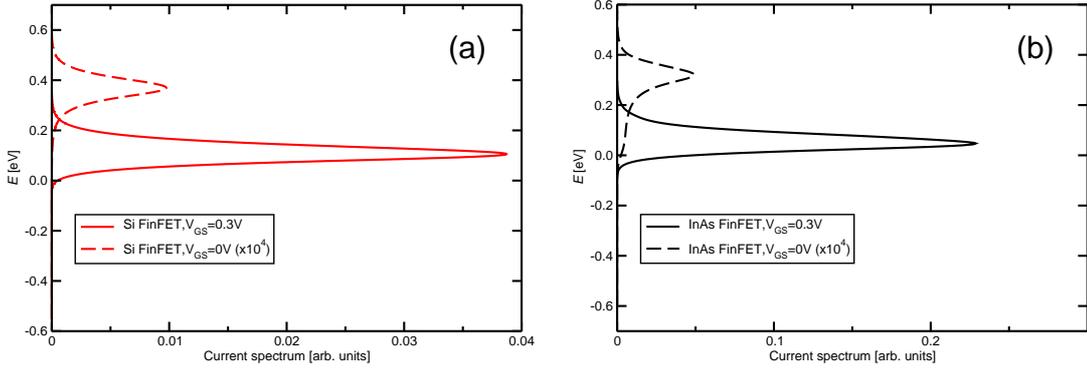


Figure 4.5: Energy-resolved current spectrum for Si (a) and InAs (b) FinFETs at  $V_{DS}=0.4V$  for  $V_{GS}=0V$  and  $V_{GS}=0.3V$ . In both figures the  $V_{GS}=0V$  spectrum is scaled up to be visible. The  $V_{GS}=0V$  curve in fig.(b) reveals the presence of band-to-band tunneling in the off state in the InAs FinFET.

be attributed to different transmission properties.

To explain the transfer characteristics of Fig. 4.3, it is more revealing Fig. 4.6. This is a plot of the first conduction subband profile along the transport direction in both devices, for  $V_{GS}=0V$  and  $V_{GS}=0.3V$ . The same variation in gate voltage causes a stronger decrease of the top-of-the-barrier energy level in InAs than Si. The channel energy level is more strongly constrained by the source and drain levels in the silicon device. In Fig. 4.6 it is clearly visible the different barrier shape of the two devices. In fact, due to the lower doping concentration in the source and drain region, the effective barrier seen by electrons travelling from source to drain is much larger in the InAs FinFET than in the Si FinFET. The smaller doping concentration ( $N_D = 6.5 \cdot 10^{18} \text{cm}^{-3}$  w.r.t.  $N_D = 10^{20} \text{cm}^{-3}$ ) used in order to obtain the same subband degeneracy of about 100meV in the doped regions, implies a better electrostatic control of the channel by means of the gate voltage in the InAs FinFET. Hence, the Si FinFET is more strongly affected by short-channel effects than the III-V FinFET.

This property could be eventually exploited in future devices for low-operating power applications with  $V_{DD}$  smaller than 0.7V.

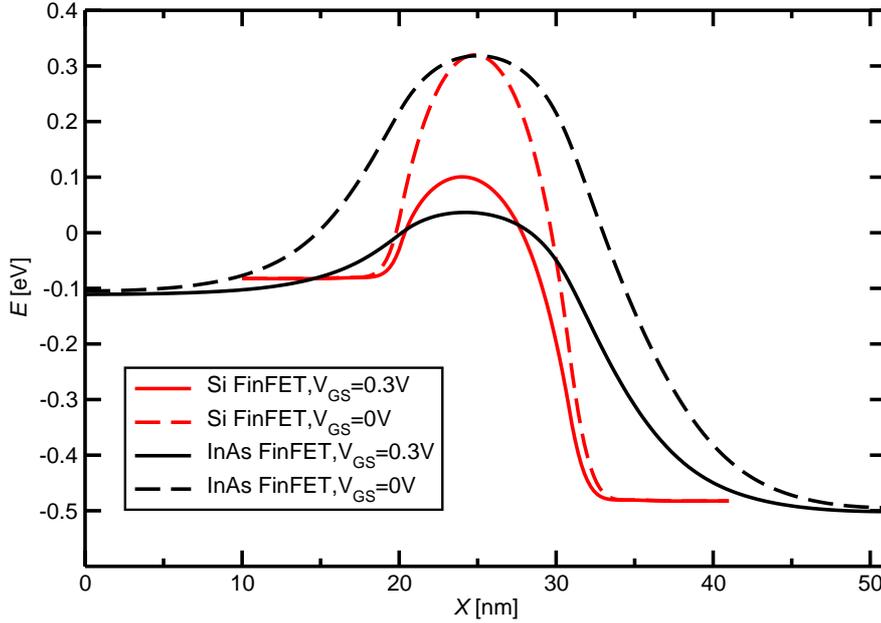


Figure 4.6: First conduction subband in the Si and InAs FinFETs, with  $V_{DS}=0.4V$ , at  $V_{GS}=0V$  and  $V_{GS}=0.3V$ , so as to be in the linear regime for both devices at both  $V_{GS}$  values.  $X$  values in InAs have been shifted by 10nm in order to have the channels in both devices aligned. The source-channel junction is at  $X = 20nm$ . The InAs device displays a larger amplitude in  $\Delta E_{max}$  for the same  $\Delta V_{GS}$ . The source and drain in the InAs FinFET reach electric neutrality at a longer distance from the channel than in the Si FinFET.

### 4.3.2 Supply-voltage scaling

It is interesting to compare the electrical performance of Si and InAs FinFETs for different values of supply voltages.

This is done in Fig.4.7 showing the resulting  $I_{on}$  values for different supply voltages. For  $V_{DD}$  values smaller than 0.7V, the InAs FinFET displays an improved  $I_{on}$  compared to the Si FinFET. For  $V_{DD}$  values larger than 0.7V the strong tunneling of electrons from the valence to the conduction subbands due to the small gap present in the InAs MOSFET, largely deteriorate their off current and it makes these devices unuseful to reach the ITRS specifications. This suggests that the use of InAs as channel material in FinFETs with similar sizes

is envisageable only at low  $V_{DD}$  values.

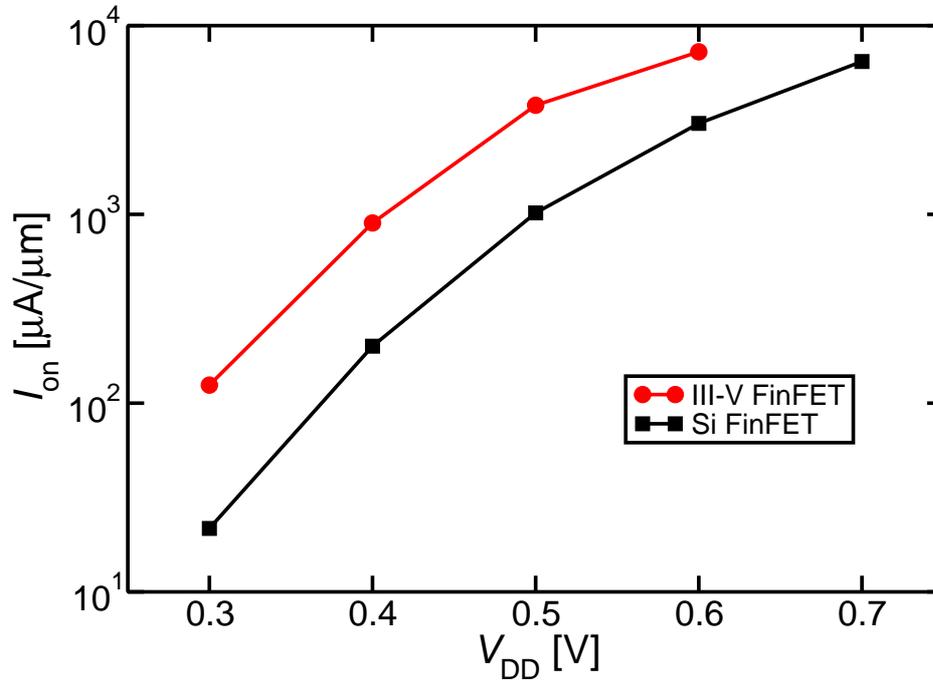


Figure 4.7:  $I_{on}$  against  $V_{DD}$  of InAs and Si FinFETs for a fixed  $I_{off}=5\text{nA}/\mu\text{m}$  target. The use of InAs instead of Si could provide a reduction of 100 mV of the supply voltage.

The strong band-to-band tunneling which characterize InAs devices will be in any case exploited to design steep slope devices working differently w.r.t. the MOSFET. This study will be the goal of the next Chapters.

## 4.4 Conclusion

This chapter has presented the results of the comparison of Si-based and InAs-based FinFETs via 3D numerical simulations. They have shown that III-V based CMOS transistors with a low doping profile have the potential to take better advantage of the improved electrostatic control provided by FinFET technology, and retain better performances under low supply voltage conditions than silicon based CMOS transistors. As an outcome of the results obtained in this chap-

ter, III-V compounds can be considered a promising candidate for the long-term downscaling of CMOS technology.

# Chapter 5

## Design of nanowire TFETs

This chapter is focused on the simulation of nanowire Tunnel-FETs (TFETs), which are one of the most promising technology options for low-power transistors. While MOSFET performances hit a hard theoretical limit of 60mV/decade for their inverse sub-threshold slope (SS), Tunnel-FETs can attain more competitive SS values [45], which can imply faster switching frequency, and lower off current. The main drawback of such a device architecture is that, since the on-current ( $I_{\text{on}}$ ) value is dominated by the band-to-band tunneling mechanism, it is typically much lower than in a conventional MOSFET for devices of similar dimensions, and quite below the specifications set by the ITRS. Consequently, there has been an important research effort to increase the on-current of TFETs. Two technological boosters were predicted to increase TFETs performances: the use of mechanical strain, and the use of type II hetero-junctions TFETs (H-TFET). They will be both explored in this Chapter.

Obtaining devices with both a serviceable  $I_{\text{on}}$  current and an inverse SS well below the MOSFET limit is the main challenge that the TFET technology has to take on, as discussed in Section 5.1. This chapter provides an extensive benchmark of the design options available to optimize the performance of TFETs [62]. Section 5.2 describes the geometry of the devices under consideration. Section 5.3 investigates the effect of strain on n-type InAs nanowire TFETs, whereas Section 5.4 is focused on the performances of GaSb-InAs heterojunctions. Finally, in Sec. 5.5 the influence of a large interface trap density at the interfaces between the semiconductor and the dielectrics is briefly discussed.

## 5.1 Technological context

The design of most electronic integrated circuits and systems is nowadays power or energy limited, and the performance practically attainable is strongly influenced by the energy efficiency of the signal processing [63]. Consequently, in the last decade VLSI processing at minimum energy has gained an ever growing interest for systems with aggressive requirements on the battery size and lifetime, as well as for energy-autonomous applications. Operation at nearly the minimum energy point, in turn, requires a very aggressive scaling of the supply voltage  $V_{DD}$ , down to only few hundreds mVs [64, 65]. At such extremely low  $V_{DD}$  values conventional MOSFETs work in weak-inversion or sub-threshold regime, and the abruptness of the transition between the off and the on state becomes of crucial importance for the ratio of on-current,  $I_{on}$ , to off-current,  $I_{off}$ , namely for the tradeoff between performance and standby leakage. As briefly said in Chapter 1, the Tunnel-FETs have been singled out by the International Technology Roadmap for Semiconductors as the most promising emerging transistors that, by reducing the sub-threshold swing,  $SS$ , to below 60mV/dec (which is a fundamental limit for MOSFETs at room temperature), may enable a voltage scaling in CMOS based integrated circuits to below 0.5V [45].

Quite many device architectures, material and technological options have been proposed in recent years, including graphene or 2D materials [66, 67, 68] and a review of many relevant achievements is available for example in [69, 70, 71]. In our PhD program we have worked on Tunnel FETs based on III-V semiconductors. Indeed, Tunnel FETs have been experimentally demonstrated by using both silicon and III-V semiconductors as a channel material. Given the maturity of silicon based CMOS technologies, it is probably not surprising that it was with silicon devices that both n-type and p-type transistors were first implemented in the same fabrication flow (albeit with a non optimized design), and that some vehicle circuits consisting of ring oscillators were reported and experimentally analyzed [72]. Given the relatively large energy bandgap of silicon and its indirect bandgap nature, however, the  $I_{on}$  in silicon Tunnel FETs is quite small for  $V_{DD}$  below 1V [73, 74], and, in particular, they are small compared to the corresponding currents in MOSFETs. In the context of group IV materials, on-current improvements are expected by the use of strained silicon [75], SiGe and Ge channel material [76, 77],

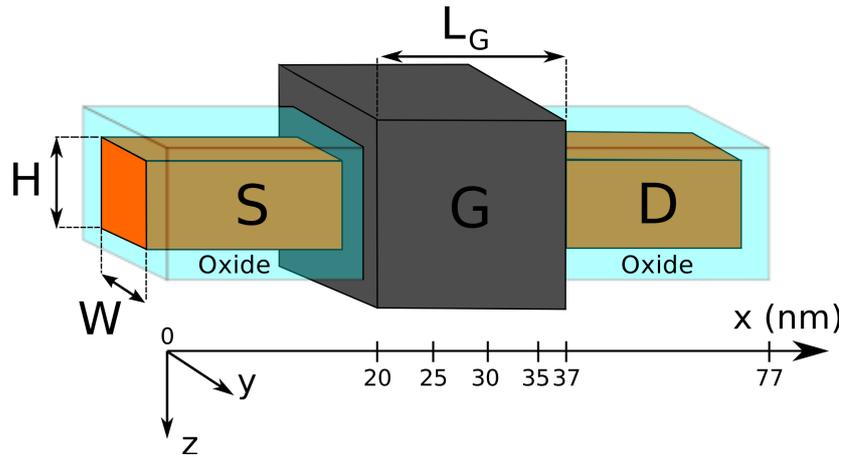


Figure 5.1: Sketch of the simulated nanowire transistors (not to scale) and for  $L_G=17\text{nm}$ . Transport direction  $x$  is assumed to be [100].

and direct bandgap GeSn [78, 79]. Some III-V semiconductors have a bandgap substantially smaller than silicon and are direct bandgap materials, hence III-V Tunnel FETs have raised legitimate expectations for  $I_{\text{on}}$  improvements compared to silicon or group IV material transistors [80]. III-V semiconductors also offer several options for staggered or broken bandgap hetero-junction with a reasonably small lattice mismatch and, in particular, the InAs/ $\text{Al}_x\text{Ga}_{1-x}\text{Sb}$  system has been identified as potentially interesting for band-to-band-tunneling transistors [81]. A main advantage of this kind of hetero-junction is that the InAs and  $\text{Al}_x\text{Ga}_{1-x}\text{Sb}$  have very close lattice constants [11].

Large  $I_{\text{on}}$  in hetero-junction Tunnel FETs was confirmed in experiments [82, 83, 84, 85, 86, 87], but with quite unsatisfactory values of sub-threshold swing. Very recent results, however, have reported InGaAs Tunnel FETs with SS values close to 60mV/dec at room temperature [88], which suggests that large SS values in III-V Tunnel FETs are not a fundamental problem, but are instead likely to be linked to larger interface defects compared to silicon devices (see also Sec.5.5), and may be improved by technology and device design optimization [89].

The inclusion of all these design option is important to provide simulations guiding the community in choosing the most promising ways to obtain efficient LOP devices [62].

## 5.2 Nanowire geometry

In order to analyze the various design options previously described, it is first important to identify an optimal device geometry.

	$L_S/L_D$ [nm]	$T_{ox}$ [nm]	$N_S$ [cm <sup>-3</sup> ]	$N_D$ [cm <sup>-3</sup> ]
Tunnel FET	20/40	1	$5 \cdot 10^{19}$	$10^{18}$
MOSFET	20	1	$3 \cdot 10^{19}$	$3 \cdot 10^{19}$

Table 5.1: Device parameters used in the simulations. The gate dielectric is SiO<sub>2</sub>.  $L_S$  and  $L_D$  are respectively the length of source and drain regions,  $N_S$  and  $N_D$  are the doping concentrations in the source and drain (respectively acceptor and donor type).  $N_D = 10^{18} \text{cm}^{-3}$  suppresses the ambipolar behavior [80], and  $L_D=40\text{nm}$  assures charge neutrality close to the drain contact for all the  $V_{GS}$  values.

The Tunnel-FETs work by exploiting an energy level mismatch at the source-channel junction. In an ideal Tunnel-FET, one would go from a non-broken bandgap to a broken bandgap configuration by varying the gate voltage value. In reality, this is not possible, because the energy mismatch at the interface is entirely determined by materials, doping, and strain conditions. What it is possible to achieve via electrostatic control is a bending of the subband profile in the channel. In order to maximize that electrostatic control of the channel, the transistors considered in this chapter are gate-all-around (GAA) nanowires as shown in Fig. 5.1. Transport takes place along the  $x$  axis, whereas the carrier confinement is along the  $y$ - $z$  plane. In the following all the simulated TFETs are systematically compared to an homojunction InAs MOSFET with the same size. The device parameters used in our simulations are summarized in Tab. 5.1. The calculations are performed with varying gate length ( $L_G$ ) values: 14, 17 and 25nm. The nanowire width  $D_W$  is set varying from 5, 7 and 10nm. Unless stated otherwise, the gate work-function is set independently in the MOSFETs and TFETs in order to have  $I_{\text{off}}=5\text{nA}/\mu\text{m}$  in the  $D_W=5\text{nm}$  nanowires, which is the LOP specification suggested by the ITRS [45].

