STOCHASTIC MODELS OF SURFACE LIMITED ELECTRONIC AND
HEAT TRANSPORT IN METAL AND SEMICONDUCTOR
CONTACTS, WIRES, AND SHEETS—
MICRO TO NANO

BY

PIERRE NICOLAS MARTIN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

      Professor Umberto Ravaioli, Chair
      Professor Andreas C. Cangellaris
      Professor Narayana R. Aluru
      Professor Kent D. Choquette
      Assistant Professor Eric Pop

ABSTRACT

We introduce novel statistical simulation approaches to include the effect of surface roughness in coupled mechanical, electronic and thermal processes in N/MEMS and semiconductor devices in the 10 nm - 1 μm range. A model is presented to estimate roughness rms Δ and autocorrelation L from experimental surfaces and edges, and subsequently generate statistical series of rough geometrical devices from these observable parameters. Using such series of rough electrodes under Holm's theory, we present a novel simulation framework which predicts a contact resistance of 80 mΩ in MEMS gold-gold micro-contacts, for applied pressures above 0.3 mN on 1 μm $\times$ 1 μm surfaces. The non-contacting state of such devices is simulated through statistical Monte Carlo iterations on percolative networks to derive a time to electro-thermal failure through electrical discharges in the gas insulating metal electrodes. The observable parameters L and Δ are further integrated in semi-classical solutions to the electronic and thermal Boltzman transport equation (BTE), and we show roughness limited heat and electronic transport in rough semiconductor nanowires and nano-ribbons. In this scope, we model for the first time electrostatically confined nanowires, where a reduction of electron - surface scattering leads to enhanced mobility in comparison to geometrical nanowires. In addition, we show extremely low thermal conductivity in Si, GaAs, and Ge nanowires down to 0.1 W/m/K for thin Ge wires with 56 nm width and $\Delta = 3$ nm. The dependency of thermal conductivity in $(D/\Delta)^2$ leads to possible application in the field of thermoelectric devices. For rough channels of width below 10 nm, electronic transport is additionally modeled using a novel non-parabolic 3D recursive Green function scheme, leading to an estimation of reduced electronic transmission in rough semiconductor wires based on the quantum nature of charge carriers. Electronic and thermal simulation schemes are finally extended to such 2D semiconductor materials as graphene, where low thermal conductivity is approximated below 1000 W/m/K for rough suspended graphene ribbons in accordance with recent experiments.

"The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It is not an attempt to violate any laws; it is something, in principle, that can be done; but in practice ... it has not been done because WE *are too big."*
–R. P. Feynman, 1986 Dirac Memorial Lectures.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

As electronic devices reached the nanometer scale, the semiconductor industry entered an era of new challenges where technology computer aided design (TCAD) has an important role to play. In 1965, Moore predicted that the number of transistors integrated on a chip would double every eighteen months. Today, while the dimension of electronic devices reaches atomistic scales, the idea of following this trend by simply reducing the size of conventional complementary metal oxide semiconductor (CMOS) technologies has been abandoned. Besides, the field of electronics has evolved toward a range of applications which is now much broader than the initial logic and arithmetical operations it was first intended to address. New electronic devices are integrated today in biological systems, allow mechanical, optical and chemical sensing or actuating at the nanometer scale, and actively contribute to pushing our scientific perception further in the twenty-first century. Many of these new devices now rely on breakthrough technologies such as nano-electromechanical systems (NEMS), which constitute the leading edge in mastering the coupled mechanical, electrical and thermal behavior of materials at the quasi-atomistic scale. In addition, a large class of these new technologies utilizes novel materials, such as graphene or carbon nanotubes, which make full profit of the quantized behavior of energy transport processes in low-dimensional systems. At the forefront of size reduction, graphene constitutes the most advanced state of the art, where the dimension of the device is effectively reduced to a single atomic layer. In that respect, the impact of the appearance of such devices in the laboratory drives theoretical and experimental exploration to new fronts, where the quantum nature of electronic and heat conduction, as well as the atomistic nature of material deformation, is pivotal in understanding the physics of new devices.

As an example, the complexity of the multi-scale and multi-physics nature of low dimension devices is well illustrated by the mere scenario of conduction in a real semiconductor wire of a few nanometers in diameter. At this

scale, the wave nature of charge carriers must be taken into account in order to capture the quantization of allowed energies resulting from electronic confinement within the nanowire (NW) structure. As electronic waves propagate through the imperfect wire, roughness patterns at its surface interfere with these waves in non-negligible proportions, thus hindering conduction. Similarly, electronic waves exchange energy with the atomic crystal lattice and generate heat waves. Subsequently, as thermal conduction occurs, heat waves will also interfere with imperfections of the wire surface. Hence, the latter problem which, at the macro scale, would be mostly captured by Kirchhoff's and Joule's equations, requires at the nano-scale careful examination of atomic configurations and a detailed comprehension of the low scale geometry of the device surfaces.

Nonetheless, pushing scientific innovation to this level of accuracy presents tremendous challenges. Many uncertainties remain due to the difficult theoretical exploration of atomistic phenomena and the enormous investment that the manufacturing of such devices represents. An accurate understanding of the interaction between matter and mechanical and electromagnetic waves should be carried further down to smaller sizes. And most importantly, based on a solid understanding of these phenomena, common, standard and reliable design patterns need to be identified for this novel class of devices. In order to meet that need, the typical response of the industry has been catalyzed through repetitive prototypical iterations in the laboratory, until best of breed solutions emerged. In this scope, a key step in reducing the design time of nanoscale devices is the development of efficient simulation tools that will allow experimentalists to test their thought experiments in a few hours instead of months [1]. In that respect, due to the inadequacy of current simulation tools, algorithmic innovation is necessary in several areas. Consequently, new research initiatives have recently emerged such as the IMPACT center for advancement of MEMS/NEMS VLSI at the University of Illinois, in order to "develop and enhance the understanding of experimentally observed, fundamental multi-physics and multi-scale principles and processes governing MEMS/NEMS performance." In such initiatives, we cannot but insist on the multi-physics nature of the scientific protocol. Remarkably, the major limitation to device scaling today is the heat generation arising from the transport of considerable electrical currents at small scales. Besides, much uncertainty in the assessment of device performance arises from the considerable effect of the interaction between random fluctuations of the systems interfaces and the wavelike nature of the particles involved in conduction. The multi-scale aspect of the problem is

non negligible as well, as the interesting properties extracted from atomistic behaviors eventually need to be integrated in large scale systems.

We introduce in this dissertation a series of novel simulation schemes that allow computationally efficient inclusion of the atomic scale surface geometry in statistical models of coupled thermal conduction, electronic transport, and mechanical deformation in M/NEMS, thermoelectric, and transistor applications. In each of these areas, the ability to efficiently model surface roughness is key to understanding the subsequent coupling between propagation of charge and heat carriers and, in the case of MEMS, the associated mechanical deformation resulting from asperities at the boundary of these devices. In the area of M/NEMS the model we introduce shows good agreement with experimental observations on the coupling between applied pressure and contact resistance in gold micro-contacts. In the field of thermal transport, inclusion of similar roughness models in semi-classical simulations of heat conduction shows unprecedented agreement with experiments on nanowires. In particular, this study holds strong promise for application of artificialy rough semiconductor nanowires as efficient thermoelectric materials, where the geometry of the wire surface is engineered to decouple heat and electronic transport. In the area of electronic transport, we introduce in this dissertation a quantum model based on a non-parabolic 3D recursive Green function approach which allows one to accurately account for the electronic quantum resistance caused by the asperities at the boundary of real semiconductor nanowires of diameter $D < 10$ nm. Based on the validation of these models with experimental measurements, the simulation frameworks are extended to novel materials and structures, such as graphene based applications, quantum modulated transistors (QMTs), and patterned semiconductor sheets.