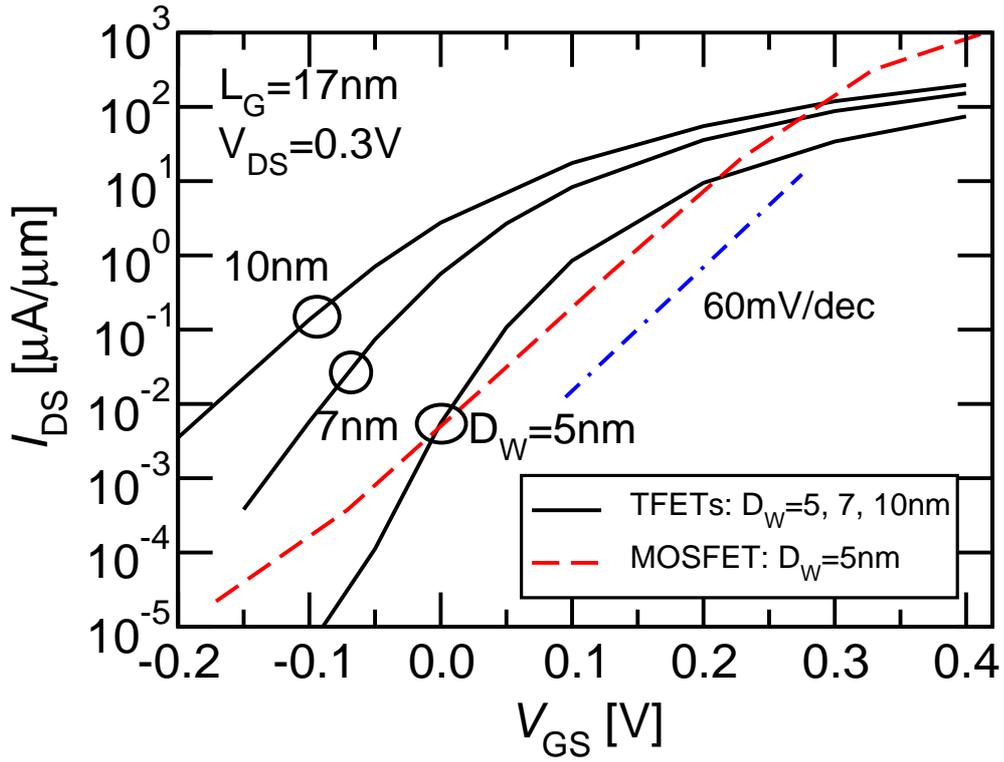


Figure 5.2:  $I_{DS}$  versus  $V_{GS}$  curves at  $V_{DS}=0.3\text{V}$  for InAs TFETs with different diameters  $D_W=W=H$  (see Fig. 5.1), and for the InAs MOSFET with  $D_W=5\text{nm}$ . The gate workfunction is  $WF=4.66\text{eV}$  for TFETs and  $WF=4.88\text{eV}$  for the MOSFET. For  $D_W=5\text{nm}$  the  $I_{\text{off}}$  is  $5\text{nA}/\mu\text{m}$ .

### 5.2.1 Nanowire width comparison

Usually, nanowire with a small width provide a better electrostatic control of the channel, while losing current throughput. Therefore, in our search for optimal TFET parameters,  $D_W$  is the first parameter we should take into consideration.

Fig. 5.2 shows the  $I_{DS}$  versus  $V_{GS}$  curves of InAs TFETs with varying values of  $D_W$ , as well as of an InAs MOSFET with  $D_W=5\text{ nm}$ , and the theoretical MOSFET limit of  $60\text{mV}/\text{decade}$ . The  $I_{DS}$  is normalized to  $D_W$ . The gate workfunction is identical in all TFETs, and set so that  $I_{\text{off}}=5\text{nA}/\mu\text{m}$  in the  $D_W=5\text{nm}$  devices. The first result to be pointed out is the shift towards smaller  $V_{GS}$  values of the IV curves of larger  $D_W$  values, which is a clear effect of the electrostatic degradation induced by the smaller confinement. As expected, the Tunnel-FETs have a better inverse subthreshold slope and, in particular, the  $D_W=5\text{nm}$  TFET

has a SS value of around 40mV/decade. However, although the MOSFET displays a higher inverse subthreshold slope than the TFETs, it has a better  $I_{on}$  for the same  $I_{off}$ , even with a  $V_{DD}$  value as low as 0.3V. Furthermore, if we try to compensate that low output current by increasing  $D_W$ , the graph shows that this would degrade the SS value significantly.

### 5.2.2 Nanowire length comparison

It is interesting to perform also a study on the channel-length dependence of the electrostatic integrity of InAs TFETs. Figs. 5.3 and 5.4 compare the SS and drain-induced barrier thinning (DIBT) in InAs nanowires for different values of gate length  $L_G$ . The DIBT is a metric for the evaluation of short channel effects in TFETs [90] and it is defined in a fashion similar to the DIBL in MOSFET. For a variation  $\Delta V_{DS}$  of the drain voltage it reads out

$$\text{DIBT} = -\frac{\Delta V_T}{\Delta V_{DS}}, \quad (5.1)$$

where  $V_T$  is the threshold  $V_{GS}$  value.

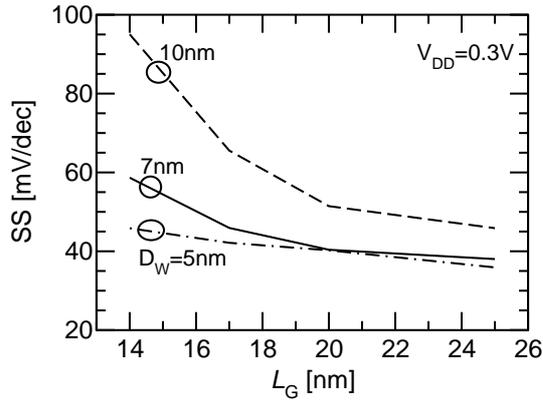


Figure 5.3: Inverse subthreshold slope SS (averaged for  $I_{DS}$  between  $1\text{nA}/\mu\text{m}$  and  $100\text{nA}/\mu\text{m}$ ) versus  $L_G$  and  $D_W$  for InAs TFETs.

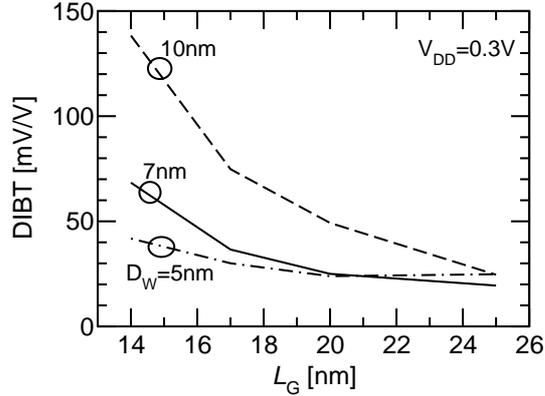


Figure 5.4: Drain induced barrier thinning defined as DIBL in MOSFETs, that is as the  $V_T$  reduction for  $V_{DS}$  increasing from 50mV to  $V_{DD}=0.3\text{V}$  [91].  $V_T$  was here defined as the  $V_{GS}$  yielding  $I_{DS}=100\text{nA}/\mu\text{m}$ .

Fig. 5.3 shows us that, below  $L_G=17\text{ nm}$ , the  $D_W=10\text{ nm}$  TFET cannot reach below the  $\text{SS}=60\text{ mV/decade}$  limit. Nanowires with width of 5 nm and 7 nm,

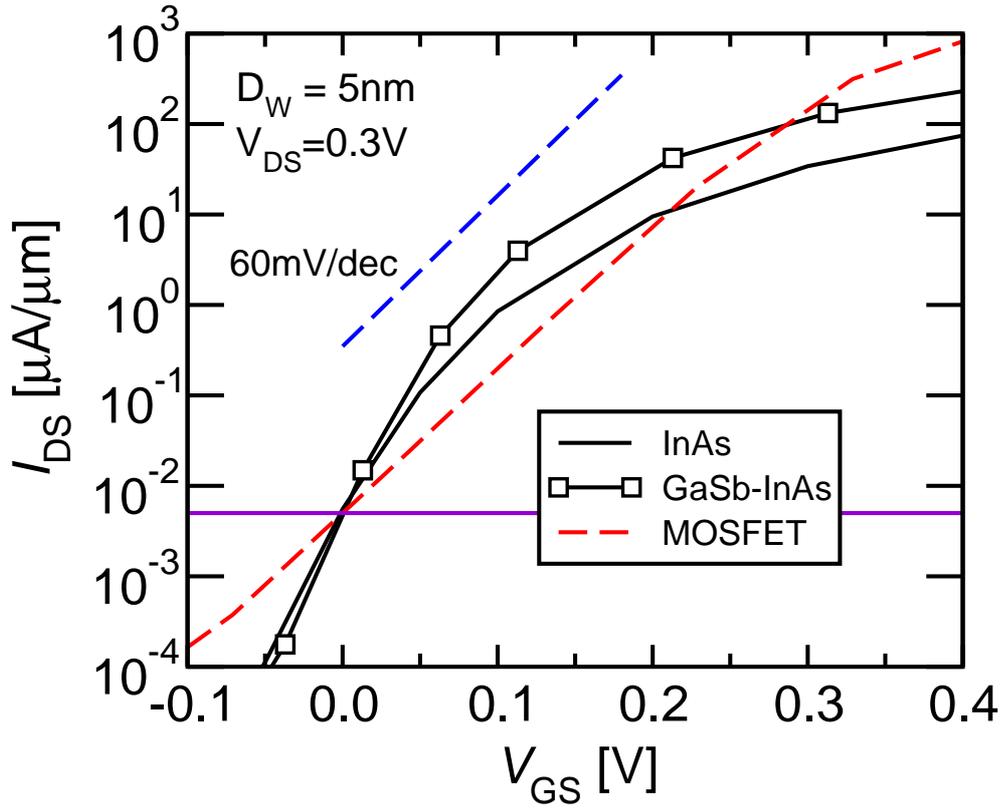


Figure 5.5:  $I_{DS}$  versus  $V_{GS}$  curves at  $V_{DS}=0.3\text{V}$  for two TFETs design options and for the MOSFET ( $L_G=17\text{nm}$ ,  $D_W=5\text{nm}$  and  $I_{\text{off}}=5\text{nA}/\mu\text{m}$ ). H-TFETs can improve  $I_{\text{on}}$  compared to an InAs TFET, however the  $I_{\text{on}}$  remains smaller than the MOSFET  $I_{\text{on}}$ .

however, show adequate SS and DIBT, so our analysis will focus on them. A further decrease in  $D_W$  would lead to too low current values. As noted in [80], a reduction of the gate length below 15nm leads to a severe degradation of both the SS and DIBT.

Since the homojunction InAs TFETs do not exhibit good enough performances w. r. t. the InAs MOSFET, other technological boosters have to be considered to make the TFET a competitive architecture.

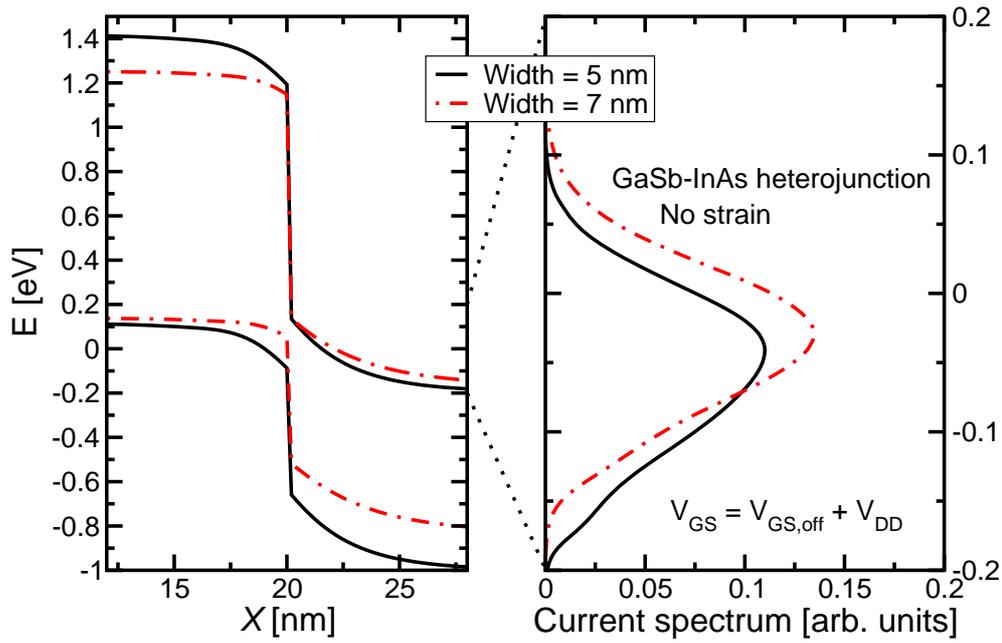


Figure 5.6: Spatial profile of the highest valence and of the lowest conduction subband (Left) and current spectrum as a function of energy (Right) for a GaSb-InAs H-TFET with  $D_W=5\text{nm}$  and  $7\text{nm}$ . The source Fermi level  $E_{F,s}$  is taken as zero.  $V_{GS}=V_{GS,off}+V_{DD}$  with  $V_{DD}=0.3\text{V}$ .  $L_G=17\text{nm}$ .

### 5.3 Impact of the hetero-junction

In the InAs TFET, the energy subband mismatch at the source-channel interface is determined solely by the difference in dopant density in the source and the channel. It is possible to obtain a much more abrupt mismatch by using two different semiconductors in the source and channel. Doing so brings us closer to an ideal subband profile configuration for a TFET. In this study we have computed the performance of GaSb-InAs heterojunction TFETs (H-TFETs), as we expect the energy levels mismatch between the two materials to create a broken, or near-broken bandgap configuration at the source-channel interface, thus increasing  $I_{DS}$  significantly.

The  $I_{DS}$  vs.  $V_{GS}$  curve for a GaSb-InAs H-TFET is shown in Fig. 5.5. As expected, the  $I_{on}$  value is greatly increased when compared to the InAs TFET, while retaining a virtually identical SS value. This means that the use of H-TFETs as on-current boosters can be more convenient w. r. t. other options

such as the use of mechanical strain, which results also in a SS deterioration [80]. Although previous studies [81] reported degraded SS values for heterojunction TFET, the present study is not in contradiction with them, because the device in [81] presents a broken bandgap configuration, while the present one does not due to quantum confinement effects [62]. This can be clearly observed in Fig. 5.6, which shows the potential profile for the GaSb-InAs heterojunction.

## 5.4 Impact of strain engineering

The evolution of the lowest conduction and of the highest valence subbands in the flat band condition as a function of the strength of the applied strain is illustrated in Fig. 5.7. A tensile biaxial strain raises the GaSb valence energy level, and reduces the InAs conduction energy level, pushing the system toward a broken bandgap configuration. This increases the tunneling current.

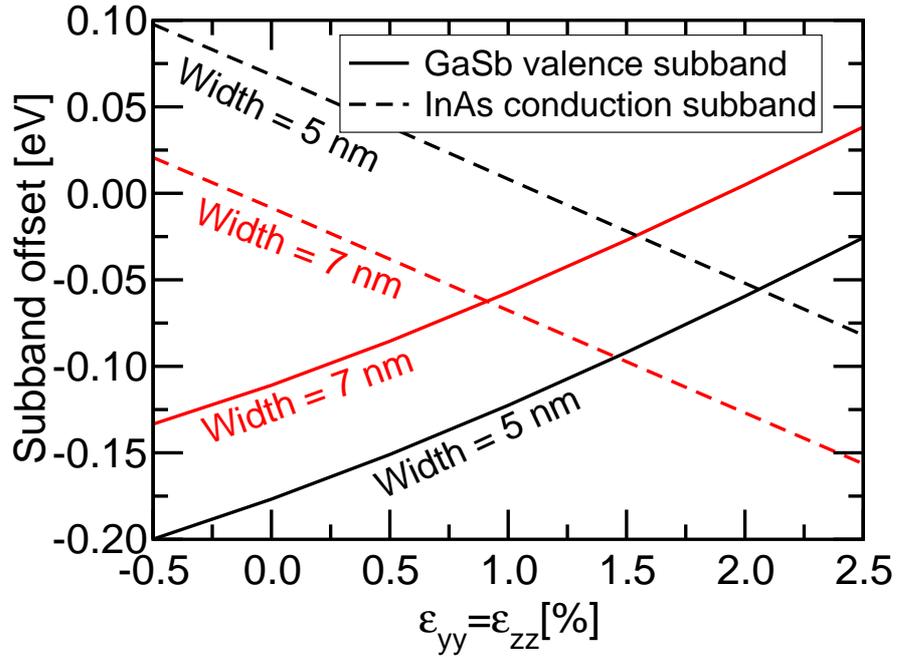


Figure 5.7: Valence subband edge for GaSb and conduction subband edge for InAs versus biaxial strain ( $\epsilon_{yy} = \epsilon_{zz}$ ) in the plane normal to the transport direction  $x$  and for a nanowire with  $D_W = 5\text{nm}$  and  $7\text{nm}$ . Deformation potentials from [11]. A tensile strain (i.e.  $\epsilon_{xx} = \epsilon_{yy} > 0$ ) helps induce a broken bandgap alignment in a GaSb-InAs H-TFET.

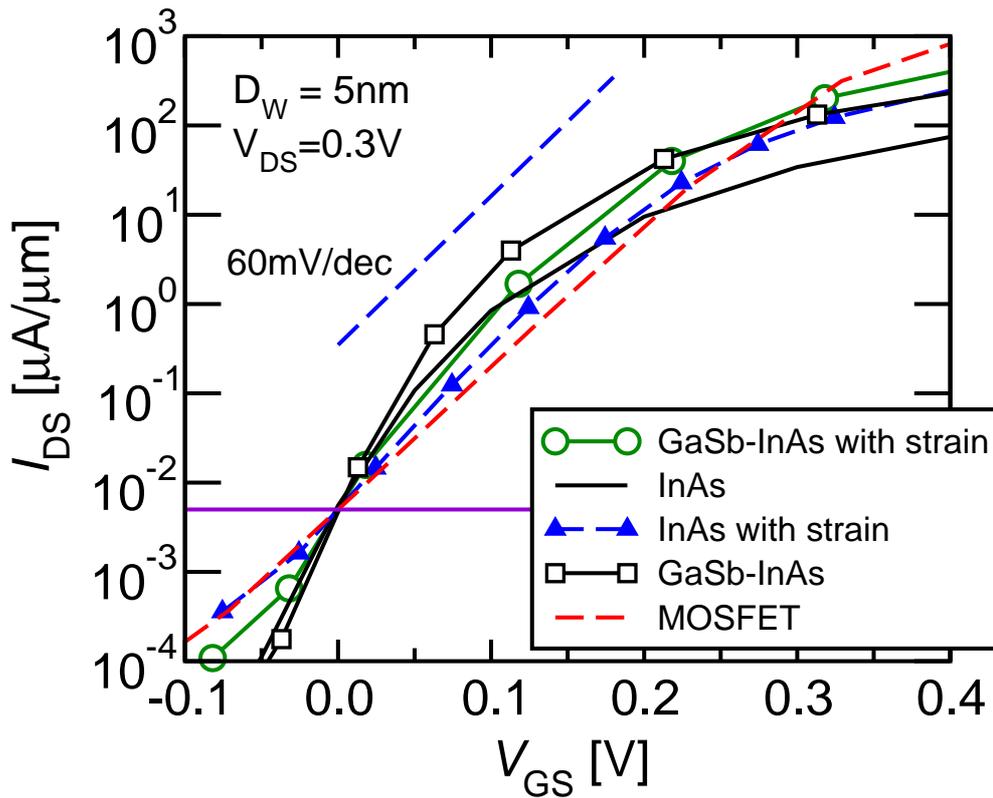


Figure 5.8:  $I_{DS}$  versus  $V_{GS}$  curves at  $V_{DS}=0.3\text{V}$  for four TFETs design options and for the MOSFET ( $L_G=17\text{nm}$ ,  $D_W=5\text{nm}$  and  $I_{\text{off}}=5\text{nA}/\mu\text{m}$ ). The strain is a 2% biaxial tensile strain. H-TFETs and strain can improve  $I_{\text{on}}$  compared to an unstrained InAs TFET, applying both allows to get a competitive  $I_{\text{on}}$ .

Fig. 5.8 illustrates such an on-current increase with strain, in both the InAs TFET and the H-TFET. As already stated, the use of strain considerably degrades the SS value. For the H-TFET, the strong SS degradation is explained by the fact that the strain brings the system in a fully broken bandgap configuration, as shown in Fig. 5.9. Moreover, the tensile strain has the effect to reduce the energy gap not only in the source region, which is important to increase the current in the ON state, but also in the channel region, which has the effect of increasing also the off current. A possible solution, even if technologically complicated, is to use a strain distribution which is not zero only in the source region [92].

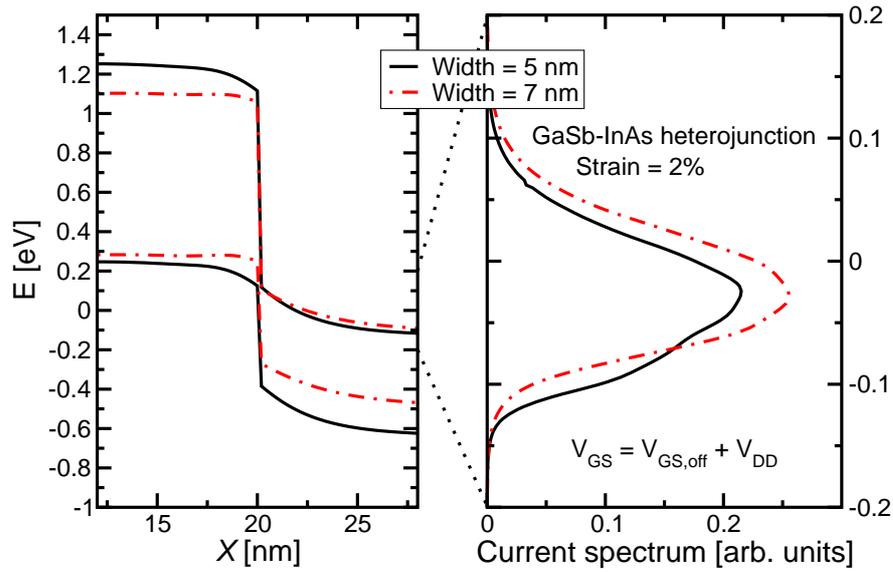


Figure 5.9: Subband profile (Left) and current spectrum (Right) as in Fig. 5.6 for a GaSb-InAs H-TFET with a 2% biaxial tensile strain and for  $D_W=5\text{nm}$  and  $7\text{nm}$ .  $V_{GS}=V_{GS,off}+V_{DD}$  with  $V_{DD}=0.3\text{V}$ .  $L_G=17\text{nm}$ .

### 5.4.1 Analysis results

The results obtained in this chapter are collected in Fig. 5.10. The  $I_{on}$  SS and DIBT for all devices under consideration are represented, for  $D_W=5, 7$  and  $10\text{nm}$ , and  $L_G=17\text{nm}$ . For devices with applied strain, only the  $D_W=5\text{nm}$  option is shown. This is because nanowires with a higher width cannot reach the target  $I_{off}$  value of  $I_{off}=5\text{nA}/\mu\text{m}$  for any  $V_{GS}$  value and therefore they have not been considered any more. The energy overlap in those strained TFET is such that ambipolar behavior becomes dominant for negative values of  $V_{GS}$ . The same applies for the MOSFET with  $D_W=10\text{nm}$ .

It should be noted that for all design option, increasing  $D_W$  actually results in a lower effective  $I_{on}$ . This apparently unexpected result is not in contradiction with the results shown in Fig. 5.2, because while in Fig. 5.2 we used a constant gate workfunction for all  $D_W$  values, in Fig. 5.10 we are comparing designs with different gate workfunctions, set in such a way that  $I_{DS}=I_{off}$  when  $V_{GS}=0\text{V}$ . Since the SS is severely degraded in wider nanowires, this leads to a lower effective  $I_{on}$  for a constant  $V_{GS,on}=V_{DD}=0.3\text{V}$ , even if the saturation current is higher. For

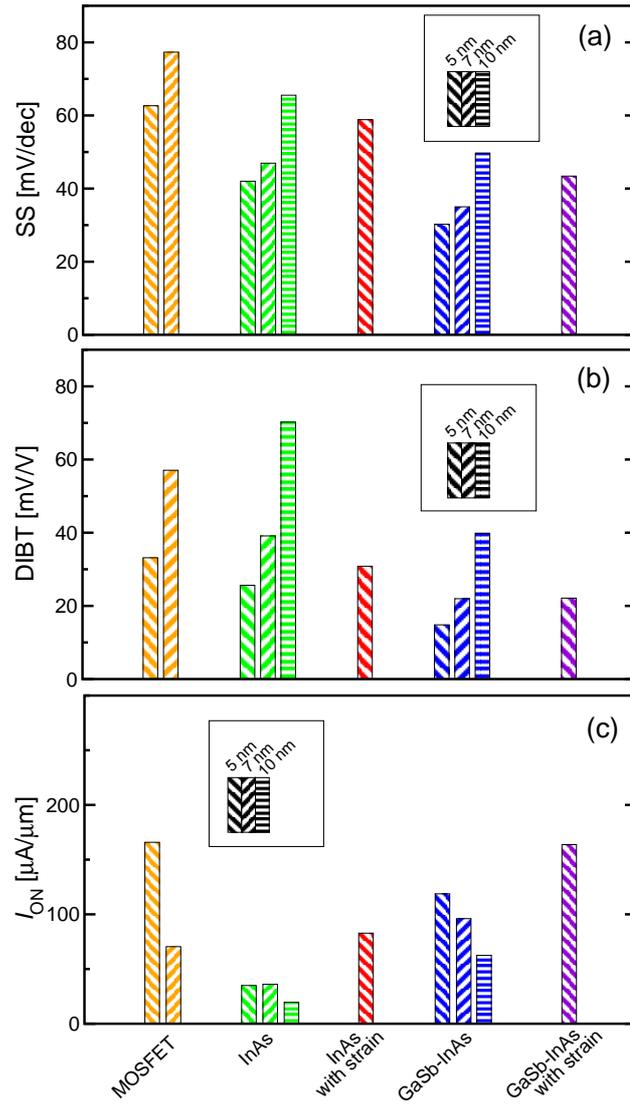


Figure 5.10: (a) Inverse subthreshold slope SS (averaged for  $I_{DS}$  between  $1nA/\mu m$  and  $100nA/\mu m$ ) for different design options of TFETs and for the InAs MOSFET. Different  $D_W=5, 7$  and  $10nm$  are shown for some options. For all devices the SS is degraded when increasing  $D_W$ .  $L_G=17nm$ . (b) Drain induced barrier thinning (DIBT) for different design options of TFETs and DIBL for the InAs MOSFET. Different  $D_W=5, 7$  and  $10nm$  are shown for some design options. For all devices the DIBT is degraded when increasing  $D_W$ .  $L_G=17nm$ . (c) On current  $I_{on}$  at  $I_{off}=5nA/\mu m$ , defined as the  $I_{DS}$  at  $V_{DS}=V_{DD}=0.3V$  and  $V_{GS}=(V_{GS,off}+V_{DD})$ , with  $V_{GS,off}$  being the  $V_{GS}$  giving  $I_{DS}=I_{off}$  [16]. For each design option a clear correlation exists between SS degradation and  $I_{on}$  reduction with increasing  $D_W$ .  $L_G=17nm$ .

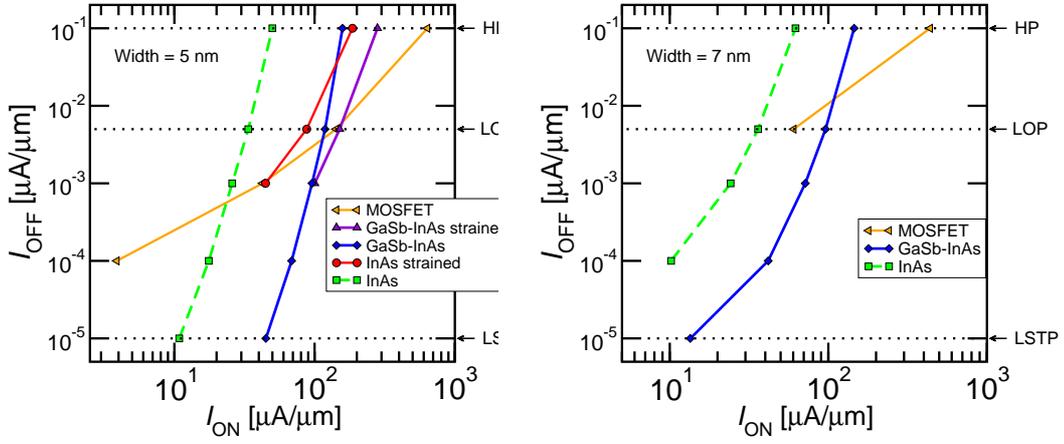


Figure 5.11:  $I_{\text{on}}$  vs. target  $I_{\text{off}}$  curves for all design options, for  $D_W=5\text{nm}$  and  $D_W=7\text{nm}$  respectively. The dotted horizontal lines represent the ITRS target  $I_{\text{off}}$  values for high power (HP), low power (LOP) and low standby power (LSTP) requirements. Missing points for low  $I_{\text{off}}$  values mean the device cannot reach the target  $I_{\text{off}}$  for any gate voltage condition.

that reason, nanowires with  $D_W=5\text{nm}$  appear to be the best design choice.

By applying both technology boosters, namely the use of tensile biaxial strain, and of a III-V heterojunction, it is possible to obtain a TFET with a SS below the theoretical MOSFET limit of 60 mV/decade, while keeping a current value comparable to MOSFET designs, albeit not improving on it.

One can legitimately wonder how those devices compare with different values for the target  $I_{\text{off}}$ . This is answered in Fig. 5.11, which compares the effective  $I_{\text{on}}$  for different target values of  $I_{\text{off}}$ . We can see that MOSFET performances (in orange) are the best for high power applications, however they severely degrade for lower power applications, so much so that nanowire MOSFET cannot reach the  $I_{\text{off}}$  condition for low standby power applications.

For LOP applications, the strained H-TFET displays the best performances. However strained devices cannot reach low standby power requirements either. Finally, unstrained H-TFETs are the vastly superior option when it comes to LSTP applications. Their optimal subthreshold slope guarantees a good effective  $I_{\text{on}}$  even at the lower power applications.

## 5.5 SS degradation induced by traps

It is important to remark that the analysis of the different design options proposed in the previous Sections have considered ideal material and interfaces without taking into account any kind of defects such as surface roughness [93, 94, 95] or traps at the interfaces between the semiconductors and the oxide [96, 97].

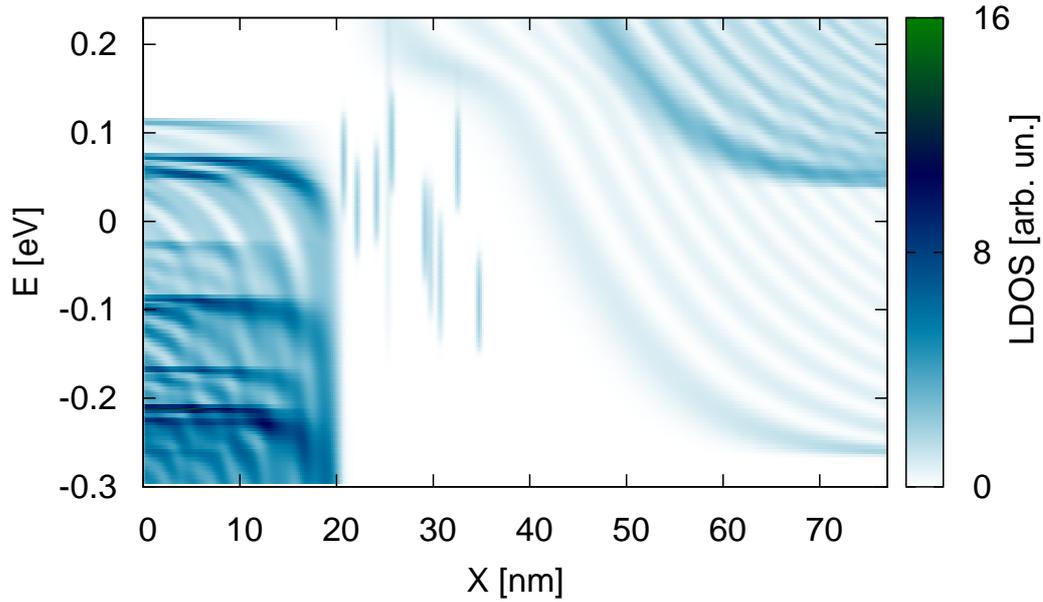


Figure 5.12: Energy-resolved LDOS at  $V_{DS}=0.3\text{V}$  and  $V_{GS}\approx-0.1\text{V}$  of an InAs Tunnel-FET in the presence of randomly distributed interface traps in the channel region ( $20\text{nm} < x < 37\text{nm}$ ) with a trap energy  $E_{T,FB}=-94\text{meV}$  and a trap density  $D_{IT}=2\cdot 10^{12}\text{cm}^{-2}$  [89]. The LDOS produced by the traps is clearly visible at different spatial positions and at different energies.  $L_G=17\text{nm}$ .

As previously stated, the presence of interfaces states are a serious concern for III-V based devices. They act as stepping stones in band-to-band tunneling processes, eventually enhancing phonon-assisted processes, and therefore they are probably the cause of the deteriorated SS often measured in experiments [98, 99].

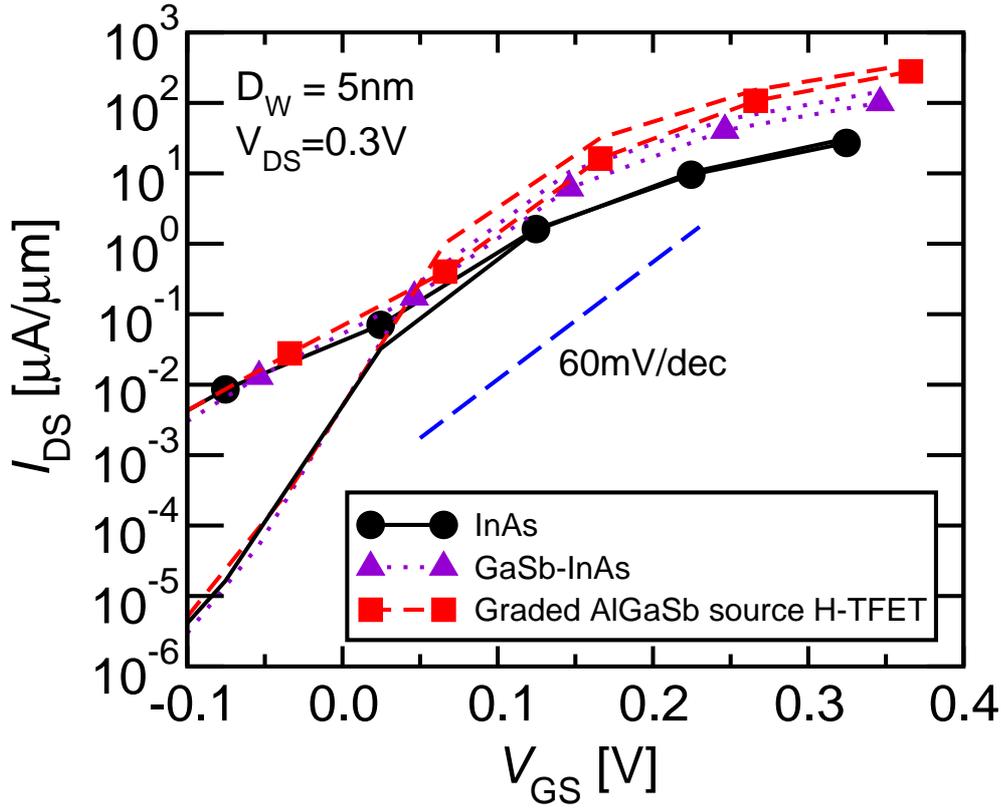


Figure 5.13:  $I_{DS}$  versus  $V_{GS}$  at  $V_{DS}=0.3V$  for a homo-junction InAs TFET, a GaSb-InAs H-TFET and a graded  $Al_{x_M}Ga_{(1-x_M)}Sb$  source H-TFET. IV curves are shown both for devices with no traps (lines) and for a device realization with a trap energy  $E_{T,FB}=-94meV$  and a trap density  $D_{IT}=2\cdot 10^{12}cm^{-2}$  (lines plus symbols). In all the device designs the traps degrade remarkably the SS value.  $L_G=17nm$ .