In the latter applications, the essence of the device physical behavior is tightly bound to surface roughness. For instance, in the area of M/NEMS, metallic micro-contacts are frequently used in radio-frequency systems as variable capacitors and controllable electrical switches [2, 3], with particular advantages over solid state technologies for tunable filter applications. Indeed, micro-fabricated metal switches offer high off-state and low on-state impedance, and a wide bandwidth of operation. Nevertheless, both in the off and on states, failures are still frequent due to overheating and discharges between contacts. A careful study of electronic and heat exchange as a function of the pressure applied between contacts reveals that the roughness of the surface is principally responsible for resistivity at the contact and subsequent heating due to Joule processes. In addition to a statistical study of

these phenomena, we propose a multiscale framework that models gas discharges between non-contacting electrodes, based on a percolation approach which reproduces heat and electronic exchange in a gas. Additionally, for semiconductor applications, nanowires have demonstrated great efficiency in the areas of field effect transistors [4], interconnects [5], and heterostructures [6]; and they most recently have drawn much interest for their possible thermoelectric applications [7, 8, 9, 10, 11]. With widths ranging from a few hundreds of nanometers down to 10 nanometers and below, the variations of the nanowire geometry with surface roughness range from a few percent to almost 20% in the case of the thinnest devices. As such, these fluctuations have non-negligible impact on the resistance of the device during heat and electronic conduction. In the most extreme cases, surface roughness may become the principal mode of scattering of charge or heat carriers, making entire applications revolve around careful nano-engineering of surface roughness properties [7]. In the pioneering applications of field effect transistors (FET), electron-surface interaction was studied early on due to its major impact on conduction in MOSFET channels. Incidentally, the necessity to engineer thermally efficient devices only appeared in the early 2000s, and naturally drew attention to the phonon-surface interaction. When the characteristic dimensions of asperities at a rough surface come to the order of the phonon or electron wavelength (5-30 Å), it is furthemore expected that the surface scattering rate will be altered to reflect the effect of the interface roughness. Such mechanisms are crucial to the understanding of electron transport in transistor inversion layers, where accurate models have been developed based on perturbation theory [12, 13]. We will present such models based on a semi-classical approach to electronic transport and apply our knowledge to engineer efficient FET devices that minimize the intrinsic effect of surface roughness. In addition, we introduce a novel semi-classical treatment of the interaction between heat waves and surface roughness scattering based on perturbation theory. The efficiency of the computational approach allows for a rapid investigation of thermoelectric applications, where engineering of surface roughness enables a decoupling between electronic and heat transport. This study is further carried to nanowire scales below 10 nm, where the dimensions explicitly call for a quantum treatment of the interference between electronic wave functions and the repetitive patterns in the rough surface of real 3D nanowires. We propose a treatment of this problem through a novel non-parabolic 3D recursive Green function approach that we introduce in this dissertation.

The knowledge of surface geometry is central to understanding the physics

of new electronic applications, ranging from the microscopic scale to the atomic scale. Within this scope, we should emphasize the statistical nature of the following study. Indeed, material interfaces are intrisically random, and no two devices have the exact same geometry when it comes to the detailed configuration of nanoscale surface fluctuations. As a result, the following discussion is based on statistical averages, equivalently in time or in space in the steady state, as per the principle of ergodicity. In view of this, computational simulations appear as a powerful alternative, where a series of geometries can be artificially generated on a computer, and physical properties averaged over this ensemble. In the following chapters, we will use this statistical Monte Carlo approach in many occasions in order to derive the average behavior of devices with realistic geometries. In addition, the quantum nature of carrier transport at small scales intrinsically revolves around probability distribution functions, which is well illustrated by Heisenberg's uncertainty principle. In that respect, the semi-classical semiconductor Monte Carlo approach will be used to derive average device characteristics based on quantum processes, by collection of statistics on time series and ensemble of carriers.

At this stage, technology computer aided design has an important role to play in the understanding of the multi-physics processes involved in novel electronic devices. Nevertheless, the multi-scale approach faces a specific challenge, as atomistic computations are extremely expensive in terms of computational time, and are difficult to generalize to large systems. In this document, we adopt a new simulation framework under which the behavior of novel integrated NEMS can be efficiently predicted. The central idea resides in a hierarchy of stochastic simulation techniques which can be seamlessly integrated in a systems approach (Fig. 1.1). At the top of the hierarchy, highly non-linear coupled thermo-electrical behaviors of integrated systems are modeled by means of percolation theory. This approach hinges on the prior knowledge of the statistical characterization of heat and electronic transport in nano-channels. To serve this purpose, Monte Carlo simulation has already shown great performance in achieving an optimal trade-off between accuracy and computational efficiency. At the bottom of the hierarchy, highly non-linear boundary effects are treated by means of molecular dynamics or full quantum simulations, and integrated in the Monte Carlo simulation. The result of this hierarchy is an efficient simulation tool of large scale systems based on an accurate understanding of the relevant low dimensional coupled physical phenomena which are exploited in novel devices.
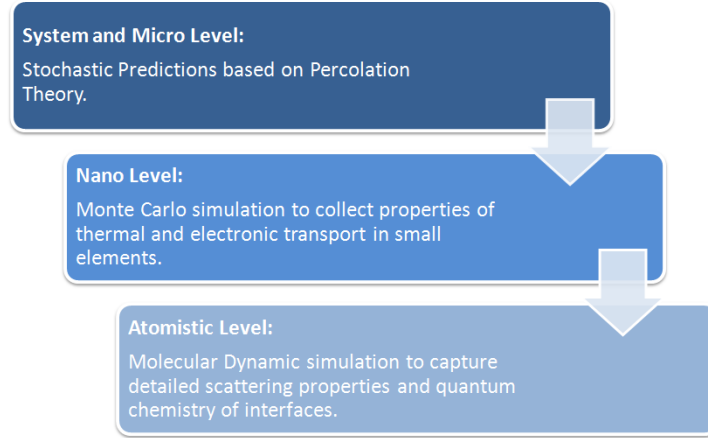
Figure 1.1: Multi-scale framework introduced for the stochastic prediction of large scale integrated novel nano-electromechanical systems.

Following this multiscale approach, this disseration progresses from an investigation of the effect of surface roughness at the micro-scale to associated effects in the sub-10 nm range. At first, we will show utilization of random geometries with Kirchhoff's, Joule's and Euler's classical laws in metallic micro-contacts. Subsequently, we will narrow down our scope to semi-classical effects in electronic and heat transport due to surface roughness scattering mechanisms in the 20-500 nm range. Finally, we will refine our understanding with an accurate quantum treatment of electronic interferences due to random geometry fluctuations below 10 nm.

This dissertation is organized as follows. In chapter 2, we present the statistical base for the study of surface roughness, on which we will further develop our physical models of mechanical, thermal and electronic processes. In particular, we will see how a parametric model can be inferred from experimental observations of device geometries in the 2D and 3D case, and how one may generate computerized random surfaces and edges based on the same set of parameters. In chapter 3, we will directly apply the latter surface models to analyze electro-mechanical coupling in gold-gold micro-contacts, under an extension of Holm's theory allowing for collection of averages over a series of rough contacts. This approach will be validated by a good agreement with experimental observations on such contacts at the micro-scale. In chapter 4, we complete our study of MEMS micro-contacts with models of failure mechanisms by electrical discharges in the OFF state of such devices. In particular, this study will illustrate the ideas behind the higher level of our multi-scale framework. In that respect, we will present a novel percolation approach as an efficient way to capture the highly non-linear electro-thermal

failure of non-contacting MEMS. Through an approximation based on a statistical description of small 1D conducting channels of gas, we will further study how the geometry and the roughness of metal electrodes alter the reliability of these devices. Because this model revolves around a statistical understanding of quasi-1D conducting channels, we present, in chapter 5, the Monte Carlo method to stochastic assessment of electronic performance in quasi-1D semiconducting nano-channels. In the scope of MEMS sensors and actuators, ultimate down-scaling will lead, in the near future, to the use of beams and cantilevers based on nano-whiskers, nanotubes or nanowires. Hence, it is important to understand the new electronic transport processes which arise from the interaction of the wave-like nature of electrons and random fluctuations of the channel interfaces in devices with intrinsically high surface-to-volume ratio. The Monte Carlo method will be demonstrated as an efficient tool to determine an optimal way to achieve electrical performance in quasi-1D silicon nano-channels. In addition, we will extend this investigation to FET devices, and apply our knowledge of statistical surface roughness to efficient design of devices limiting surface induced resistivity. The method discussed in this scope allows inclusions of energy dependent scattering mechanisms, the full electronic band structure in silicon based nanowires, and quantum correction in order to most accurately approximate surface and size effects in the latter devices. In chapter 6, we extend this approach to thermal conduction in such nano-channels. In particular, we introduce a novel formalism under which one can treat the coupling of heat waves with the interface roughness of semiconductor nanowires, using perturbation theory in wires of diameter ranging from 20 - 500 nm. The resulting predicted behavior shows excellent agreement with experimental data and sets high hopes for future efficient nano-engineering of nanowire surfaces for thermoelectric applications. After validating this new theory on silicon nanowires, we will explore the case of Ge and GaAs nanowires, which exhibit extremely low thermal conductivity as a result of surface roughness. In addition, we apply the latter perturbation theory to derive scattering rates in graphene, which will allow consideration of graphene nano ribbons (GNR) for nano-engineering of rough edges in the scope of thermal applications. Finally, in chapter 7, we conclude our discussion with more visionary approaches to quasi-1D electronic transport for channels ultimately scaled below 10 nm. At these extreme sizes, the coupling of waves and interfaces becomes strongly modified by the quantum nature of the problem, and we present a new approach to computing quantum-enhanced surface roughness scattering for electronic and heat transport in ultra-scaled nano-channels.

In particular, we extend the recursive green function (RGF) approach to non-parabolic bands and 3D geometries, for investigation of transport in ultra-thin silicon nanowires. In conclusion, we examine the effect of real boundaries on visionary devices, based on a quantum operation of the transistor effect, such as quantum modulated transistors (QMT) and asymmetrically patterned semiconductor nano-ribbons.

CHAPTER 2

STATISTICAL CHARACTERIZATION OF ROUGH SURFACES FOR NANOSCALE ELECTRICAL AND THERMAL APPLICATIONS

2.1   Introduction

Describing atomic scale geometry of solid interfaces has drawn much attention for applications to thermal and electrical conduction in semiconductor field effect transistors (FET) [14, 15, 16, 17], semiconductor nanowires (NW) [18, 7, 19, 20], contacting MEMS switches [21, 22, 23, 24], and the mechanical properties of contacting regions in MEMS switches [23]. In all cases, it was experimentally and theoretically demonstrated that electrical and thermal conduction in rough devices is drastically reduced in comparison to the theoretical case of perfectly smooth devices [23, 16, 7, 8, 20]. Modeling rough boundaries presents challenges in the accuracy of the statistical description of randomly fluctuating device boundaries. In the following discussion, we present a set of statistical tools which enable both the extraction of relevant statistical parameters from the experimental observation of rough surfaces, and a methodology to model artificially rough surfaces for their accounting in computer simulations of low scale device physics.

In view of this, one of the most compelling cases is the electronic transport across a contact formed by two rough metallic interfaces (Fig. 2.1.a). In this case, as an electrical current flows through constrictions formed by contacting asperities at the surface of each electrode, the resulting interfacial resistance depends directly on the nano-scale fluctuations of the respective surface geometries [25, 26]. As those constrictions limit the total flux of current from one electrode to the other, the throughput resistance rises across increasingly rough contacts. In addition, mechanical properties of such contacts depend themselves on the nano-scale geometrical fluctuations of the surface of the electrodes. Indeed, the effective contact area is a direct result of the number of asperities that achieve physical contact, their shape, and their plasticity [25, 26]. The pressure required to put electrodes in contact is a function of the same effective contact area, and subsequently depends on the same nano-scale physical properties of the surfaces. Hence,
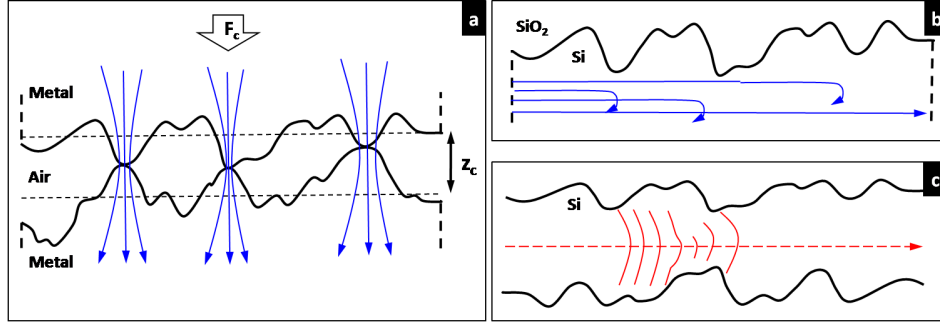
Figure 2.1: (a) Model of conduction across a contact formed by rough metallic electrodes, (b) model of electrical conduction parallel to a rough semiconductor boundary, (c) model of heat propagation in a rough nanowire. Blue lines depict flux of electrical current, red lines depict propagation of the crystal lattice vibrations.

drawing an accurate physical picture of micro-contacts becomes increasingly complex when considering interface roughness, and strongly relies on an accurate description of surface geometry.

Similarly, conduction in the vicinity and parallel to boundaries is strongly limited by interface roughness (Fig. 2.1.b). In silicon FETs for instance, asperities at the silicon interface perturb the electric field in the electronic conduction channel [15, 16]. As a consequence, charge carriers scatter from perturbations in the electrical potential near the boundary, and a surface roughness induced electrical resistance appears in the device [27]. It was again theoretically and experimentally demonstrated that strong fluctuations of the interface geometry increase resistivity in comparison to smoother boundaries [14, 28].

Very recently, rough nanowires attracted the attention of the scientific community for their promise as efficient thermoelectric devices. In this case, it was shown [7, 8, 19, 20] that electronic conduction and thermal conduction could be decoupled in nanowires by careful nano-engineering of arbitrarily rough surfaces. In this case, heat carriers, which take the form of vibrations of the crystal lattice, are strongly perturbed by the series of constrictions along the nanowire (Fig. 2.1.c). As a result, these nanowires have very poor thermal performances, while keeping a relatively high electronic conductivity due to careful sizing of the boundary fluctuations. In this case again, accurate description of boundary geometry is required.

While it is natural to picture a device boundary as a spatially fluctuating two-dimensional (2D) nappe, advances in nano-technology recently enabled the appearance of purely 2D semi-metallic or semiconducting devices, such

10

as monolayer graphene [29, 30]. In these devices, essentially made of a mono-atomic crystalline layer, boundaries are effectively a 1D randomly varying line. In order to describe surface limited conduction in this new class of devices, we need to further consider the geometrical description of rough 1D boundaries.

In the subsequent sections, we provide a theoretical statistical framework to analyze and describe rough boundaries for applications in nano- and micro-scale conduction and mechanics. Beyond the possibility to theorize on hypothetical boundary shapes, it must be noticed that considerable advances in nano-scale interface physics have been driven by the improvement of experimental observation techniques, such as transmission electronic microscopy (TEM) [31, 15] and scanning tunneling microscopy (STM) [32], which now allow resolution of surface geometry down to the atomic scale. Consequently, the following sections will also respond to the need of extracting statistical roughness parameters from experimental observations. In section 2.2, we present our general statistical framework necessary to the description of randomly varying interfaces. Sections 2.3 and 2.4 respectively apply this knowledge to the analysis and random generation of 1D and 2D boundaries.

2.2 Statistical Description of Rough Surfaces

Regardless of the application, real semiconductor or metal interfaces present asperities linked to the imperfections of the growth or deposition process, and the intrinsic discrete nature of the organization of atoms in the crystal structure. While modern observation techniques such as transmission electron microscopy (TEM) allow the observation of the exact geometry of each individual samples under consideration, average properties of electronic and heat conduction at these interfaces may only be deduced using a statistical approach. Hence, stochastic tools may be used to describe the distribution of asperities at a surface [14, 16, 33, 34, 35, 36], their height, and the geometry of the surface as a whole. In this section, we provide a set of such tools which may independently be used on 2D surfaces or 1D edges, and allow one to either retrieve the statistics of experimentally observed interfaces or generate digitized random surfaces for applications in computer simulations.

For this purpose, it is natural to define the root mean square (rms) roughness, which in essence measures the average rms height of the geometrical

fluctuations of a surface defined by a height profile $h(\mathbf{r})$.

$$\Delta_{rms} = \sqrt{\frac{1}{A_\Omega} \int_\Omega (h(\mathbf{r}) - \overline{h})^2 \, d\mathbf{r}} \tag{2.1}$$

where A_Ω is the area of the domain of integration Ω. For computer applications, the domain of integration is commonly discretized in a set of N points $\{\mathbf{r}_i\}_i$. In this case, the formula above becomes

$$\Delta_{rms} = \sqrt{\frac{1}{N} \sum_1^N (h(\mathbf{r}_i) - \overline{h})^2} \tag{2.2}$$

Here we have used the average height of the profile $\overline{h}$:

$$\overline{h} = \frac{1}{N} \sum_1^N h(\mathbf{r}_i) \tag{2.3}$$

In addition, several applications, especially in the area of micro-contacts, hinge on the notion of a height distribution function $p(h)$. The height distribution function gives the probability of finding a point in the domain Ω at a given height h. It is a common assumption to use Gaussian statistics for the height distribution function:

$$p(h) = \frac{1}{\Delta_{rms}} \exp\left(-\frac{h^2}{2\Delta_{rms}^2}\right) \tag{2.4}$$

Notice here that all previous definitions may be equally used in the case of a 2D surface or a 1D edge.