According to [96, 97], the traps were included in the transport model by generating electrically active 0-D states. This was obtained by writing the complete Hamiltonian in real space as:

$$\hat{H}(\mathbf{r}, -i\nabla_{\mathbf{r}}) = \hat{H}_{\mathbf{k}\cdot\mathbf{p}}(-i\nabla_{\mathbf{r}}) + E_{C(V)}(\mathbf{r}), \quad (5.2)$$

where  $\hat{H}_{\mathbf{k}\cdot\mathbf{p}}(-i\nabla_{\mathbf{r}})$  is the kinetic operator given by the  $8\times 8$   $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian [10], and the band profiles  $E_V(\mathbf{r})$  and  $E_C(\mathbf{r})$  are given by

$$\begin{aligned} E_V(\mathbf{r}) &= E_{V0} - q\phi(\mathbf{r}) \\ E_C(\mathbf{r}) &= E_{C0} - q\phi(\mathbf{r}) + \sum_{t=1}^{N_T} V_t(\mathbf{r}, \mathbf{r}_t) \end{aligned} \quad (5.3)$$

with  $\phi(\mathbf{r})$  being the electrostatic potential, and  $E_{C0}$  ( $E_{V0}$ ) the material dependent valence and conduction band edge. The position of a trap is here denoted by  $\mathbf{r}_t=(x_t,y_t,z_t)$  and  $N_T$  is the number of traps in the device. The  $E_{C0}$  discontinuity at the III-V – oxide interface was set to 3.0eV. A trap enters Eq. (5.3) as a cubic potential well  $V_t(\mathbf{r}, \mathbf{r}_t)$  superimposed to the conduction band profile, where  $V_t(\mathbf{r}, \mathbf{r}_t)$  is  $[-V_{\text{depth}}]$  if  $(\mathbf{r}-\mathbf{r}_t)\in C_t$  and it is null otherwise, with  $C_t$  being the cube representing the trap. This approach for the trap modeling has the advantage to account for the discrete nature of traps, and to provide zero-dimensional electrically active states, which can both modify the device electrostatics and play a direct role in the carrier transport.

The depth  $V_{\text{depth}}$  of the quantum well describing the trap can be used to modify the trap energy level. In fact the NEGF simulations neatly show the local density of states (LDOS) corresponding to a trap. For example, Fig. 5.12 shows the LDOS in one statistical realization of a GaSb-InAs H-TFET with a  $D_{\text{IT}}=2\times 10^{12}\text{cm}^{-2}$  located in the channel region going from 20 to 37 nm (corresponding to around 10 traps) [62].

The influence of traps on the I-V curves of the different design options studied in this Chapter is shown in Fig. 5.13. As can be seen, at least for the trap density considered in these simulations, the interface states severely degrade the SS and thus undermine all design optimizations. In order to realize efficient steep slope devices based on tunneling effects it will be therefore crucial to find technological recipes able to strongly reduce similar impurity densities.

## 5.6 Conclusion

This chapter has presented the results of the numerical simulation and optimization study of InAs nanowire Tunnel-FETs, and GaSb-InAs heterojunction-TFETs, with and without the application of tensile biaxial strain, with a benchmark against nanowire InAs MOSFETs of similar dimensions. Some tradeoffs between a good subthreshold swing and a large drive current have pointed out fundamental aspects in the working principle of these devices such as the inherent interplay between good electrostatic control and large quantum confinement effects.

In the light of these results, while H-TFETs are most certainly the preferable

---

option for LSTP applications, their performances, even after optimization, are not enough to replace nanowire MOSFETs as the driving technology for low operating power applications. A more efficient solution to improve the electric performance of Tunnel-FETs has to be considered, which is the topic of the next chapter.



# Chapter 6

## Source region engineering of III-V Tunnel-FETs

In this chapter we present two original options to modify the design of hetero-junction III-V TFETs with the goal of largely increasing their on-current values at least for small values of  $V_{DD}$  and eventually outperform the standard MOSFET.

Sec. 6.1 details the physical mechanism inspiring these kinds of approach. Sec. 6.2 presents the basic idea behind the first design option consisting in the use of a grading of the molar fraction  $x_M$  of an AlGaSb ternary in the source region, as well as the the simulation results and their physical explanation. Sec. 6.3 focuses on the second technological option studied in this Chapter, consisting in the use of a InAs/GaSb/InAS quantum well at the interface between the source and the channel region. Finally, the conclusion is given in Sec. 6.4.

## 6.1 Tunneling path modification

From the analysis in the previous Chapter, we determined that a significant portion of the tunneling path in a GaSb-InAs heterojunction TFET is located in the depletion region of the source. This can be seen in Fig. 6.1, showing the tunneling path for the carriers near the maximum of the transmission spectrum, for a H-TFET in the ON state. Because of the source depletion, a noticeable part of the source highest valence subband is below the source Fermi level set at 0eV. A large fraction of the transmission spectrum remains unexploited, because of the absence of charge carriers in that energy range [100].

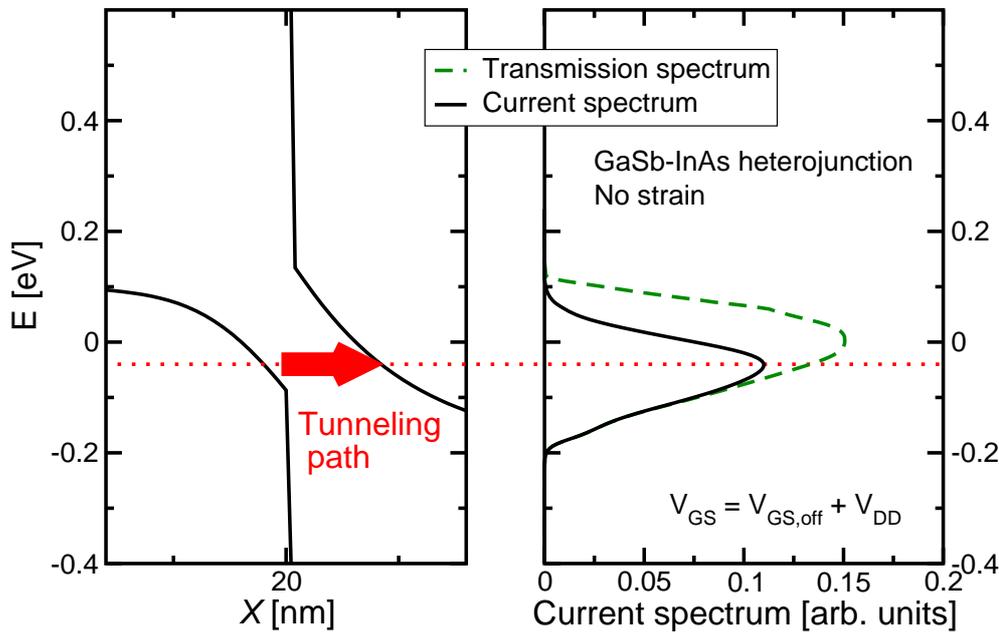


Figure 6.1: Subband profile (Left) and current spectrum and transmission spectrum as a function of energy (Right) for a GaSb-InAs H-TFET with  $D_W=5\text{nm}$ . The source Fermi level  $E_{f,s}$  is taken as zero.  $V_{GS}=V_{GS,off}+V_{DD}$  with  $V_{DD}=0.3\text{V}$ .  $L_G=17\text{nm}$ . The tunneling path giving the maximal transmission probability is marked in red.

Repopulating the source region with carriers with an energy above the Fermi level would lead to a better exploitation of the transmission spectrum, as well as a shorter depletion length and tunneling path, and thus a further improved transmission. This would benefit the on current value significantly. On the other hand, increasing the dopant density in the source region is not envisageable

because this would affect the SS value negatively.

## 6.2 Use of the molar grading

Another way to shorten the depletion length is to make use of the fortunate properties of the AlGaSb ternary alloy [101]. In  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$ , the molar fraction  $x_M$  decreases the valence band edge, as illustrated in Fig. 6.2. The molar fraction  $x_M$  in a simple nanowire is spanned from 0 to 1, while calculating the subband edges value. There is an almost linear energy differential of about 0.4eV between pure GaSb and pure AlSb. It has to be pointed out that the lattice constant of  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$  is almost unaffected by the molar fraction: GaSb lattice parameter is 0.6136nm, while AlSb lattice parameter is 0.6096nm [11]. This is particularly important in order to avoid undesired stress effect and generation of dislocation defects at the junction.

Fig.6.2 suggests us that by designing a device with a molar fraction grading in the source, it is possible to compensate for the valence band bending in the source depletion region, thus resulting in a valence band edge with a flatter profile than in the GaSb-InAs heterojunction-TFET. This principle is illustrated in Fig. 6.3. The proposed device is similar to a GaSb-InAs H-TFET, however instead of a pure GaSb source, the source near the contacts is pure AlSb, and there is a interfacial region of length  $T_{\text{grad}}$  between the AlSb layer and the InAs channel. That interfacial region is made of  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$  with a molar fraction  $x_M$  varying linearly from 1 (at the AlSb layer) to 0 (at the interface with the InAs channel). The fact that the grading in molar fraction does not change the lattice parameter significantly means no undesired strain is introduced in the device.

It should be noted that it is alternatively possible to modulate the molar fraction grading profile in order to obtain an almost perfectly flat valence band profile in the source. For the sake of simplicity, however, the present study focuses on linear molar fraction grading profiles [101].

### 6.2.1 Simulation results

In order to verify the validity of the molar fraction grading concept, we simulated gate-all-around nanowire TFETs, with an InAs channel and drain, and an AlSb-

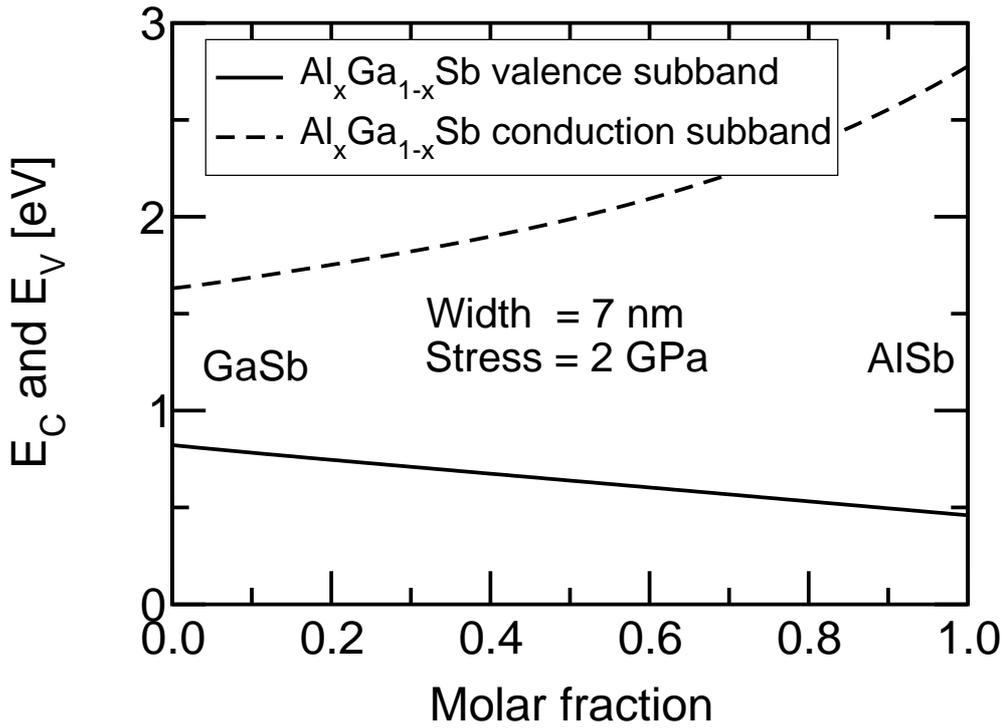


Figure 6.2: First valence subband and first conduction subband energy levels in a  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$  nanowire, as a function of the molar fraction  $x_M$ . The nanowire width is  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}=7\text{nm}$ , and under a tensile biaxial strain of 2GPa.

$\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$  source, with a progressive grading of the molar fraction near the source-channel interface, over a length  $T_{\text{grad}}$ , as shown in Fig. 6.3. Different  $T_{\text{grad}}$  values from 4nm to 10 nm have been tested in order to optimize the device. In all cases,  $x_M=0$  at the source-channel interface, thus it is a GaSb/InAs heterojunction in all cases. The transport direction is [100]. Based on the conclusions from the previous chapter for optimal geometrical parameters, the nanowires have a gate length  $L_G=17\text{nm}$ , and a width  $D_W=5\text{nm}$  in order to minimize the SS value through better electrostatic control. Electrostatic control is performed through an all-around  $\text{SiO}_2$  gate with a  $T_{\text{ox}}=1\text{nm}$  thickness. The source has a doping donor concentration  $N_S=5\cdot 10^{19}\text{cm}^{-3}$  and the drain has a doping acceptor concentration  $N_D=10^{18}\text{cm}^{-3}$ . The channel remains undoped, and we assume abrupt doping profiles with no diffusion of the dopants into the channel. In order to get charge neutrality at the source and drain extremity for all gate voltage  $V_{GS}$  values, the

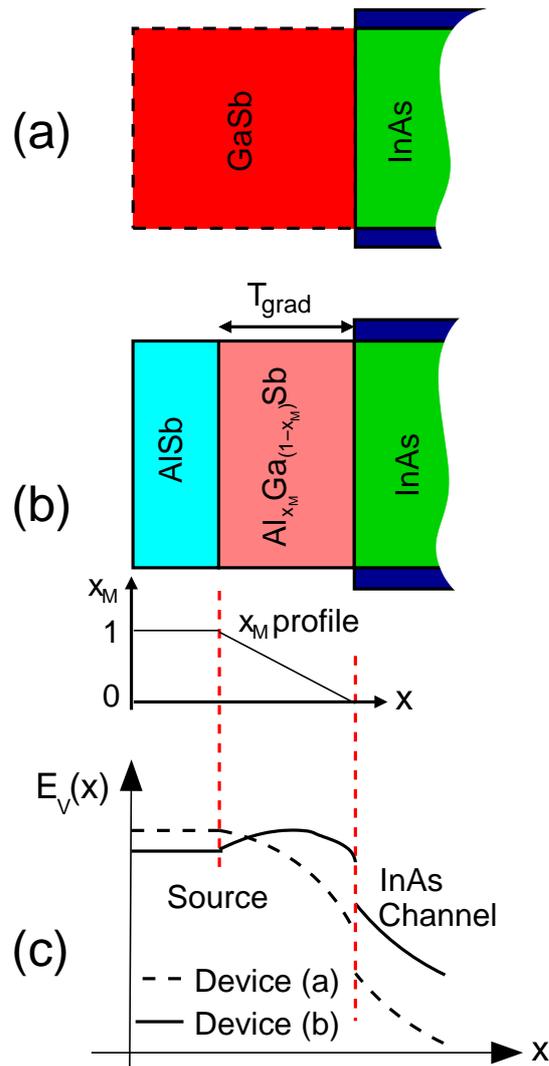


Figure 6.3: Sketch illustrating the source region of (a) a hetero-junction GaSb/InAs Tunnel-FET and (b) a hetero-junction  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}/\text{InAs}$  Tunnel-FET featuring a molar fraction grading in region of thickness  $T_{\text{grad}}$ . Part (c) shows the qualitative valence band profile for device (a) (dashed line, w/o molar fraction grading) and device (b) (solid line, with molar fraction grading). The thickness of the AlSb layer in (b) is given by  $(L_S - T_{\text{grad}})$ , with  $L_S = 20\text{nm}$  being the length of the source region.

source has to have an extension  $L_S = 20\text{nm}$  and the drain an extension  $L_D = 40\text{nm}$ .

The graded devices are benchmarked against the GaSb/InAs heterojunction-TFET and to the InAs MOSFET from previous chapter, with parameters de-

scribed in Table 5.1, and a width  $D_W=5\text{nm}$ . The  $\mathbf{k}\cdot\mathbf{p}$  parameters for  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$  depending on the molar fraction  $x_M$ , and for the other materials, have been drawn from [11].