It is also necessary to have information about the corellation between heights at two points $\mathbf{r_1}$ and $\mathbf{r_2}$. Indeed, two surfaces may have the same height distribution function and roughness rms, but drastically different frequencies of spatial fluctuations. To account for the frequency properties, we define the height-height correlation function as the following spatial average:

$$\tilde{C}(\mathbf{r}') = \left\langle [h(\mathbf{r}) - \overline{h}][h(\mathbf{r} + \mathbf{r}') - \overline{h}] \right\rangle_{\mathbf{r}} \tag{2.5}$$

It is worth noticing the connection of the height-height correlation function and roughness rms as

$$\tilde{C}(0) = \Delta_{rms}^2 \tag{2.6}$$

In the case of an isotropic 2D surface, it is useful to have an expression of the correlation function which does not depend upon the direction of the

space vector $\mathbf{r}$, but only the distance from an arbitrary center. We use for that the circular average of the height-height corellation function:

$$C(r) = \frac{1}{2\pi} \int_0^{2\pi} C'(r \cos \phi, r \sin \phi) \, d\phi \qquad (2.7)$$

In the case of a discrete representation of the surface, the integral becomes a sum over the discrete mesh points and the values of the correlation function may be approximated by its value at the nearest grid point to $\mathbf{r} = (r \cos \phi, r \sin \phi)$.

Finally, it is useful to introduce Fourier transforms to manipulate quantities related to the spatial frequency of the geometrical fluctuations. The Fourier transform of the height profile is

$$H(\mathbf{k}) = \int_\Omega h(\mathbf{r}) \exp(+i\mathbf{k} \cdot \mathbf{r}) \, d\mathbf{r} \qquad (2.8)$$

and the height spectral density (SD) is subsequently

$$S(\mathbf{k}) = |H(\mathbf{k})|^2 \qquad (2.9)$$

In this case, by the Wiener-Khinchin theorem [37], the height-height correlation function is conveniently defined as the inverse Fourier transform of the height spectral density

$$\tilde{C}(\mathbf{r}) = \frac{1}{(2\pi)^d} \int_{\Omega'} S(\mathbf{k}) \exp(-i\mathbf{k} \cdot \mathbf{r}) \, d\mathbf{k} \qquad (2.10)$$

where d is the order of dimensionality (1 or 2). This expression is numerically advantageous, as it allows the use of fast divide-and-conquer computational schemes of Fourier transforms - called fast Fourier transform (FFT) - in order to compute the height-height correlation function from the height SD.

In this section, we have reviewed all major statistical tools necessary to analyze the roughness of random surfaces. In the next two sections, we present schemes to numerically compute those quantities, and generate discretized random rough 1D edges and 2D surfaces relevant to the statistics presented above.

2.3   One-Dimensional Case

The case of 1D edges not only presents the advantage of allowing simpler statistical description, but is also suited to the description of physical processes

at the boundaries of emerging two-dimensional materials such as single layer graphene. In this section we will see how Gaussian statistics can be used to describe and numerically generate random rough graphene edges.

In the scope of the description of such physical behaviors as the scattering of heat or charge carriers at a rough interface, a quantity of frequent interest is the height spectral density, as used for instance in the reference work of Goodnick et al. [14, 15]. In section 2.1, we have seen how the SD could be obtained from accurate imaging of a surface. In fact, experimental work often describes 2D surfaces in terms of quantities associated with the SD, such as the roughness rms and the auto-correlation length [33, 34, 15, 31]. Here we will extend these common two-dimensional concepts to the 1D case.

At the center of the following discussion is the not so trivial idea of generating a random surface on a numerical mesh which is continuous enough to represent a real surface and at the same time abides by the analytical form which is chosen a priori for a spectral density. By "continuous enough", we imply here the mathematical meaning of the term, where we expect at least the lowest order derivatives to be continuous and smooth.

In order to be consistent with the work on carrier scattering by Goodnick et al. and followers [14, 16, 15, 19, 38, 39], we use a Gaussian model for the height correlation function. This model has shown particularly good agreement with experimental data in various cases of metallic and semiconducting crystals [15, 31].

$$C(r) = \Delta^2 e^{-r^2/L^2} \tag{2.11}$$

where L is called the correlation length, or autocovariance length. As will be shown below, L is a measure of the average width of the peaks of the geometrical fluctuations at the interface.

A Gaussian correlation function presents the additional advantage of yielding a Gaussian spectral density. Indeed, using the FT formalism of equation 2.10, we have in one dimension:

$$
\begin{align}
S_{1D}(k) &= \int_{-\infty}^{\infty} \Delta^2 e^{-r^2/L^2} e^{-ikr}\, dr \tag{2.12} \\
S_{1D}(k) &= \sqrt{\pi}\Delta^2 L e^{-k^2 L^2/4} \tag{2.13}
\end{align}
$$

We would like to emphasize here the fact that, for one-dimensional edges, the SD varies proportionally to the autocovariance length L and has units of cubic meters.

We will now address the problem of finding a randomly rough height function $h(r)$ which has the assumed correlation function from equation 2.11,

with given rougness rms Δ and autocovariance length L. As an important reminder for the following discussion, let us note the following property of a Gaussian function

$$\int_{-\infty}^{\infty} e^{-\frac{r^2}{\sigma^2}} \, dr = \sqrt{\pi}\sigma \tag{2.14}$$

Let $\eta(r)$ be a continuous uniform random distribution on $\left[-\sqrt{3}\Delta, +\sqrt{3}\Delta\right]$. This means that along the length of the domain, the probability that the height of $\eta(r)$ is in the interval $[z - dz, z + dz]$ is the same and equal to $\frac{dz}{2\sqrt{3}\Delta}$ for any z in the interval above, and zero otherwise. For this random distribution, the average height is theoretically 0 for an infinitely long domain, and the roughness rms is

$$\begin{aligned} \Delta^2_{rms}(\eta) &= \langle(\eta(r) - \bar{\eta})^2\rangle_r \\ &= \Delta^2 \end{aligned} \tag{2.15}$$

In addition, let $g(r)$ be a Gaussian function such that

$$g_{1D}(r) = \frac{1}{(\sqrt{\pi}\sigma)^{1/2}} \cdot e^{-\frac{r^2}{2\sigma^2}} \tag{2.16}$$

and take the height function

$$h_{1D}(r) = g_{1D} * \eta(r) \tag{2.17}$$

where $*$ represents the spatial convolution. For speed of computation, it is possible to use the properties of Fourier transform to express the convolution in terms of Fourier transform (FFT) and inverse Fourier transform (IFFT) of the respective functions

$$h(r) = \text{IFFT}(\text{FFT}(g) \cdot \text{FFT}(\eta)) \tag{2.18}$$

The roughness rms of this height function is

$$\begin{aligned} \Delta^2_{rms}(h) &= \int_{-\infty}^{\infty} \Delta^2_{rms}(\eta) \cdot g^2(r) \, dr \\ &= \frac{\Delta^2}{\sqrt{\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{r^2}{\sigma^2}} \, dr \\ &= \Delta^2 \end{aligned} \tag{2.19}$$
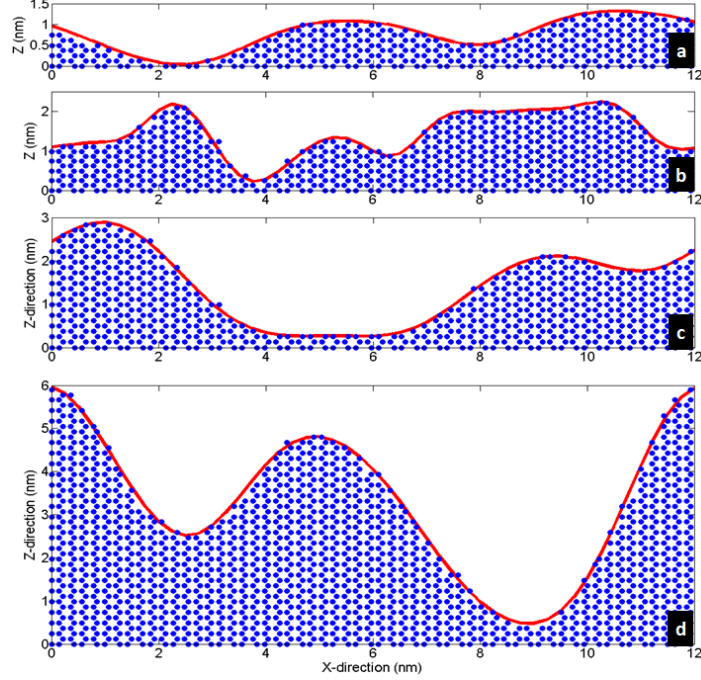
Figure 2.2: 12 nm rough graphene edges randomly generated using a Gaussian height correlation function with (a) $\Delta = 3$ Å, $L = 22$ Å, (b) $\Delta = 3$ Å, $L = 11$ Å, (c) $\Delta = 5$ Å, $L = 22$ Å, and (d) $\Delta = 10$ Å, $L = 22$ Å. Blue dots represent the position of carbon atoms in a real armchair graphene nano-ribbon.

and its height-height correlation function is

$$
\begin{aligned}
C(r) &= \frac{\Delta^2_{rms}(\eta)}{\sqrt{\pi}\sigma} \int_{-\infty}^{\infty} -\frac{(x+r)^2+x^2}{2\sigma^2}\, dx \\
&= \frac{\Delta^2}{\sqrt{\pi}\sigma} e^{\frac{r^2}{4\sigma^2}} \int_{-\infty}^{\infty} -\frac{(\sqrt{2}x+r/\sqrt{2})^2}{2\sigma^2}\, dx \\
&= \Delta^2 e^{-r^2/4\sigma^2}
\end{aligned}
\tag{2.20}
$$

Using $\sigma = L/2$, the autocovariance length associated to the height profile $h(r)$ takes the desired value L. Hence, we have defined a random rough 1D height profile with roughness rms Δ, autocovariance length L, and the desired Gaussian correlation function. It is important to notice in this case that, due to the convolution approach which was used, the roughness peaks have a Gaussian shape. The width of the peak is given by the autocovariance L. In Fig. 2.2, 12 nm of randomly generated graphene edge is represented for such a Gaussian model with varying Δ and L. As expected, the parameter L effectively controls the width of the peaks while Δ adjusts the average height of the geometrical fluctuations. In addition, for each ribbon

edge, we give an indication of where the carbon atoms would be located in armchair ribbons. While the statistical parameters chosen in this example do not appear in contradiction with the graphene lattice, this picture nevertheless raises the question of the actual definition of *edge* for physical applications. Indeed, one may simply consider that the physical edge is the ensemble of lines connecting the nearest neighbor carbon atoms located at the edge. For electrostatic or mechanical purposes, the definition of *edge* would rather be function of a critical distance from each atom linked to the Lennard-Jones potential. In the case of electronic conduction near the edge, the actual limit of the ribbon could also be defined from the dimension of orbitals of carbon atoms at the edge. In the latter cases, the interface would appear smoother, thus making the continuous Gaussian model more relevant to physical application than a discrete description by the position of edge atoms.

The question remains of the amount of statistical knowledge that one can gain from a ribbon edge of finite size. This point is central to determine the quality and the confidence in estimation of the roughness statistics from a real, finite size 1D edge. For this purpose, we compare the correlation function computed from equation 2.10 for randomly generated edges of 50 nm and 1 μm respectively (Fig. 2.3). The edges are generated in such a way that for an infinitely long edge, the auto-correlation length is exactly the Gaussian function as specified by 2.20. It is apparent that an estimation of the roughness on a 50 nm edge undershoots the value of the height rms and overshoots the value of the auto-correlation length. In our scope, the value of the auto-correlation length is obtained from a least square optimization scheme which finds the best fitting Gaussian function to the correlation function computed from the real surface. In addition, we compare the auto-correlation function obtained from the randomly generated edge to one obtained from the edge resulting exactly from a mapping of carbon atoms to the surface in an graphene nano-ribbon (GNR) armchair configuration. In both cases, we can see that both definitions of the edge lead to a similar value of roughness rms. Nevertheless, due to the natural atomic periodicity, the correlation function of the atomic edge does not converge to zero and rather keeps a finite positive value. This results in a larger correlation length of the atomic edge in comparison to the smoothened randomly generated edge. Not surprisingly, longer edges yield a more accurate estimation of the edge roughness statistical parameters.

In summary of this section, we have described stochastic procedures to both generate 1D rough edges with a given Gaussian auto-correlation length,
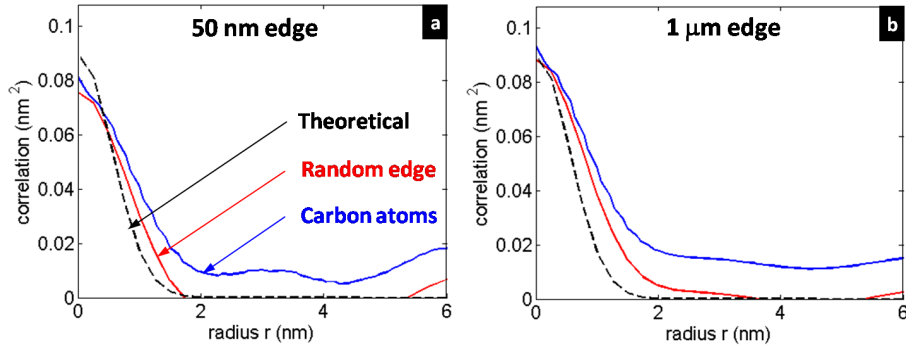
Figure 2.3: Comparison of the correlation of a randomly generated rough edge and the associated "carbon-atom" discrete edge to the theoretical Gaussian correlation with $\Delta = 3$ Å and $L = 22$ Å for integration over (a) a 50 nm edge, and (b) a 1 μm edge.

and estimate the height root mean square and auto-correlation length of a 1D edge. We have applied this methodology to the case of graphene ribbon edges. In this case, the atomic representation of the edge and a randomly generated edge of Gaussian auto-correlation function yield comparable height rms. In addition, we showed that edges described by the direct position of atoms yield larger correlation lengths. The next section will address similar schemes in the case of 2D surfaces.

2.4 Two-Dimensional Case

The theoretical case of 1D surfaces only appears in the limit of single layer materials, and the majority of real devices, semiconductor or metal, naturally present 2D boundaries. In the current regime of device down-scaling and nano-engineering, typical dimensions become comparable to those of asperities at the interfaces. Hence, the effects of geometrical non-linearities on mechanical, thermal and electrical processes are of increasing importance in the overall performance of real nano-engineered devices. In addition, due to boundary fluctuations, the single device behavior may vary from one device to another. Consequently, in order to understand the physics of modern devices and model their performance, it is necessary to have theoretical tools which describe interfacial randomness. In a similar way to what we previously modeled in the case of 1D edges, we present here a set of statistical recipes to randomly generate 2D surfaces of a given spectral density.

Here also, the question of generating a digital random 2D surface on a computer, with a given SD, and enough smoothness and continuity to reflect

real geometries, is a central and far from trivial issue. We follow here the approach of the previous section, based on the convolution of two functions presenting appropriate statistics. Here again, let us assume $\eta(\mathbf{r})$ a uniform continuous distribution on $[-\sqrt{3}\Delta, +\sqrt{3}\Delta]$, and $\mathbf{r}$ on the surface domain. In the 2D case, let us define the new function

$$g_{2D}(\mathbf{r}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-r^2/2\sigma^2} \tag{2.21}$$

Similarly to what has been derived in section 2.3, defining $h(\mathbf{r}) = g_{2D}(\mathbf{r}) * \eta(\mathbf{r})$, the resulting roughness rms is

$$
\begin{aligned}
\Delta_{rms}^2(h) &= \int_0^{2\pi} \int_0^\infty \Delta_{rms}^2(\eta) \cdot g^2(r)\, dr\, d\theta \\
&= \Delta^2
\end{aligned}
\tag{2.22}
$$
$$\tag{2.23}$$

and the height-height correlation function is similarly

$$C(\mathbf{r}) = \Delta^2 e^{-r^2/4\sigma^2} \tag{2.24}$$

which leads to the choice of $\sigma = L/2$ in order to guarantee that the autocovariance length associated to h is L. Hence, we defined a height function with a given roughness rms Δ and a Gaussian correlation function. In 2D, using the Fourier transform formalism introduced in section 2.2, the resulting spectral density from such a Gaussian correlation function is Gaussian by the Wiener-Khinchin theorem:

$$
\begin{aligned}
S(\mathbf{k}) &= \int_{-\infty}^\infty \int_{-\infty}^\infty \Delta^2 e^{-(x^2+y^2)/L^2} e^{-i(k_x x + k_y y)}\, dx\, dy \\
&= \pi \Delta^2 L^2 e^{-k^2 L^2/4}
\end{aligned}
\tag{2.25}
$$

Incidentally, the formula above is in accordance with the work of Goodnick et al. and followers [14, 15, 16, 19, 38].