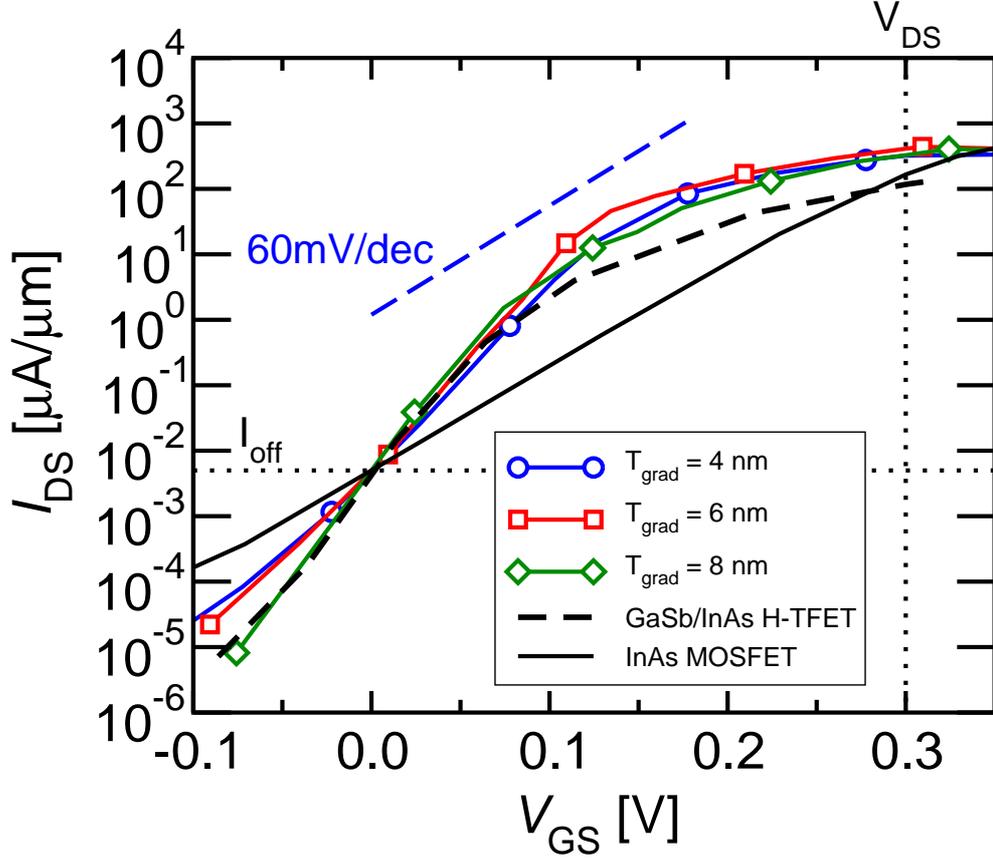


Figure 6.4:  $I_{DS}$  versus  $V_{GS}$  curves at  $V_{DS}=0.3\text{V}$  for three advanced grading options in  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$ -InAs H-TFET and at fixed  $I_{\text{off}}=5\text{nA}/\mu\text{m}$ . The GaSb-InAs H-TFET with no grading and the InAs MOSFET are also shown for comparison.

Fig. 6.4 displays the  $I_{DS}$  vs.  $V_{GS}$  curves for the graded source device for select  $T_{\text{grad}}$  values, as well as a comparison to the regular GaSb-InAs H-TFET, and the InAs MOSFET. Like in the previous analysis, the gate work function is set independently for each transistor so that  $I_{DS}=I_{\text{off}}=5\text{nA}/\mu\text{m}$  when  $V_{GS}=0\text{V}$ . The striking result of Fig. 6.4 is that the molar fraction grading in the source region can provide a significantly higher (around 80%) on current in the  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}$  H-TFETs, compared to the GaSb-InAs H-TFET, with virtually no subthreshold

slope degradation. A physical explanation for such a result will be given in Sec. 6.2.2.

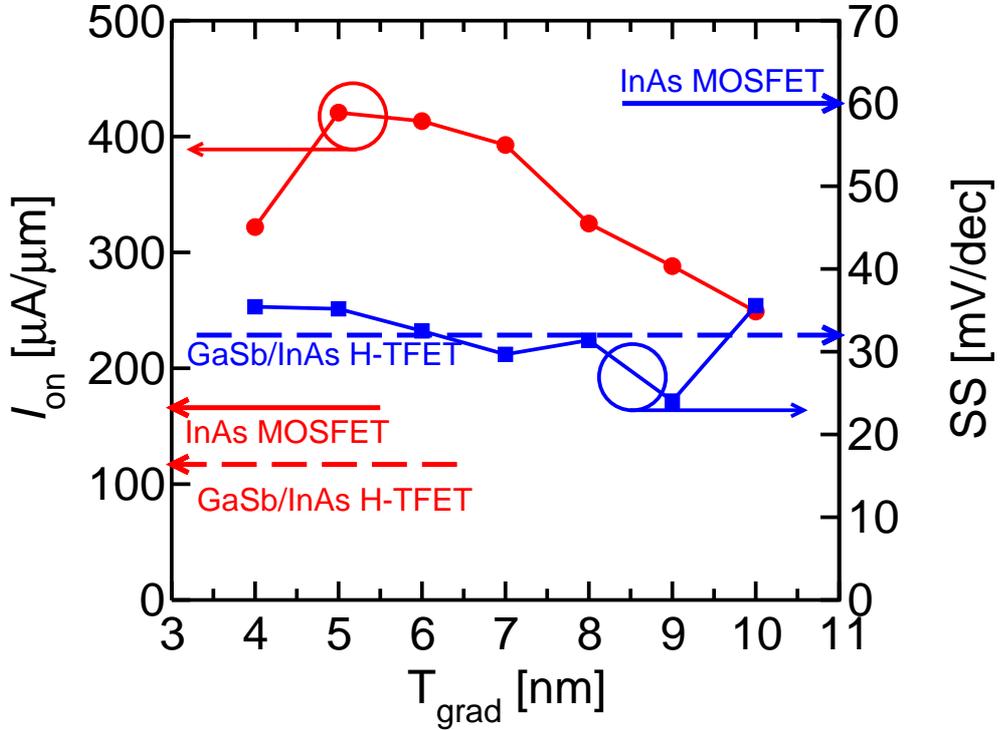


Figure 6.5:  $I_{\text{on}}$  and subthreshold swing at  $V_{DD}=0.3\text{V}$  and versus the thickness  $T_{\text{grad}}$  of the grading region in an  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb-InAs}$  H-TFET.  $I_{\text{on}}$  is maximum for a  $T_{\text{grad}}$  of about 5nm. The  $I_{\text{on}}$  and SS for the GaSb-InAs H-TFET with no grading and the InAs MOSFET are also shown for comparison.

The on current and inverse subthreshold slope SS for different values of  $T_{\text{grad}}$  can be seen in Fig. 6.5, as well as the reference values for the benchmark transistors. There is an optimal value for the grading depth at  $T_{\text{grad}}=5\text{nm}$ , to maximize the on current at over  $400\mu\text{A}/\mu\text{m}$ . On the other hand, the SS value is fairly independent of  $T_{\text{grad}}$ . This result suggests that the use of the molar fraction grading in the source region of H-TFETs is preferable to the use of mechanical strain, which usually degrades the sub- $V_T$  slope. On the other hand, we remark that the use of the molar fraction grading has been already proved to be a feasible design option in InP-based hetero-junction bipolar transistors [102, 103], which is an encouraging precedent for its possible exploitation in hetero-junction Tunnel

FETs.

### 6.2.2 Physical interpretation

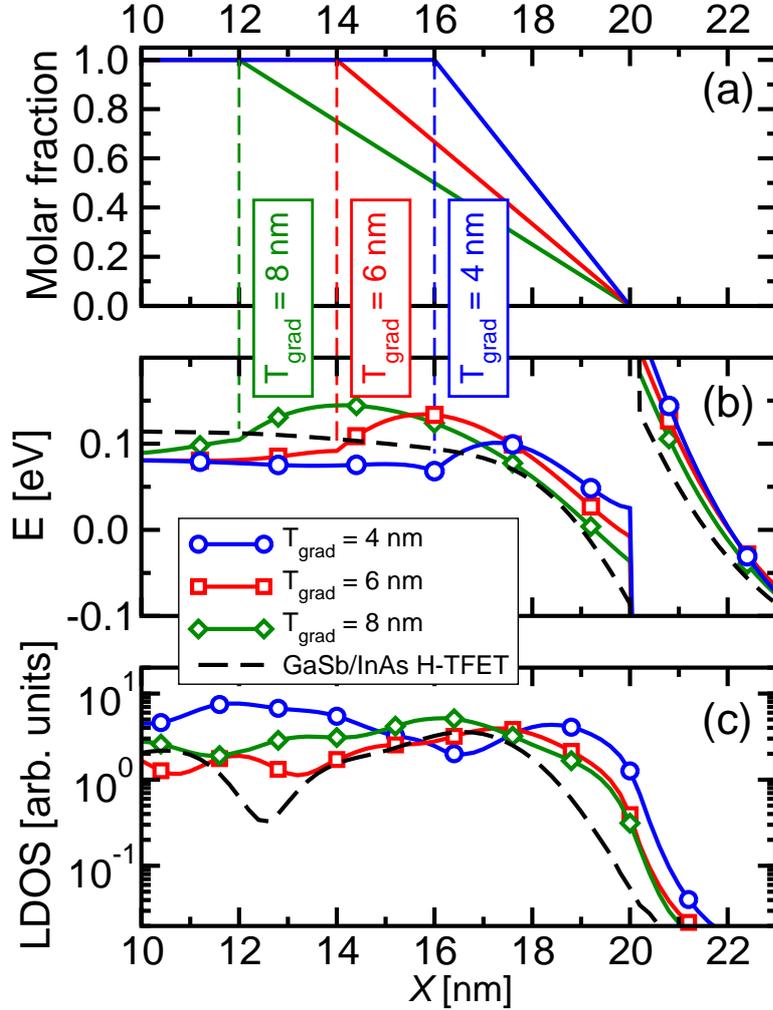


Figure 6.6: Molar fraction profile (a), corresponding subband profile (b), and valence band local density of states (c) in the region close to the source-channel junction for  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}/\text{InAs}$  Tunnel-FETs with different  $T_{\text{grad}}$  values. For each  $T_{\text{grad}}$  the LDOS is plotted at the energy for maximum transmission identified in Fig. 6.8. The source-channel junction is at  $x = 20$  nm and the source Fermi level is set to  $E_{F_s} = 0$  eV. A GaSb/InAs H-TFET with no grading is shown for comparison. The biases are  $V_{GS} = V_{DS} = V_{DD} = 0.3$  V.

Beyond the intuitive principle exposed in section 6.2, we aimed to analyze further the reasons behind the on current increase in the graded source devices, and the physical explanation for the existence of a  $T_{\text{grad}}$  optimum at  $T_{\text{grad}}=5\text{nm}$ . In order to do so, Fig. 6.6 plots the subband profiles and local density of states (LDOS) for three select grading profiles, and a comparison to the GaSb-InAs H-TFET.

Fig. 6.6(b) does indeed show a flattening of the valence band when compared to the non-graded H-TFET, as expected from the principle illustrated in Fig. 6.3. It should be noted that the energy gap at the source-channel junction is a constant ( $\sim 0.22\text{eV}$ ) which does not depend on the source grading configuration. It is only a material parameter which is determined entirely by the semiconductors juxtaposition at the source-channel interface. For all grading configurations, the junction is always a GaSb-InAs junction.

As a consequence, straightening the band profile via source molar fraction grading has the undesirable drawback of displacing the conduction band edge further into the channel, lengthening the tunneling distance. Thankfully this aftereffect magnitude is not enough to counteract the tunneling length reduction brought on by the straightening of the valence band.

Fig. 6.6(c) shows the local density of state for valence carriers, at the energy  $E_{\text{max}}$  which corresponds to the maximum carrier transmission. In graded devices the LDOS penetrates deeper into the channel region, whereas in the GaSb-InAs TFET the valence carriers are almost entirely constrained in the source region. The source grading and the induced triangular potential well confine the carriers against the source-channel junction.

We can take a look at the whole LDOS picture for a better understanding of the carriers confinement phenomenon. Fig. 6.7 shows the local density of states over the whole energy spectrum for three graded devices and for the GaSb-InAs TFET. The confinement of the carriers against the source-channel junction induced by the grading is made apparent. It is also noticeable that the shorter  $T_{\text{grad}}$  configuration (at  $T_{\text{grad}}=3\text{nm}$ ) has the undesirable effect of driving the carrier density downward in terms of energy, thereby depopulating the energy region relevant to tunneling. This contributes to the decrease in  $I_{\text{on}}$  for short values of  $T_{\text{grad}}$  seen in Fig. 6.5.

To fully extend our understanding of the current improvement, we include

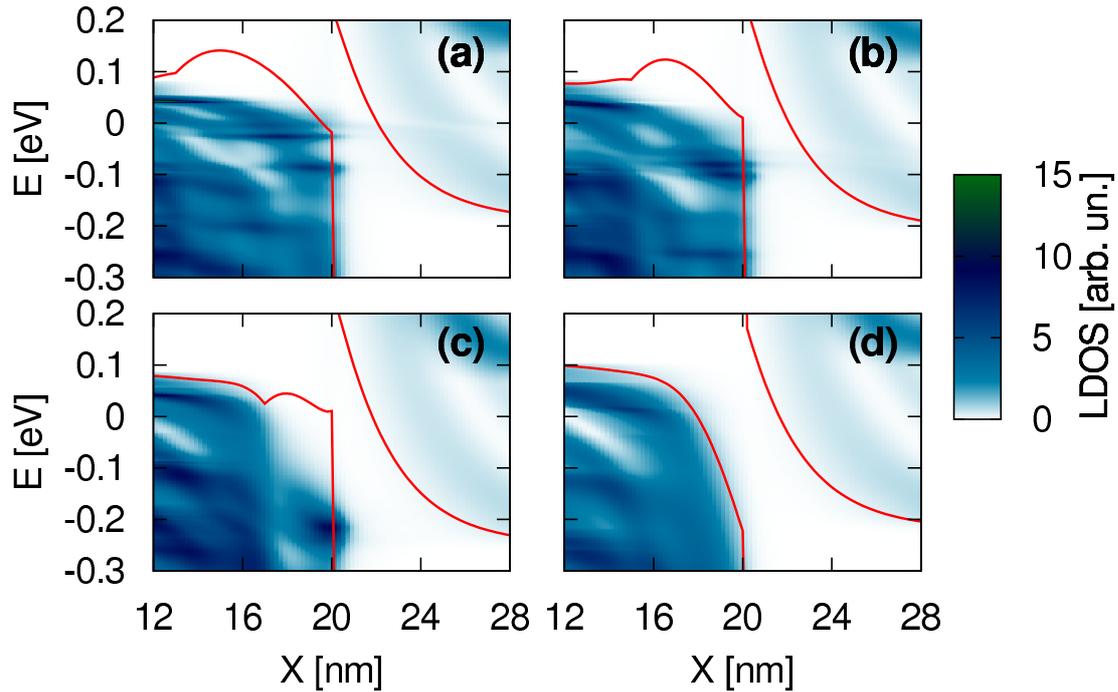


Figure 6.7: Local density of states for all carriers in the region close to the source-channel junction for  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb-InAs}$  Tunnel-FETs with  $T_{\text{grad}}=7\text{nm}$  (a),  $T_{\text{grad}}=5\text{nm}$  (b),  $T_{\text{grad}}=3\text{nm}$  (c) and for the GaSb-InAs H-TFET (d). The corresponding subband edge profiles are shown in superimposition in red. The source-channel junction is at  $x = 20\text{nm}$  and the source Fermi level is set to  $E_{F_s} = 0\text{eV}$ . The biases are  $V_{GS} = V_{DS} = V_{DD}=0.3\text{V}$ .

Fig. 6.8, which shows the the distribution of the carriers transmission and of the current over the energy spectrum, as given by the simulations. As can be expected, the  $x_M$  grading introduces a very significant increase of the transmission spectrum amplitude, albeit with a reduction of the energy range bandwidth: whereas the GaSb-InAs H-TFET shows a transmission spread over the whole  $-0.2\text{eV} < E < 0.1\text{eV}$  energy range, the graded source devices have a distinct transmission peak. What is more interesting is that both the energy and the amplitude of that transmission peak increases with  $T_{\text{grad}}$ , because of the influence of the  $T_{\text{grad}}$  over the general subband profile and local density of state as shown in Figs.6.6 and 6.7. For  $T_{\text{grad}}$  values over 6nm, the transmission peak energy displaces to-

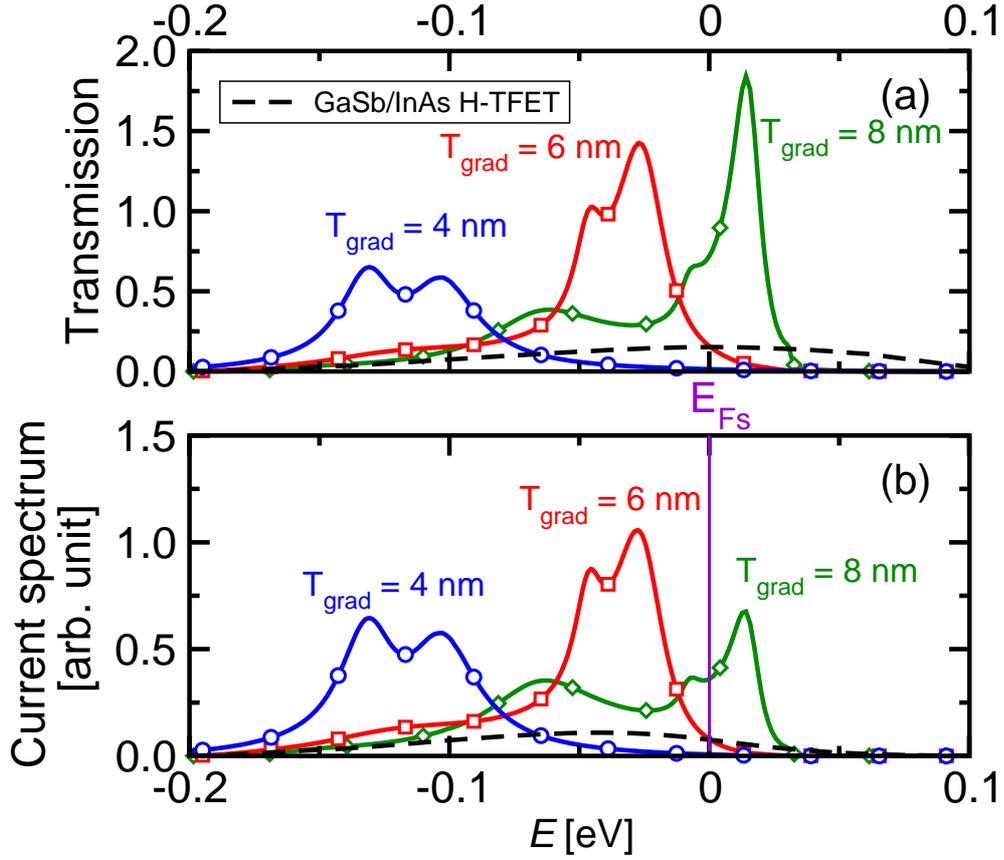


Figure 6.8: Transmission (a) and current density (b) versus energy for the  $\text{Al}_{x_M}\text{Ga}_{(1-x_M)}\text{Sb}/\text{InAs}$  Tunnel-FETs with different  $T_{\text{grad}}$  analyzed also in Fig. 6.6. A GaSb/InAs H-TFET without grading (dashed line) is shown for comparison. The biases are  $V_{GS} = V_{DS} = V_{DD} = 0.3\text{V}$ . The source Fermi level  $E_{F_s} = 0\text{eV}$  is also marked.

ward higher and higher energies that are less and less populated, because they are located above the source Fermi level  $E_{F_s} = 0\text{eV}$ . In other words, the current spectrum is proportional to the transmission spectrum multiplied by the Fermi-Dirac occupation function  $f_{\text{FD}}(E - E_{F_s}) = \{1 + \exp[(E - E_{F_s})/k_B T]\}^{-1}$  at the source contact, so that the final  $I_{\text{on}}$  is the result of a tradeoff between the overall amplitude of the transmission spectrum  $T(E)$ , and the occupation of states at the energy range corresponding to the transmission maximum.

The combination of those two factors explains the existence of the  $I_{\text{on}}$  optimum for  $T_{\text{grad}} = 5\text{nm}$  in Fig. 6.5. For larger molar fraction grading depth  $T_{\text{grad}}$ , the

transmission peak is unexploited because of the low states occupancy at the transmission peak. For shorter  $T_{\text{grad}}$ , the amplitude of the transmission peak itself is lower, because of the lower amount of available states at the source-channel junction.

### 6.3 Insertion of a quantum well

In this Section we study another alternative and original configuration able to provide a large on-current enhancement in H-TFETs [104].

The main technological problem of the device proposed in Sec. 6.2 is probably due to the difficulty in realizing a very steep molar grading of the AlGaSb layer going from  $x_M = 1$  to  $x_M = 0$  within a thin region of a few nanometers (see Fig. 6.3).

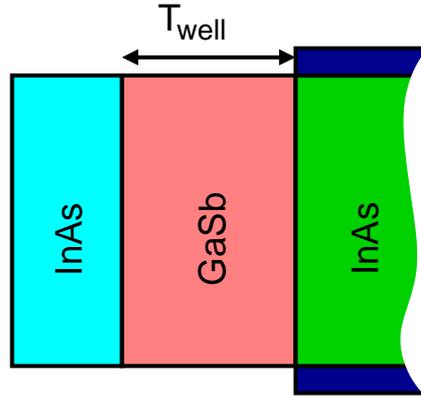


Figure 6.9: Sketch of the source region of the InAs-GaSb-InAs-QW TFET. Transport direction  $x$  is assumed to be  $[100]$ .

The idea here is to create a strong concentration of electronic density of states at the edge between the source and the channel region with a simpler approach. Such a DOS concentration is expected to help the penetration of the electronic wave function in the channel region and hence to boost the band-to-band current. This condition can be obtained by the presence of a thin quantum well (QW) for the holes at the end of the source region by inserting a few atomic layers of GaSb between two InAs regions. The technological realization of two abrupt GaSb/InAs junctions should be much easier than the aggressive grading proposed in Sec. 6.2.

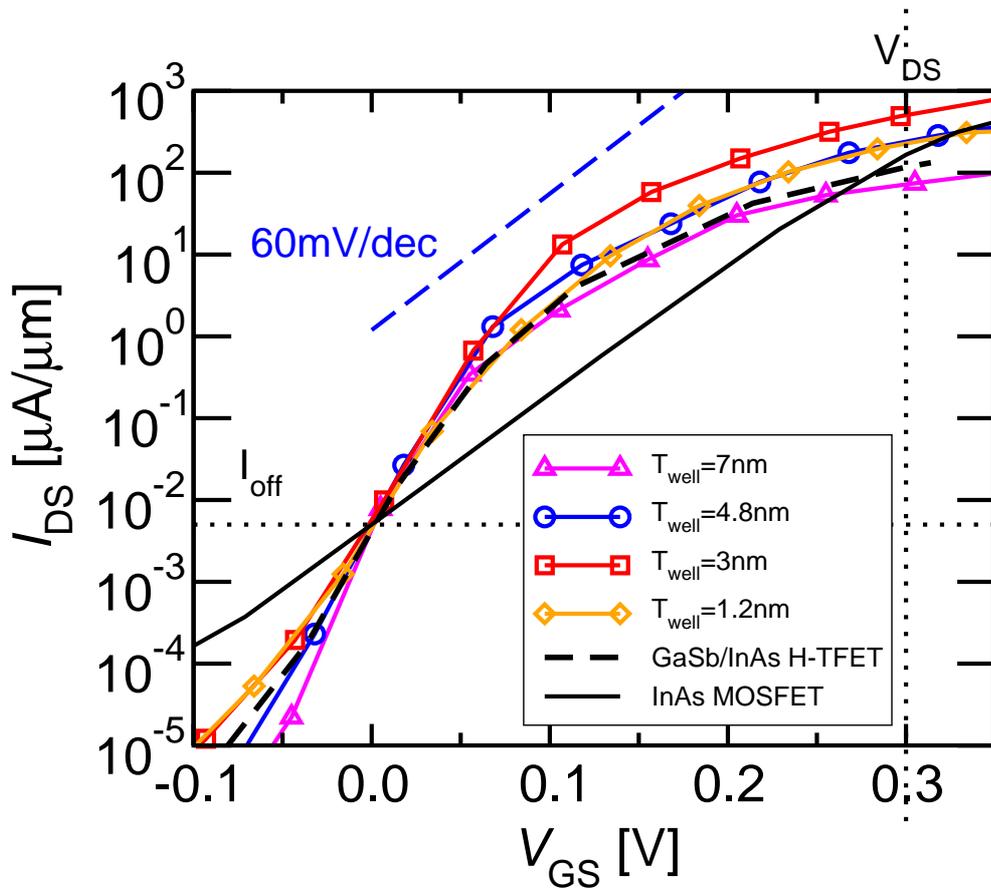


Figure 6.10:  $I_{DS}$  versus  $V_{GS}$  curves at  $V_{DS}=0.3V$  for 4 different values of the well thickness  $T_{well}$  of the QW-TFE and at fixed  $I_{off}=5nA/\mu m$ . The GaSb-InAs H-TFET with no grading and the InAs MOSFET are also shown for comparison.

The proposed device is sketched in Fig. 6.9. The source is composed by a first p-doped InAs region and a by a thin GaSb undoped layer of thickness  $T_{well}$ , while the channel and the drain are composed by undoped InAs and weakly n-doped InAs, respectively. In this study we have considered quantum wells of different thickness varying from 1.2 to 7 nm and investigated the efficiency in improving the I-V characteristics of the GaSb-InAs H-TFET.

Fig. 6.10 shows the transfer characteristics of the InAs-GaSb-InAs QW-TFET for different values of the well thickness  $T_{well}$ . The I-V curves of the homo-junction InAs MOSFET and of the GaSb-InAs H-TFET are also shown for comparison. The gate work-function were set to have the same  $I_{off}=5nA/\mu m$  for all the transis-

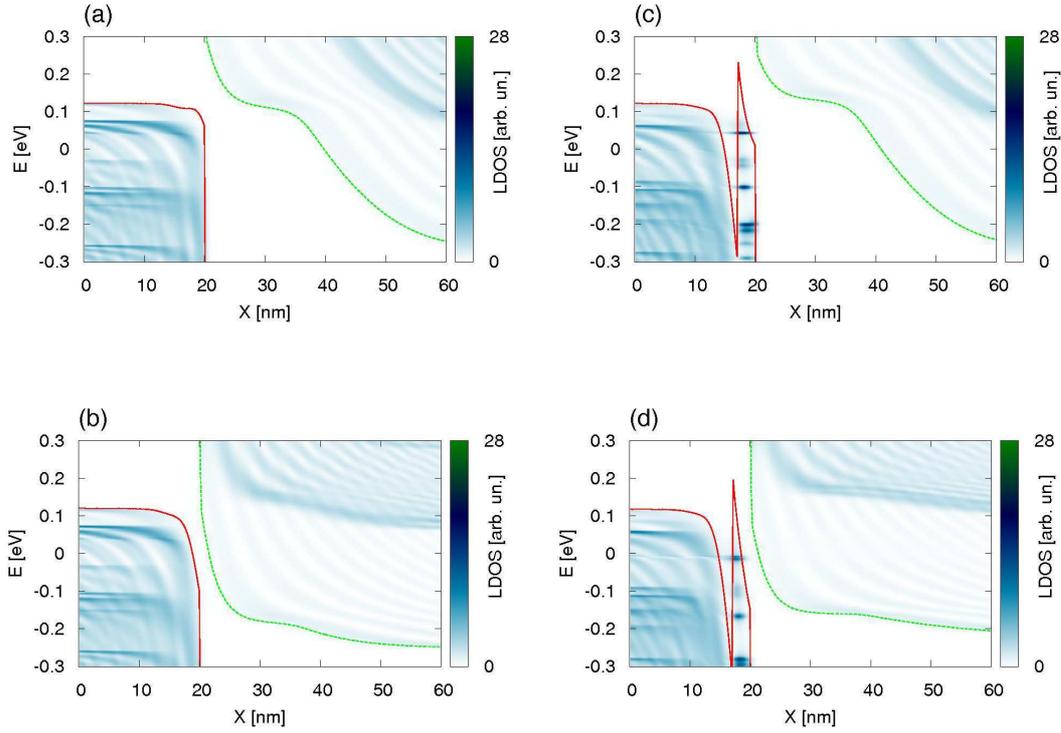


Figure 6.11: Spatial profiles for the lowest conduction and of the highest valence sub-band and local-density-of-states energy spectra. (a),(c) correspond to the  $I_{off}$  bias condition (i.e.  $V_{GS}=0V$  and  $V_{DS} = V_{DD} =0.3V$ ), while (b),(d) are for the  $I_{on}$  bias condition (i.e.  $V_{GS} = V_{DS} = V_{DD} =0.3V$ ). The figure compares the results for the GaSb-InAs H-TFET in (a),(b) with those for the InAs-GaSb-InAs QW-TFET with  $T_{well} =3nm$  in (c),(d).

tors. The device parameters are the same used in Sec. 6.2.1. Similarly to the case of the H-TFET with molar fraction grading, the InAs-GaSb-InAs QW-TFET can outperform the H-TFET and the InAs MOSFET in terms of  $I_{on}$  for different values of  $T_{well}$ , still preserving a very steep inverse  $-V_T$  slope. According to Fig. 6.10, the maximal  $I_{on}$  value is obtained for  $T_{well} = 3nm$ , which corresponds to 5 atomic layers of GaSb (the lattice parameter of GaSb is around 0.6nm), while even the devices with wells of  $T_{well} = 1.2nm$  (2 atomic layers) and with  $T_{well} = 4.8nm$  (8 atomic layers) still present better I-V curves than the MOSFET.

### 6.3.1 Role of the quantum well

In order to deeply investigate the physical mechanisms responsible of such current improvement, we have analyzed microscopic quantities such as the energy spectra of LDOS (see Fig. 6.11) and of the current density (see Fig. 6.12) along the transport direction for both the H-TFET and the QW-TFET with  $T_{\text{well}} = 3\text{nm}$ . These quantities were plotted together with the spatial profiles of the lowest conduction subband and the highest valence subband.

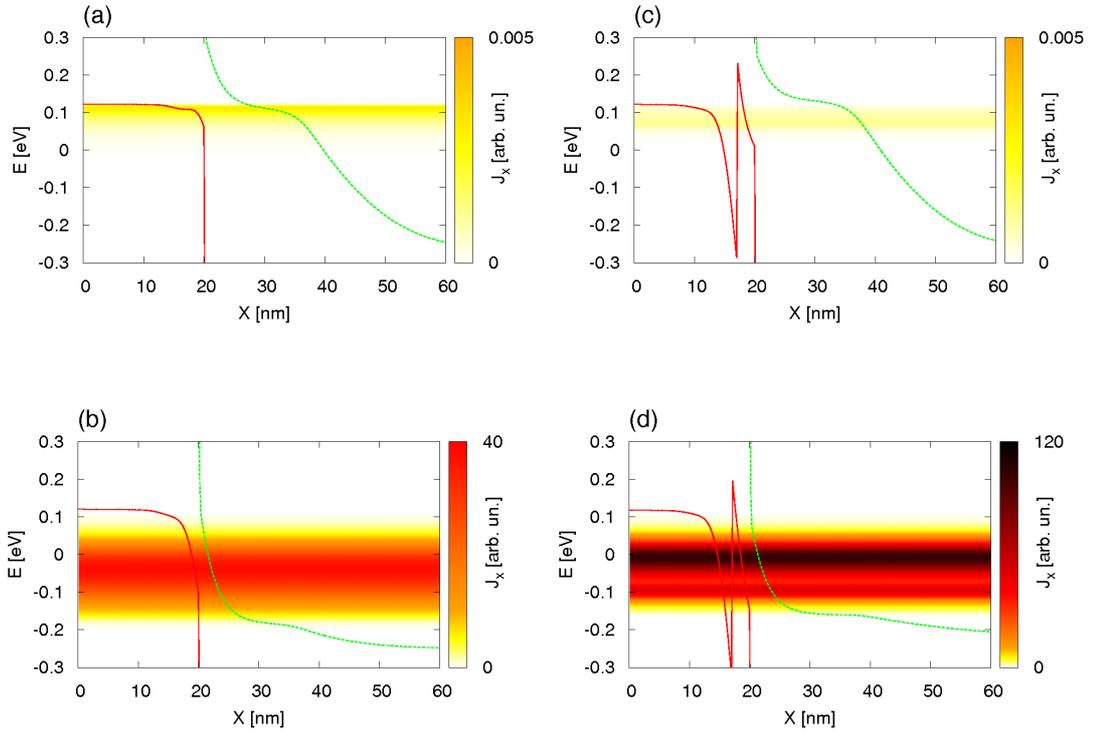


Figure 6.12: Spatial profiles for the lowest conduction and of the highest valence subband and current density energy spectra. (a),(c) correspond to the  $I_{off}$  bias condition (i.e.  $V_{GS}=0\text{V}$  and  $V_{DS} = V_{DD} = 0.3\text{V}$ ), while (b),(d) are for the  $I_{on}$  bias condition (i.e.  $V_{GS} = V_{DS} = V_{DD} = 0.3\text{V}$ ). The figure compares the results for the GaSb-InAs H-TFET in (a),(b) with those for the InAs-GaSb-InAs QW-TFET with  $T_{\text{well}} = 3\text{nm}$  in (c),(d).

First, a clear increase of the DOS concentration is visible in the QW-TFET (in Figs. 6.11(c),(d)) as an effect of the quantum well with respect to the GaSb-InAs

H-TFET (in Figs. 6.11(a),(b)). As shown in Figs. 6.12(a),(c), when the device is in the OFF state ( $V_{GS} = 0V$ ;  $V_{DS} = V_{DD} = 0.3V$ ), the energies important for the band-to-band tunneling are just below the top of the valence subband in the source, says between 0 and 0.1 eV. In this case the DOS concentration of the quantum well is energetically located below this energy window and it is therefore not active in increasing the current, thus avoiding to deteriorate the sub- $V_T$  slope.

In the ON state ( $V_{GS} = V_{DS} = V_{DD} = 0.3V$ ), however, the change in the LDOS induced by the quantum well (shown in Figs. 6.12(b),(d)) has a direct effect on the current. In this case the DOS concentration in the QW is energetically located in the energy window important for the transport, below  $E_{Fs} = 0eV$ , and actively participate to the tunneling process. In fact, the QW counteracts the repulsion of the hole LDOS at the source-channel junction when  $V_{GS}$  increases, hence boosting the probability of electron band-to-band tunneling [104].

## 6.4 Conclusion

This chapter has presented the concepts of two novel architectures for hetero-junction Tunnel FETs, the first based of a grading of the molar fraction of a ternary alloy in the source, specifically AlGaSb, the second based on the use of InAs/GaSb/InAs quantum well in the source-channel junction.

In Sec. 6.2, we have exposed how in principle the introduction of a linear  $x_M$  grading in an  $Al_{x_M}Ga_{(1-x_M)}Sb$  source could result in a straightening of the valence band and thus an increase of the on current in a AlGaSb-InAs H-TFET. In Sec. 6.2.1, we have shown the result of NEGF simulations of the proposed device, and shown the existence of an optimal grading depth  $T_{grad}$  which provides improved performances over the benchmark GaSb-InAs TFET without any degradation of the subthreshold slope. In Sec. 6.2.2, we have provided insight into the actual physical mechanism of the graded H-TFET, and shown that the consequences of the molar fraction grading extend beyond mere band engineering, and how they determine an optimal grading profile. Secs. 6.3 and 6.3.1 have similarly shown how the device with the quantum well in the source can also outperform the InAs MOSFET for  $V_{DD}=300mV$ .