To give a geometrical idea of surface roughness parameters, Fig. 2.4 depicts the roughness height profile h on a 50×50 nm surface for different rms Δ and autocovariance lengths L. Similarly to the 1D case, with our given random generation approach, the average height of surface peaks varies in proportion to Δ, while the average width of these peaks increases with larger L. Parameters of $\Delta = 1.7$ Å and $L = 2.2$ nm correspond to the case of Si-SiO$_2$ interfaces as reported in [14].

It must be noticed here that the Gaussian correlation function results from
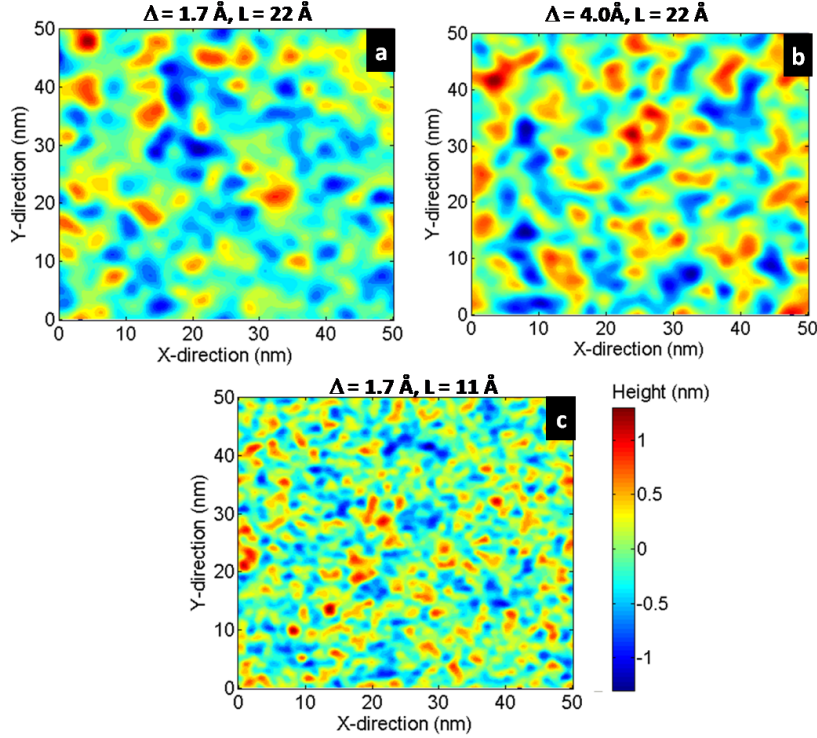
Figure 2.4: Example of rough 2D profiles for (a) $\Delta = 1.7$ Å, $L = 2.2$ nm, (b) $\Delta = 4$ Å, $L = 2.2$ nm, (c) $\Delta = 1.7$ Å, $L = 1.1$ nm.

a series of integrations over an infinite surface. Therefore, a purely Gaussian correlation function is only the physical property of a quasi-infinite surface. In estimating the roughness parameters of a real rough interface based on the Gaussian model, the question for the experimentalist remains to know what the required total area of the interface is in order to achieve a satisfactory level of confidence in the estimated roughness rms Δ and autocovariance length L. Thus, we computed the correlation function of randomly generated surfaces of different areas on Fig. 2.5. Estimating the correlation function on a small surface results to larger errors in comparison to the theoretical Gaussian model. In order to further explore this phenomenon, we represent in Fig. 2.6 the error made in approximating the correlation function of a randomly generated finite rough surface by the expected Gaussian function. For this purpose, we compute the relative error on the first 6 nm of the radius from the center of the surface, corresponding to the interval on which the autocorrelation function is represented on Fig. 2.5. It is noticeable that this error decreases with larger total area, and that the total area required to achieve a given level of accuracy appear to be independent of the roughness rms Δ. Thus, an estimation error of 1.5% is achieved by averaging statistics

20
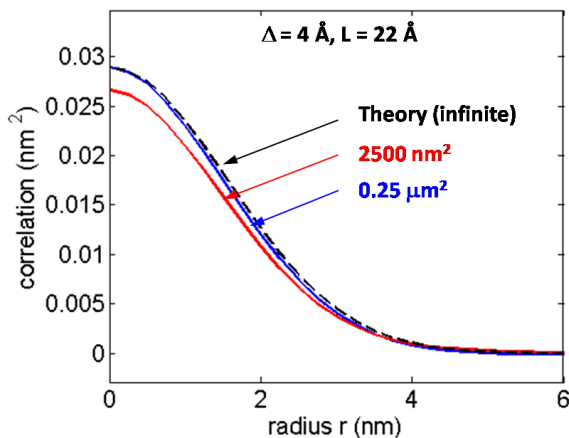
Figure 2.5: Estimation of the 2D correlation function on randomly generated surfaces of varying total area.

on a 400 nm surface, regardless of the roughness height rms. Collecting roughness parameters by fitting the correlation function on smaller surfaces will degrade this accuracy.

A methodology to extract Δ and L naturally appears in the light of the discussion above. First, one starts with the computation of the radius averaged autocorrelation function, as described in section 2.2. The factor Δ is directly given by the values at radius zero of the computed autocorrelation function $C(0) = \Delta^2$. Next, one finds the best Gaussian fit to the estimated autocorrelation function. This can be achieved by a least square minimizing scheme where one seeks the parameter λ which minimizes

$$L = \min_{\lambda} \left\{ \sum_i \left[C(r_i) - \Delta^2 e^{r_i/\lambda} \right]^2 \right\} \tag{2.26}$$

Then, the autocorrelation length L is defined as described in the equation above. As discussed above, the accuracy of such an estimation depends on the total area over which the correlation function was averaged. It is important to remember here that the statistics used above all hinge on the assumption that the geometry of the surface is essentially isotropic. Nevertheless, it is conceivable to have different (Δ, L) pairs in orthogonal directions along the surface. In this case, each pair would be estimated similarly to the method described above, but the autocorrelation function would only be averaged in the direction of interest.

Finally, we would like to bring to the reader's attention that additional significant geometrical features may be estimated from higher order prop-
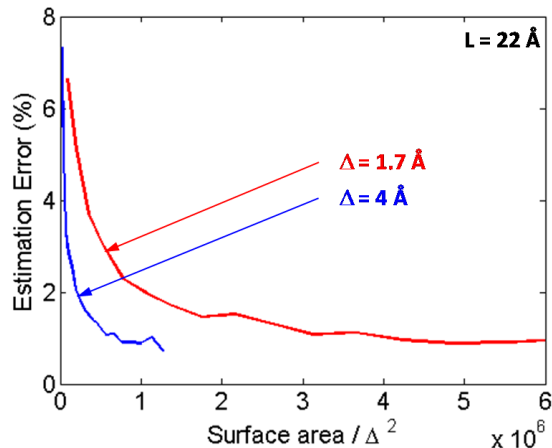
Figure 2.6: Error in estimating the correlation of a finite size rough 2D surface by a Gaussian function. Results are compared between surfaces of varying roughness rms.

erties of the autocorrelation function. While the correlation function of a strictly Gaussian noise is Gaussian, periodicity in the surface geometry introduces damped oscillations in the correlation function. In Fig. 2.7.a and Fig. 2.7.d, the correlation function was computed for a periodical lattice of Gaussian peaks with a nearest neighbor distance of 40 nm. The position of the order 1 maximum gives directly the nearest neighbor distance in a perfectly periodic surface pattern.