In order to finally compare the performance of all the relevant design options studied in Chapters 5 and 6, in Fig.6.13 we have plotted the on current and the

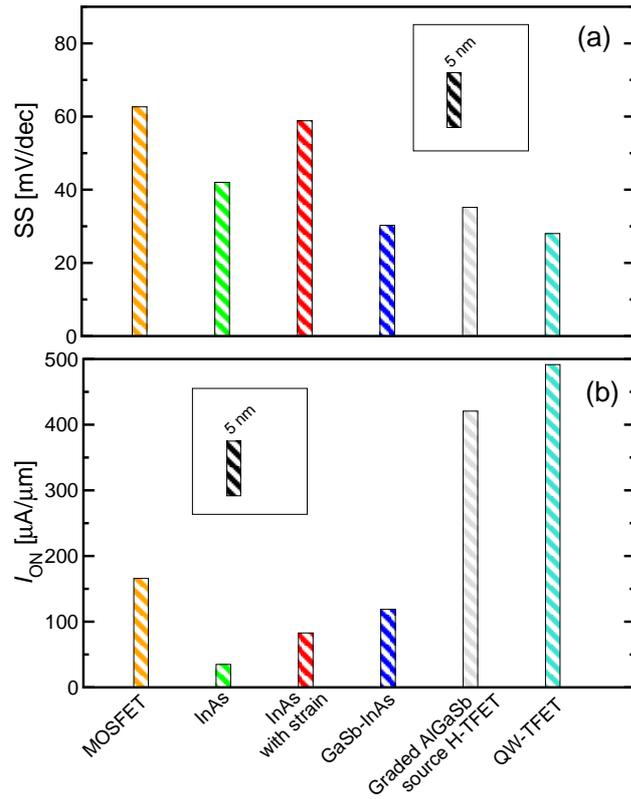


Figure 6.13: (a) Inverse subthreshold slope SS (averaged for  $I_{DS}$  between  $1nA/\mu m$  and  $100nA/\mu m$ ) for the graded source H-TFET (with  $T_{grad}=5nm$ ) and for the quantum well TFET (with  $T_{well}=3nm$ ), as well as for TFET design options and for the InAs MOSFET from chapter 5 for comparison. All devices are at  $D_W=5nm$ . Both devices presented in this chapter show a low SS similar to the H-TFET.  $L_G=17nm$ . (b) On current  $I_{ON}$  at  $I_{off}=5nA/\mu m$ , defined as the  $I_{DS}$  at  $V_{DS}=V_{DD}=0.3V$  and  $V_{GS}=(V_{GS,off}+V_{DD})$ , with  $V_{GS,off}$  being the  $V_{GS}$  giving  $I_{DS}=I_{off}$  [16]. The graded devices greatly outperforms the MOSFET, and the QW-TFET, despite its simplicity, outperforms the graded device.  $L_G=17nm$ .

inverse sub- $V_T$  slopes of these devices.

As can be seen the last two technological design options can present largely improved on current w. r. t. standard devices, while preserving very steep sub-threshold voltage slopes.



# Chapter 7

## Conclusions

The work accomplished during this PhD program consisted in the development and exploitation of nanodevices simulation tools, based on advance band structure and transport calculation techniques.

In Chapter 2, we presented the  $\mathbf{k} \cdot \mathbf{p}$  and AEPM methods used in the rest of the study. The atomistic empirical pseudopotential method, which benefits from the highest physical accuracy, was improved upon in order to be usable for transport simulations for at least a number of realistic configurations. For situations where the AEPM computational burden remains too high, such as the presence of surface roughness, we used the 8-bands  $\mathbf{k} \cdot \mathbf{p}$  method, whose accuracy remains satisfactory after being validated against AEPM results.

In Chapter 3, we presented the non-equilibrium Green's function formalism used for the simulation of quantum transport. A self-consistent Poisson-Schrödinger solver was employed. Phonon scattering was taken into account in the framework of the self-consistent Born approximation.

Chapter 4 exploited the predictive power of the aforementioned transport model to investigate the viability of III-V compounds as channel materials for the most recently introduced CMOS technology, namely FinFETs. We compared the performances of a theoretical Si-based short channel FinFET against its InAs-based counterpart. Supply voltage downscaling being a concern for CMOS technologists, this comparison was taken to low voltage values. Our study has shown that an InAs device takes better advantage of the strong electrostatic control offered by the FinFET architecture, which leads the InAs FinFET to provide the same on-state current as the Si FinFET for 100mV lower supply voltage values.

This confirms the place of III-V compounds as a serious contender for upcoming CMOS technological nodes.

In Chapter 5, we used the same simulation tools to tackle questions concerning a radically different architecture. Although a very promising post-CMOS technology, tunnel field effect transistors have been struggling against their inherent issue regarding low values of on-state current. First, our simulations confirmed that conventional p-i-n TFETs do not perform well enough to constitute an improvement against existing CMOS technology, despite their steep subthreshold slope. Then, we investigated the value of two technology boosters in the form of strain engineering, and hetero-junction TFETs. The results showed that devices using these boosters could compete with standard CMOS technology. These results however had to be tempered by the exposition of the harmful consequences the presence of surface defects have on the slope of tunnel-FETs, which indicates real-life devices might very well still underperform.

In light of this problematic, in Chapter 6, we used the full extent of the versatility of our simulation tools to investigate a radically novel improvement to the TFET architecture. At first, we introduced a grading of the molar fraction in a ternary alloy in the source region of a hetero-junction TFET as a way to perform band engineering. This method exhibited excellent results, making the graded source device a clear improvement over its MOSFET benchmark competitor. Being aware of the difficulty of fabricating such a graded device, and having exposed that the main physical mechanism behind the performance improvement was the confinement of carriers against the source/channel junction, we presented another design exploiting the same mechanism, in the form of a quantum well at the channel interface. This option retains the performance increase of the graded source architecture, while being significantly easier to fabricate. These results should invigorate the status of TFETs as a potential MOSFET successor.

Such encouraging results confirm the importance of full-quantum physical simulations in predicting and exploring the performances of design options for transistors in either the close or far future. Coherent phenomena like electron tunneling cannot be adequately taken into account by semiclassical models, without which charge transport cannot be properly modeled, especially in quantum effect-reliant devices such as TFETs.

On the topic of the author's work during this PhD program, I will discuss

---

some of the more technical and numerical aspects of this thesis. The simulations were performed by improving an existing homemade code for the NEGF-Poisson solver, as well as for the  $\mathbf{k} \cdot \mathbf{p}$  bandstructure calculation. I have developed the code for AEPM bandstructure calculations, which accommodates arbitrary crystal orientation of the channel, arbitrary strain conditions, either unary or binary materials, and surface roughness with appropriate surface passivation. I also wrote a companion code to extract pseudopotential form factors for arbitrary materials using bulk bandstructure characteristics as an input.

The steps taken to improve the computational speed of the AEPM calculations included predicting the nanowire gap energy level from bulk calculations and using a shift-invert method (preceded by a LU factorization of the system's Hamiltonian) to acquire the first conduction and valence bands faster. This was made possible by the fact that the real-space Hamiltonian has a highly sparse matrix representation.

On that topic, the choice of going for a real-space representation of the pseudopotentials rather than the more common reciprocal space representation was motivated by a work of research of optimal calculation parameters which put into light that reciprocal space approaches struggled to describe adequately the confinement conditions on the nanowires edges, the real space approach allowing for a finer spacial tuning of wavefunctions.

The work of actually exploiting the transport code involved a process of tuning simulation parameters and device parameters to obtain physically credible results, and interpreting these results to verify their physical sanity, by way of extracting relevant data and checking their compliance to known physical constraints.

As for perspectives, minor changes to the AEPM code could make it able to calculate bandstructures in ternary alloys, as well as the presence of defects. A drastic step that was developed during this program to improve the AEPM computation speed was to compute the bandstructure for a small number of wavevector values, and using the resulting wavefunctions as a reduced basis for the computation of the other values. This method has shown excellent results for the calculation of conduction bands, but not so much for valence bands. Understanding why could lead to a very fast AEPM implementation.

Finally, these calculations have shown that the same level of spacial accuracy is not needed in all parts of the channel. A finite element-based code instead of a

grid-based one could contribute significantly to the reduction of the computation effort, and is being investigated in our laboratory.

# Bibliography

- [1] H. M. Fahad and M. M. Hussain, “Are nanotube architectures more advantageous than nanowire architectures for field effect transistors?,” *Scientific reports*, vol. 2, 2012.
- [2] P. Y. Yu and M. Cardona, *Fundamentals of semiconductors*. Berlin: Springer-Verlag, 2001.
- [3] E. Fermi, “Sopra lo spostamento per pressione delle righe elevate delle serie spettrali,” *Il Nuovo Cimento (1924-1942)*, vol. 11, no. 3, pp. 157–166, 1934.
- [4] J. R. Chelikowsky and M. L. Cohen, “Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors,” *Physical Review B*, vol. 14, no. 2, p. 556, 1976.
- [5] M. L. Cohen and V. Heine, “The fitting of pseudopotentials to experimental data and their subsequent application,” *Solid state physics*, vol. 24, p. 64, 1970.
- [6] L.-W. Wang, S.-H. Wei, T. Mattila, A. Zunger, I. Vurgaftman, and J. Meyer, “Multiband coupling and electronic structure of  $(\text{InAs})_n/(\text{GaSb})_n$  superlattices,” *Physical Review B*, vol. 60, no. 8, p. 5590, 1999.
- [7] J. R. Chelikowsky and M. L. Cohen, “Self-consistent pseudopotential calculation for the relaxed (110) surface of GaAs,” *Physical Review B*, vol. 20, no. 10, p. 4150, 1979.
- [8] L. W. Wang and A. Zunger, “Electronic structure pseudopotential calculations of large (apprx. 1000 atoms) si quantum dots,” *The Journal of Physical Chemistry*, vol. 98, no. 8, pp. 2158–2165, 1994.

- [9] L.-W. Wang, J. Kim, and A. Zunger, “Electronic structures of [110]-faceted self-assembled pyramidal InAs/GaAs quantum dots,” *Physical Review B*, vol. 59, no. 8, p. 5678, 1999.
- [10] T. B. Bahder, “Eight-band k·p model of strained zinc-blende crystals,” *Phys. Rev. B*, vol. 41, pp. 11992–12001, Jun 1990.
- [11] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, “Band parameters for III-V compound semiconductors and their alloys,” *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001.
- [12] R. G. Veprek, S. Steiger, and B. Witzigmann, “Ellipticity and the spurious solution problem of k·p envelope equations,” *Phys. Rev. B*, vol. 76, p. 165320, Oct 2007.
- [13] M. Shin, “Full-quantum simulation of hole transport and band-to-band tunneling in nanowires using the k·p method,” *Journal of Applied Physics*, vol. 106, no. 5, p. 054505, 2009.
- [14] G. L. Bir, G. E. Pikus, P. Shelnitz, and D. Louvish, *Symmetry and strain-induced effects in semiconductors*, vol. 624. Wiley New York, 1974.
- [15] T. B. Bahder, “Eight-band k p model of strained zinc-blende crystals,” *Physical Review B*, vol. 41, no. 17, p. 11992, 1990.
- [16] F. Conzatti, M. Pala, D. Esseni, E. Bano, and L. Selmi, “Strain-induced performance improvements in InAs nanowire tunnel FETs,” *Electron Devices, IEEE Transactions on*, vol. 59, no. 8, pp. 2085–2092, 2012.
- [17] M. Lundstrom, *Fundamentals of Carrier Transport*. New York: Addison Wesley, 1990.
- [18] D. Esseni, P. Palestri, and L. Selmi, *Nanoscale MOS transistors*. Cambridge: Cambridge University Press, 2011.
- [19] R. Landauer, “Spatial variation of Currents and Fields due to Localized Scatterers in Metallic Conduction,” *Phys. Rev. Lett.*, vol. 1, pp. 233–195, 1957.

- [20] M. Büttiker, “Four terminal phase coherence conductance,” *IBM J. Res. Dev.*, vol. 57, p. 1761, 1986.
- [21] Y. Meir and N. S. Wingreen, “Landauer formula for the current through an interacting electron region,” *Phys. Rev. Lett.*, vol. 68, pp. 2512–2515, Apr 1992.
- [22] A.-P. Jauho, N. S. Wingreen, and Y. Meir, “Time-dependent transport in interacting and noninteracting resonant-tunneling systems,” *Phys. Rev. B*, vol. 50, pp. 5528–5544, Aug 1994.
- [23] D. Ferry and S. Goodnick, *Transport in Nanostructures*. Cambridge: Cambridge University Press, 1997.
- [24] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, 1998.
- [25] G. Mahan, *Many-Particle Physics*. New York: Plenum Press, 1990.
- [26] S. Datta, *Quantum Transport - Atom to Transistor*. United Kingdom: Cambridge University Press, 2005.
- [27] S. Datta, *Quantum Phenomena*. New York: Addison Wesley, 1989.
- [28] S. Poli and M. Pala, “Channel-length dependence of low-field mobility in Silicon-nanowire FETs,” *IEEE Electron Device Letters*, vol. 30, pp. 1212–1214, nov. 2009.
- [29] K. Rogdakis, S. Poli, E. Bano, K. Zekentes, and M. Pala, “Phonon and surface roughness limited mobility of gate-all-around 3C-SiC and Si nanowire FETs,” *Nanotechnology*, vol. 20, no. 29, p. 295202, 2009.
- [30] A. Svizhenko and M. P. Anantram, “Role of scattering in nanotransistors,” *Electron Devices, IEEE Transactions on*, vol. 50, pp. 1459–1466, June 2003.
- [31] M. P. Anantram, M. Lundstrom, and D. Nikonov, “Modeling of Nanoscale Devices,” *Proceedings of the IEEE*, vol. 96, pp. 1511–1550, Sept 2008.

- [32] M. P. L. Sancho, J. M. L. Sancho, and J. Rubio, “Quick iterative scheme for the calculation of transfer matrices: application to Mo (100),” *Journal of Physics F: Metal Physics*, vol. 14, no. 5, p. 1205, 1984.
- [33] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic, “Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches,” *Journal of Applied Physics*, vol. 92, no. 7, pp. 3730–3739, 2002.
- [34] M. Luisier, A. Schenk, and W. Fichtner, “Quantum transport in two- and three-dimensional nanoscale transistors: Coupled mode effects in the nonequilibrium greens function formalism,” *Journal of Applied Physics*, vol. 100, no. 4, pp. –, 2006.
- [35] S. Poli, M. Pala, T. Poiroux, S. Deleonibus, and G. Bacarani, “Size dependence of surface-roughness-limited mobility in Silicon-nanowire FETs,” *IEEE Trans. Electron Devices*, vol. 55, pp. 2968–2976, nov. 2008.
- [36] A. Fetter and J. Walecka, *Quantum Theory of Many-Particle Systems*. New York: Dover, 2003.
- [37] M. Frey, A. Esposito, and A. Schenk, “Simulation of intravalley acoustic phonon scattering in silicon nanowires,” in *Solid-State Device Research Conference, 2008. ESSDERC 2008. 38th European*, pp. 258–261, Sept 2008.
- [38] H. Mera, M. Lannoo, C. Li, N. Cavassilas, and M. Bescond, “Inelastic scattering in nanoscale devices: One-shot current-conserving lowest-order approximation,” *Phys. Rev. B*, vol. 86, p. 161404, Oct 2012.
- [39] H. Mera, M. Lannoo, C. Li, N. Cavassilas, and M. Bescond, “Inelastic scattering in nanoscale devices: One-shot current-conserving lowest-order approximation,” *Phys. Rev. B*, vol. 86, p. 161404, Oct 2012.
- [40] N. Cavassilas, M. Bescond, H. Mera, and M. Lannoo, “One-shot current conserving quantum transport modeling of phonon scattering in n-type double-gate field-effect-transistors,” *Applied Physics Letters*, vol. 102, no. 1, pp. –, 2013.

- [41] M. Lenzi, P. Palestri, E. Gnani, S. Reggiani, A. Gnudi, D. Esseni, L. Selmi, and G. Baccarani, "Investigation of the Transport Properties of Silicon Nanowires Using Deterministic and Monte Carlo Approaches to the Solution of the Boltzmann Transport Equation," *IEEE Trans. Electron Devices*, vol. 55, pp. 2086 – 2096, nov. 2008.
- [42] M. Luisier and G. Klimeck, "Atomistic full-band simulations of silicon nanowire transistors: Effects of electron-phonon scattering," *Phys. Rev. B*, vol. 80, p. 155430, Oct 2009.
- [43] M. V. Fischetti, "Monte Carlo simulation of transport in technologically significant semiconductors of the Diamond and Zinc-Blende structures - Part I: homogeneous transport," *IEEE Trans. Electron Devices*, vol. ED-38, p. 634, 1991.
- [44] E. B. Ramayya, D. Vasileska, S. M. Goodnick, and I. Knezevic, "Electron transport in silicon nanowires: The role of acoustic phonon confinement and surface roughness scattering," *Journal of Applied Physics*, vol. 104, no. 6, pp. –, 2008.
- [45] *International Technology Roadmap for Semiconductors: 2011 Update*. <http://www.itrs.net>, 2011.
- [46] D. J. Frank, "Power-constrained CMOS scaling limits," *IBM Journal of Research and Development*, vol. 46, no. 2.3, pp. 235–244, 2002.
- [47] J. A. del Alamo, "Nanometre-scale electronics with III-V compound semiconductors," *Nature*, vol. 479, pp. 317–323, 2011.
- [48] S. Datta, "III-V field-effect transistors for low power digital logic applications," *Microelectronic Engineering*, vol. 84, no. 910, pp. 2133 – 2137, 2007.
- [49] D. Antoniadis, I. Aberg, C. Ni Chleirigh, O. Nayfeh, A. Khakifirooz, and J. Hoyt, "Continuous MOSFET performance increase with device scaling: The role of strain and channel material innovations," *IBM Journal of Research and Development*, vol. 50, pp. 363–376, July 2006.