In the physical case of conduction across a contact between electrodes, it is frequent to use a cylinder based model to represent the constrictions through which electrical current flows (Fig. 2.1). For this case of interest, we compute in Fig. 2.7.b and Fig. 2.7.e the correlation function for elevated circles of similar radius $a = 25$ nm, and center-to-center distance of $d_{cc} = 100$ nm. In the case of asperities of regular size and spacing, the first minimum of the correlation function gives the diameter of asperities, and the first order maximum gives again the separation between asperities. Finally, we test the case of randomly distributed asperities of regular, but randomly varying diameter (Fig. 2.7.c and Fig. 2.7.f). In this case we generated a distribution of cylinders of maximum diameter $a_{max} = 55$ nm and average separation $\langle d_{cc} \rangle = 200$ nm. As such, the first minimum reveals the maximum diameter, and the first order maximum approximates the average distance between asperities.
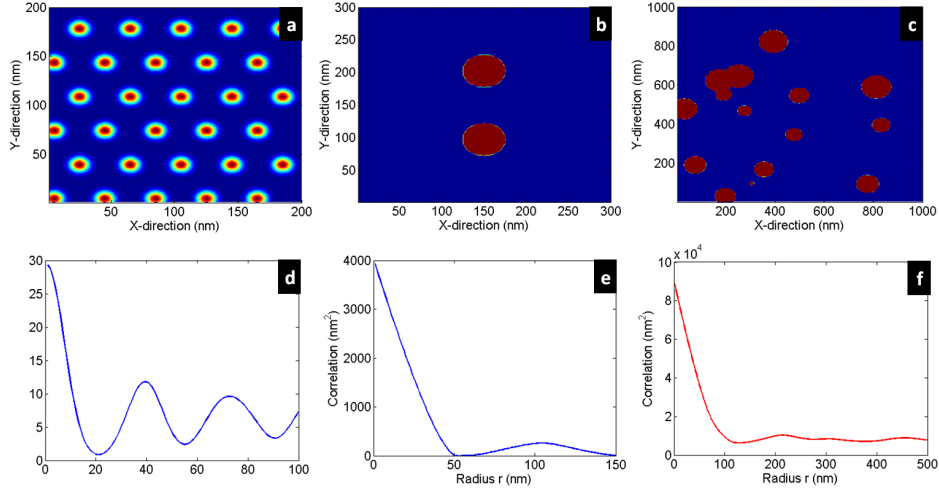
Figure 2.7: Profile of deterministic 2D surfaces in the configuration of (a) equally distributed Gaussian functions, (b) equal size and constant separation circles, (c) randomly distributed circles of bounded random radius, and (d),(e),(f) associated correlation functions.

2.5 Conclusions

At the conclusion of this analysis, we have scoped methods to both extract statistical roughness parameters from a randomly fluctuating solid interface, and numerically generate such surfaces for applications in computer simulations. Our primary here concern was conduction across and parallel to solid metallic or semiconductor device boundaries. In this case, we presented a zero order and first order statistical framework based on a Gaussian model in order to describe the random geometrical variations of the solid surface due to the presence of imperfections, or asperities, on the surface. This Gaussian model allows for a computerized reproduction of such random surfaces in both 2D and 1D cases. The 1D case is of particular relevance in the current scientific context where mono-atomic layer materials, such as graphene, are integrated in electronic devices. From a study of the correlation function of interfaces, we were able to extract important information on the sizing of surface asperities in terms of the statistical roughness height root-mean-square Δ and autocorrelation length L. These two parameters are central to further discussion on nano-scale interface processes during heat and eleconic transport. We presented here a method to extract these parameters directly from experimental observations. This pair of parameters may be used to draw a direct connection between theoretical results and experimental imaging of similar surfaces. This may be done through a direct generation of random device geometries in computer simulations, or

more evolved physical tools using directly the roughness parameters.

CHAPTER 3

STATISTICAL SIMULATION OF ROUGHNESS INDUCED ELECTRO-MECHANICAL COUPLING IN M/NEMS METAL CONTACTS

3.1 On the Need for Algorithmic Innovation in M/NEMS Modeling

Micro- and nano- fabricated sensors, actuators, and other passive or active electro-mechanical coupled devices are now successfully developed throughout the world [40, 41, 21, 42, 23]. Examples of fabrication processes for such devices, called micro- or nano-electromechanical systems or M/NEMS, include surface micro-machining, laser etching, microelectrodischarge, 3D printing, micromolding and high aspect ratio metal plating. Through this high variety of available fabrication processes, M/NEMS devices have been used toward a broad range of applications, in mechanical sensing, atomic scale sensing, micromotors, scanning probe tips for scanning tunneling microscopes, chemical microanalysis, and radio frequency devices. These devices involve several physical behaviors, classical or quantum in nature, such as mechanical motion, nano-scale heat and electronic transport, fluid dynamics, mechanical stress, surface chemistry, and capacitive electrical effects. Thus, effective and efficient design of M/NEMS requires detailed knowledge of a broad range of physical mechanisms. As a result, most new M/NEMS products are the result of repeated iterations of experimental prototypes, instead of emerging from effective computer aided design (CAD) tools. In this scope, a key step in reducing M/NEMS design time is the development of efficient simulation tools that will allow experimentalists to test their thought experiments in a few hours rather than months [1]. In view of this, due to the inadequacy of current simulation tools, it is fair to say that algorithmic innovation is necessary in several areas.

In this chapter, we introduce a new multi-scale coupled electro-mechanical simulation algorithm to study the effect of atomic scale interface roughness on the electro-mechanical coupling of micro-engineered metal contacts. Such devices are frequently used in RF systems as variable capacitors and controllable electrical switches [2, 3], with particular advantages over solid state

technologies for tunable filter applications. Indeed, micro-fabricated metal switches offer high off-state and low on-state impedance, and a wide bandwidth of operation, which makes them good candidates for high-end targeted communication systems. The function of the micro-fabricated metal switch essentially lies in two metal plates, whose positions with respect to each other repeatedly vary, bringing the plates in contact at times (Fig. 3.1.a). The contacting state constitutes what is called the ON state. Key characteristics of these devices are their reliability, lifetime and current handling ability, which are tightly bound to the properties of the metal-metal contact. In particular, poor ability to handle high current at the contact is the main source of device failure [23, 43, 44, 31, 45]. In that respect, atomic scale topology of the contacting metal-metal interface, and the associated contacting pressures, are determining factors in the device lifetime.

Due to the complex shape of the device at small scale, the finite element method (FEM) is commonly used for analysis of electromechanical coupling with randomly varying boundaries [46]. However, the use of FEM discretization often leads to systems of equations that are large, sparse, ill-conditioned and semi-definite or indefinite. The multi-scale aspect of the problem can be leveraged by alternative algorithmic schemes, such as multi-grid [47, 48] or meshless approaches [49]. Nevertheless, those calculations require a fixed shape, are computationally expensive, and can only leverage the general randomness of the atomic scale boundaries when included in Monte Carlo simulations.

In the following section, we introduce a novel statistical computational scheme based on the detailed knowledge, at the atomic scale, of the electrode surface roughness, that will allow time efficient simulations of coupled electro-mechanical properties of the micro-fabricated metal contacts. In particular, this new approach is derived from parameters directly observable from the electrode surface, using the model described in chapter 2. Based on Holm's theory, we are able to determine an average resistivity vs. applied pressure curve for gold-gold micro-contacts. Close agreement is found with previously reported experimental observations [23], as we report a resistance varying from 300 mΩ to 75 mΩ for applied contact pressure varying from 25 μN to 500 μN for sputtered gold electrodes.
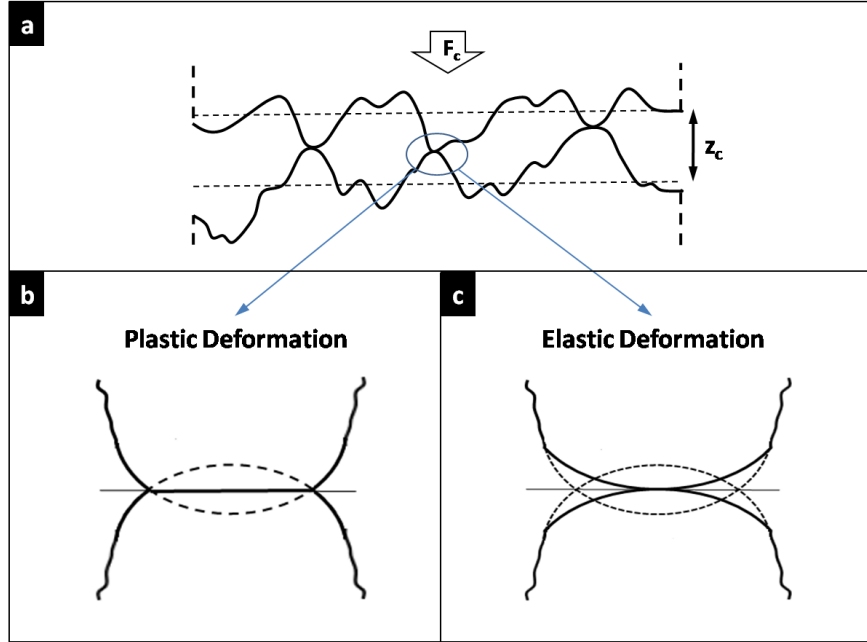
Figure 3.1: (a) Model of random surface fluctuations in a nanoscale metal contact. (b) Model of plastic deformation under applied pressure at the contact. (c) Model of elastic deformation under applied pressure at the contact.