- [50] S. Takagi, T. Iisawa, T. Tezuka, T. Numata, S. Nakaharai, N. Hirashita, Y. Moriyama, K. Usuda, E. Toyoda, S. Dissanayake, M. Shichijo, R. Nakane, S. Sugahara, M. Takenaka, and N. Sugiyama, "Carrier-Transport-Enhanced Channel CMOS for Improved Power Consumption and Performance," *Electron Devices, IEEE Transactions on*, vol. 55, pp. 21–39, Jan 2008.
- [51] S. Kim, M. Yokoyama, N. Taoka, R. Nakane, T. Yasuda, O. Ichikawa, N. Fukuhara, M. Hata, M. Takenaka, and S. Takagi, "Enhancement technologies and physical understanding of electron mobility in III–V n-MOSFETs with strain and MOS interface buffer engineering," in *Electron Devices Meeting (IEDM), 2011 IEEE International*, pp. 13.4.1–13.4.4, Dec 2011.
- [52] G. Dewey, B. Chu-Kung, R. Kotlyar, M. Metz, N. Mukherjee, and M. Radosavljevic, "III-V field effect transistors for future ultra-low power applications," in *VLSI Technology (VLSIT), 2012 Symposium on*, pp. 45–46, June 2012.
- [53] S. Kim, M. Yokoyama, N. Taoka, R. Nakane, T. Yasuda, O. Ichikawa, N. Fukuhara, M. Hata, M. Takenaka, and S. Takagi, "Sub-60 nm deeply-scaled channel length extremely-thin body InGaAs-on-insulator MOSFETs on Si with Ni-InGaAs metal S/D and MOS interface buffer engineering," in *VLSI Technology (VLSIT), 2012 Symposium on*, pp. 177–178, June 2012.
- [54] M. Poljak, V. Jovanovic, D. Grgec, and T. Suligoj, "Assessment of Electron Mobility in Ultrathin-Body InGaAs-on-Insulator MOSFETs Using Physics-Based Modeling," *Electron Devices, IEEE Transactions on*, vol. 59, pp. 1636–1643, June 2012.
- [55] D. Lizzit, D. Esseni, P. Palestri, P. Osgnach, and L. Selmi, "Performance Benchmarking and Effective Channel Length for Nanoscale InAs,  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ , and sSi n-MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 61, pp. 2027–2034, June 2014.
- [56] J. A. Del Alamo, "The high-electron mobility transistor at 30: impressive accomplishments and exciting prospects," 2011.

- [57] R. Pillarisetty, “Academic and industry research progress in germanium nanodevices,” *Nature*, vol. 479, no. 7373, pp. 324–328, 2011.
- [58] Y. Wu, R. Wang, T. Shen, J. Gu, and P. Ye, “First experimental demonstration of 100 nm inversion-mode InGaAs FinFET through damage-free sidewall etching,” in *Electron Devices Meeting (IEDM), 2009 IEEE International*, pp. 1–4, IEEE, 2009.
- [59] B. Doyle, S. Datta, M. Doczy, S. Harelend, B. Jin, J. Kavalieros, T. Linton, A. Murthy, R. Rios, and R. Chau, “High performance fully-depleted tri-gate CMOS transistors,” *Electron Device Letters, IEEE*, vol. 24, no. 4, pp. 263–265, 2003.
- [60] S. H. Park, Y. Liu, N. Kharche, M. S. Jelodar, G. Klimeck, M. S. Lundstrom, and M. Luisier, “Performance comparisons of III–V and strained-Si in planar FETs and nonplanar FinFETs at ultrashort gate length (12 nm),” *Electron Devices, IEEE Transactions on*, vol. 59, no. 8, pp. 2107–2114, 2012.
- [61] K. Nehari, N. Cavassilas, J. Autran, M. Bescond, D. Munteanu, and M. Lannoo, “Influence of band structure on electron ballistic transport in silicon nanowire MOSFETs: An atomistic study,” *Solid-State Electronics*, vol. 50, no. 4, pp. 716 – 721, 2006. Special Issue: Papers Selected from the 35th European Solid-State Device Research Conference - ESSDERC05.
- [62] S. Brocard, M. Pala, and D. Esseni, “Design options for hetero-junction tunnel FETs with high on current and steep sub-threshold voltage slope,” in *Electron Devices Meeting (IEDM), 2013 IEEE International*, pp. 5.4.1–5.4.4, Dec 2013.
- [63] J. Rabaey, J. Ammer, T. Karalar, S. Li, B. Otis, M. Sheets, and T. Tuan, “Picoradios for wireless sensor networks: the next challenge in ultra-low power design,” in *Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC. 2002 IEEE International*, vol. 1, pp. 200–201 vol.1, 2002.
- [64] S. Hanson, M. Seok, D. Sylvester, and D. Blaauw, “Nanometer device scaling in subthreshold logic and SRAM,” *Electron Devices, IEEE Transactions on*, vol. 55, pp. 175 –185, jan. 2008.

- [65] A. Wang and A. Chandrakasan, "A 180-mV subthreshold FFT processor using a minimum energy design methodology," *Solid-State Circuits, IEEE Journal of*, vol. 40, pp. 310 – 319, jan. 2005.
- [66] G. Fiori and G. Iannaccone, "Ultralow-Voltage Bilayer Graphene Tunnel FET," *Electron Device Letters, IEEE*, vol. 30, pp. 1096–1098, Oct 2009.
- [67] P. Michetti, M. Cheli, and G. Iannaccone, "Model of tunneling transistors based on graphene on SiC," *Applied Physics Letters*, vol. 96, no. 13, pp. –, 2010.
- [68] Q. Zhang, G. Iannaccone, and G. Fiori, "Two-Dimensional Tunnel Transistors Based on Bi<sub>2</sub>Se<sub>3</sub> Thin Film," *Electron Device Letters, IEEE*, vol. 35, pp. 129–131, Jan 2014.
- [69] A. Seabaugh and Q. Zhang, "Low-Voltage Tunnel Transistors for Beyond CMOS Logic," *Proceedings of the IEEE*, vol. 98, pp. 2095 –2110, dec. 2010.
- [70] A. M. Ionescu and H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches," *Nature*, vol. 479, no. 9, pp. 329–337, 2011.
- [71] H. Lu and A. Seabaugh, "Tunnel Field-Effect Transistors: State-of-the-Art," *Electron Devices Society, IEEE Journal of the*, vol. 2, pp. 44–49, July 2014.
- [72] L. Knoll, Q.-T. Zhao, A. Nichau, S. Trellenkamp, S. Richter, A. Schafer, D. Esseni, L. Selmi, K. Bourdelle, and S. Mantl, "Inverters With Strained Si Nanowire Complementary Tunnel Field-Effect Transistors," *Electron Device Letters, IEEE*, vol. 34, pp. 813–815, June 2013.
- [73] K. Moselund, M. Bjork, H. Schmid, H. Ghoneim, S. Karg, E. Lortscher, W. Riess, and H. Riel, "Silicon Nanowire Tunnel FETs: Low-Temperature Operation and Influence of High- $k$  Gate Dielectric," *Electron Devices, IEEE Transactions on*, vol. 58, pp. 2911–2916, Sept 2011.
- [74] R. Gandhi, Z. Chen, N. Singh, K. Banerjee, and S. Lee, "CMOS-Compatible Vertical-Silicon-Nanowire Gate-All-Around p-Type Tunneling FETs With 50 mV/decade Subthreshold Swing," *Electron Device Letters, IEEE*, vol. 32, pp. 1504–1506, Nov 2011.

- [75] Q. Zhao, J. M. Hartmann, and S. Mantl, “An Improved Si Tunnel Field Effect Transistor With a Buried Strained SiGe Source,” *Electron Device Letters, IEEE*, vol. 32, pp. 1480–1482, Nov 2011.
- [76] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, “Impact of SOI, Si<sub>1-x</sub>Ge<sub>x</sub>OI and GeOI substrates on CMOS compatible Tunnel FET performance,” in *IEEE IEDM Technical Digest*, pp. 1–5, dec. 2008.
- [77] D. Kazazis, P. Jannaty, A. Zaslavsky, C. Le Royer, C. Tabone, L. Clavelier, and S. Cristoloveanu, “Tunneling field-effect transistor with epitaxial junction in thin germanium-on-insulator,” *Applied Physics Letters*, vol. 94, no. 26, pp. –, 2009.
- [78] Y. Yang, S. Su, P. Guo, W. Wang, X. Gong, L. Wang, K. L. Low, G. Zhang, C. Xue, B. Cheng, G. Han, and Y.-C. Yeo, “Towards direct band-to-band tunneling in P-channel tunneling field effect transistor (TFET): Technology enablement by Germanium-tin (GeSn),” in *Electron Devices Meeting, 2012 IEEE International*, pp. 16.3.1–16.3.4, Dec 2012.
- [79] S. Wirths, A. T. Tiedemann, Z. Ikonc, P. Harrison, B. Hollnder, T. Stoica, G. Mussler, M. Myronov, J. M. Hartmann, D. Grtzmacher, D. Buca, and S. Mantl, “Band engineering and growth of tensile strained Ge/(Si)GeSn heterostructures for tunnel field effect transistors,” *Applied Physics Letters*, vol. 102, no. 19, pp. –, 2013.
- [80] F. Conzatti, M. Pala, D. Esseni, E. Bano, and L. Selmi, “A simulation study of strain induced performance enhancements in InAs nanowire Tunnel-FETs,” in *IEEE IEDM Technical Digest*, pp. 5.2.1–5.2.4, 2011.
- [81] J. Knoch and J. Appenzeller, “Modeling of high-performance p-type III-V heterojunction tunnel FETs,” *IEEE Electron Device Letters*, vol. 31, pp. 305–307, april 2010.
- [82] G. Dewey, B. Chu-Kung, J. Boardman, J. M. Fastenau, J. Kavalieros, R. Kotlyar, W. K. Liu, D. Lubyshev, M. Metz, N. Mukherjee, P. Oakey, R. Pillarisetty, M. Radosavljevic, H. W. Then, and R. Chau, “Fabrication,

- characterization, and physics of III–V heterojunction tunneling Field Effect Transistors (H-TFET) for steep sub-threshold swing,” in *IEEE IEDM Technical Digest*, pp. 33.6.1–33.6.4, 2011.
- [83] G.Zhou, R.Li, T.Vasen, M.Qi, S.Chae, Y.Lu, Q.Zhang, H.Zhu, J.-M.Kuo, T.Kosel, M.Wistey, P.Fay, A. Seabaugh, and H.(Grace) Xing, “Novel gate-recessed vertical InAs/GaSb TFETs with record high  $I_{ON}$  of  $180\mu\text{A}/\mu\text{m}$  at  $V_{DS}=0.5\text{V}$ ,” in *IEEE IEDM Technical Digest*, pp. 777–780, 2012.
- [84] R. Li, Y.Lu, G.Zhou, Q.Liu, S.D. Chae, T.Vasen, W.S.Hwang, Q.Zhang, P.Fay, T.Kosel, M. Wistey, H.Xing, and A.Seabaugh, “AlGaSb/InAs Tunnel Field-Effect Transistor With On-Current of  $78\mu\text{A}/\mu\text{m}$  at  $0.5\text{V}$ ,” *IEEE Electron Device Letters*, vol. 33, no. 3, pp. 363–365, 2012.
- [85] G.Zhou, R.Li, T.Vasen, M.Qi, S.Chae, Y.Lu, Q.Zhang, H.Zhu, J.-M.Kuo, T.Kosel, M.Wistey, P.Fay, A. Seabaugh, and H.(Grace) Xing, “Novel gate-recessed vertical InAs/GaSb TFETs with record high  $I_{ON}$  of  $180\mu\text{A}/\mu\text{m}$  at  $V_{DS}=0.5\text{V}$ ,” in *IEEE IEDM Technical Digest*, pp. 777–780, 2012.
- [86] D. Mohata, B. Rajamohanam, T. Mayer, M. Hudait, J. Fastenau, D. Lubyshch, A. Liu, and S. Datta, “Barrier-engineered arsenide-antimonide heterojunction tunnel fetts with enhanced drive current,” *Electron Device Letters, IEEE*, vol. 33, no. 11, pp. 1568–1570, 2012.
- [87] A. Dey, B. Borg, B. Ganjipour, M. Ek, K. Dick, E. Lind, C. Thelander, and L. Wernersson, “High-Current GaSb/InAs(Sb) Nanowire Tunnel Field-Effect Transistors,” *Electron Device Letters, IEEE*, vol. 34, no. 2, pp. 211–213, 2013.
- [88] M. Noguchi, S. Kim, M. Yokoyama, S. Ji, O. Ichikawa, T. Osada, M. Hata, M. Takenaka, and S. Takagi, “High ion/ioff and low subthreshold slope planar-type ingaas tunnel fetts with zn-diffused source junctions,” in *Electron Devices Meeting (IEDM), 2013 IEEE International*, pp. 28.1.1–28.1.4, Dec 2013.
- [89] M. Pala, D. Esseni, and F. Conzatti, “Impact of interface traps on the IV curves of InAs Tunnel-FETs and MOSFETs: a full quantum study,” in *IEEE IEDM Technical Digest*, pp. 6.6.1–6.6.4, 2012.

- [90] D. Mohata, S. Mookerjee, A. Agrawal, Y. Li, T. Mayer, V. Narayanan, A. Liu, D. Loubychev, J. Fastenau, and S. Datta, “Experimental Staggered-Source and N+ Pocket-Doped Channel III–V Tunnel Field-Effect Transistors and Their Scalabilities,” *Applied Physics Express*, vol. 4, no. 2, p. 024105, 2011.
- [91] L. Liu, D. Mohata, and S. Datta, “Scaling Length Theory of Double-Gate Interband Tunnel Field-Effect Transistors,” *Electron Devices, IEEE Transactions on*, vol. 59, no. 4, pp. 902–908, 2012.
- [92] F. Conzatti, M. Pala, D. Esseni, and E. Bano, “Investigation of localized versus uniform strain as a performance booster in InAs Tunnel-FETs,” *Solid-State Electronics*, vol. 88, no. 0, pp. 49 – 53, 2013. Selected extended papers from {ULIS} 2012 conference.
- [93] A. Martinez, M. Bescond, J. Barker, A. Svizhenko, M. P. Anantram, C. Millar, and A. Asenov, “A Self-Consistent Full 3-D Real-Space NEGF Simulator for Studying Nonperturbative Effects in Nano-MOSFETs,” *Electron Devices, IEEE Transactions on*, vol. 54, pp. 2213–2222, Sept 2007.
- [94] C. Buran, M. Pala, M. Bescond, M. Dubois, and M. Mouis, “Three-Dimensional Real-Space Simulation of Surface Roughness in Silicon Nanowire FETs,” *Electron Devices, IEEE Transactions on*, vol. 56, pp. 2186–2192, Oct 2009.
- [95] F. Conzatti, M. Pala, and D. Esseni, “Surface-roughness-induced variability in nanowire InAs tunnel FETs,” *IEEE Electron Device Letters*, vol. 33, no. 6, pp. 806–808, 2012.
- [96] M. Pala and D. Esseni, “Interface Traps in InAs Nanowire Tunnel-FETs and MOSFETs – Part I: Model Description and Single Trap Analysis in Tunnel-FETs,” *Electron Devices, IEEE Transactions on*, vol. 60, no. 9, pp. 2795–2801, 2013.
- [97] D. Esseni and M. Pala, “Interface Traps in InAs Nanowire Tunnel FETs and MOSFETs – Part II: Comparative Analysis and Trap-Induced Variability,” *Electron Devices, IEEE Transactions on*, vol. 60, no. 9, pp. 2802–2807, 2013.

- 
- [98] M. Passlack, R. Droopad, and G. Brammertz, "Suitability study of oxide/gallium arsenide interfaces for mosfet applications," *Electron Devices, IEEE Transactions on*, vol. 57, pp. 2944–2956, nov. 2010.
- [99] L. Lin and J. Robertson, "Defect states at III-V semiconductor oxide interfaces," *Applied Physics Letters*, vol. 98, no. 8, p. 082903, 2011.
- [100] L. De Michielis, L. Lattanzio, K. Moselund, H. Riel, and A. Ionescu, "Tunneling and Occupancy Probabilities: How Do They Affect Tunnel-FET Behavior?," *Electron Device Letters, IEEE*, vol. 34, pp. 726–728, June 2013.
- [101] S. Brocard, M. Pala, and D. Esseni, "Large On-Current Enhancement in Hetero-Junction Tunnel-FETs via Molar Fraction Grading," *Electron Device Letters, IEEE*, vol. 35, pp. 184–186, Feb 2014.
- [102] M.D.Dvorak, C.R.Bolognesi, O.J.Pitts, and S.P.Watkins, "300 GHz InP/GaAsSb/InP double HBTs with high current capability and BV/sub CEO > 6 V," *IEEE Electron Device Letters*, vol. EDL-22, pp. 361–363, 2001.
- [103] M.J.W.Rodwell, M.Le, and B. Brar, "InP Bipolar ICs: Scaling Roadmaps, Frequency Limits, Manufacturable Technologies," *IEEE Proceedings*, vol. 96, pp. 271–286, 2008.
- [104] M. Pala, S. Brocard and D. Esseni, "Exploiting hetero-junctions to improve the performance of III-V Tunnel-FETs," To be published in *J. Elec. Dev. Soc.*, 2014.