3.2 Nanoscale Mechanical and Electrical Properties of Metal Contacts

Surface roughness is represented by randomly shaped and distributed asperities at the metal interfaces. Models of these random fluctuations were presented in chapter 2. In the contacting state, asperities at each electrode form constrictions through which the flow of electrical current is constrained (Fig. 3.1.a). As a result, asperities and their shape are responsible for the effective mechanical and electrical contact resistance (Fig. 3.1.b and c). In this section, we will present the necessary theory in order to derive our statistical model.

We center our derivation on Holm's theory of the topology of electrical contacts [25]. Under applied pressure at the contact, asperities will deform under two physical regimes: elastic at low pressure, and plastic at high pressure. The contact resistance is directly a function of the effective surface of contact between each electrode. As such, for two single contacting asperities, assuming each asperity is spherical in the vicinity of the contact point and that the asperities are aligned, the contact area is either

- a circle of area exactly equal to the intersection of two asperities. This

is the case of plastic deformation at high contact pressure. One should notice in this case that the total contacting area, and subsequently the total contacting resistance, grows as the square root of the distance between each electrode.

Or

- a circle of radius less than or equal to the intersection between the two asperities, with the radius of contact a_1 given by equation 3.1. This is the case of elastic deformation at low contact pressure, where the contacting area, hence the electrical resistivity, is minimal.

$$a_1 = \left(\frac{3F_C R}{4E^*} \right)^{1/3} \tag{3.1}$$

where R is the radius of curvature of the asperities (assuming the same radius of curvature), F_C is the applied force at the contact, and E^* is the composite Young's modulus, defined as

$$(E^*)^{-1} = 2(1 - \nu^2)E^{-1} \tag{3.2}$$

where E is the material Young's modulus and ν the Poisson ration. For gold, we use $E_{Au} = 78$ GPa, and $\nu_{Au} = 0.44$. The equations above assume that the same material is used for both electrodes. Given the mechanical properties of the contact material and applied pressure, it is consequently possible to derive the contact area and the resulting electrical resistivity of the contact. Such models have been used successfully in the past [21, 25], though it is arguable whether modeling a contact surface by a distribution of spheres of varying radius yields sufficient accuracy.

Here, we propose an investigation of the next order of accuracy based on a similar model, and a statistical description of surface roughness that matches experimental observations with greater precision. In chapter 2, we introduced a statistical model to generate a geometrical rough surface whose autocorrelation curve is within a given interval of confidence from observable measurements on a real surface. In view of this, Fig. 3.2.a is obtained from scanning tunneling microscopy (STM) of a real surface of sputtered gold (after Jensen et al. [31]). Measurements give a roughness rms $\Delta = 13$ nm and autocorrelation length $L = 60$ nm. The following model is similar to Holm's approach described above. Nevertheless, it is now possible to use:

- two simulated rough surfaces to estimate the real geometrical contact map between electrodes
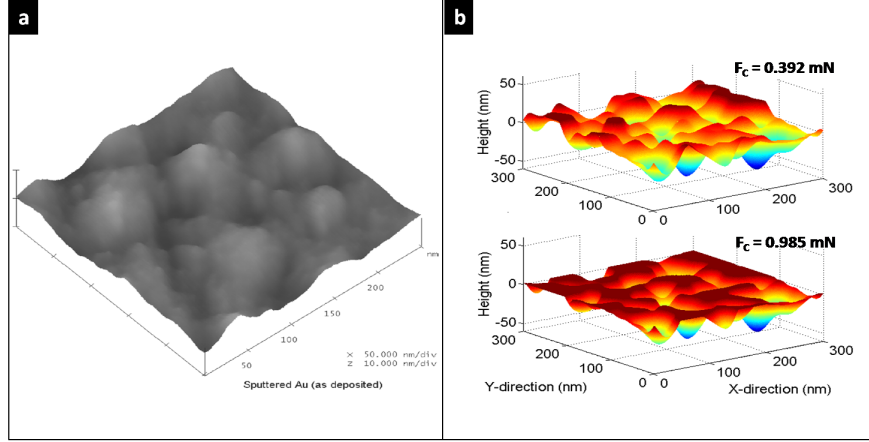
Figure 3.2: (a) STM picture of a gold contact, with measured surface rms $\Delta = 13$ nm. Picture after Jensen et al. [31]. (b) Computer generated gold surfaces while contacting under applied pressure. $\Delta = 13$ nm and $L = 60$ nm.

- the simulated shape of observed asperities to estimate the asperity radius of curvature at contact

- a series of surfaces generated in a Monte Carlo run in order to obtain the statistical distribution of the contact parameters

As an example, Fig. 3.2.b depicts a computer generated surface with the same roughness as sputtered gold (Au). The asperities were truncated from intersection with another rough surface separated by a distance related to a given applied pressure at the contact. The model described below will help understand how to relate the latter contact force to the electrode separation, using the material specific mechanical and electrical properties.

Using the Gaussian model of surface roughness, we derive below the relationship between contact radius and applied force (Fig. 3.3). Each asperity has essentially a Gaussian shape, with radius of curvature

$$R = \frac{L^2}{\Delta} \tag{3.3}$$

Injecting this result into equation 3.1, one obtains the relationship between effective contact radius and applied force for one single Gaussian shaped asperity:

$$F_C = \frac{4E^* a_1^3 \Delta}{3L^2} \tag{3.4}$$

Finally, assuming that two objects of similar radius of curvature are brought
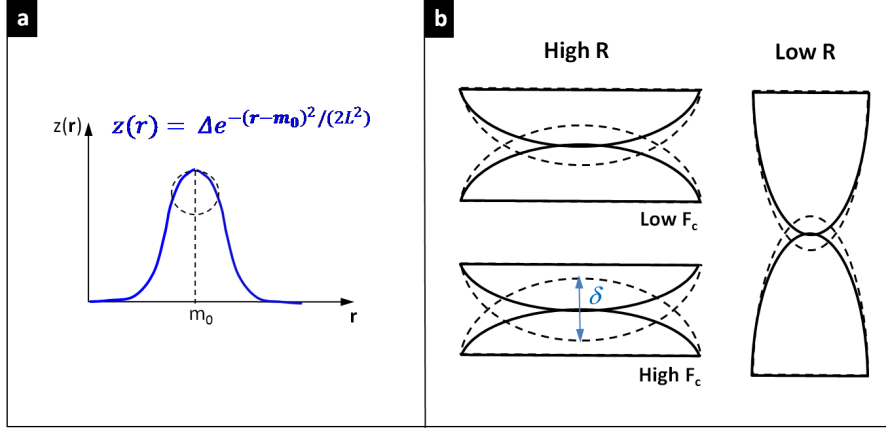
Figure 3.3: (a) Model of a single Gaussian asperity. (b) Contact radius and penetration distance for asperities for small and large radius of curvature, and under high and low contact force.

in contact, the radius of intersection between asperities follows:

$$a_2 = \left| \delta(R - \frac{\delta}{4}) \right|^{1/2} \sim (RF_C)^{1/3} \tag{3.5}$$

where δ is the distance of penetration (Fig. 3.3.b). We should emphasize the difference between a_1, the radius of effective contact area, and a_2 the radius of asperities intersection area, which corresponds to the imaginary case where asperities penetrate one another (represented as dashed lines in Fig. 3.3). As such, the quantity a_2 is easily measurable from two computer generated surfaces, while a_1 is necessary in order to compute the relationship between contact force and resistance. Under small applied pressure and with the assumption that the radius of curvature is large in comparison to the distance of penetration ($R >> \Delta$), we can hypothesize a linear dependency between a_1 and a_2

$$a_1 \approx \alpha a_2 \tag{3.6}$$

with α a constant of proportionality.

Finally, the dependency of the contact resistance on the radius of effective area follows directly from Holm's derivation [25]:

$$r_{\text{asperity}} = \frac{\rho}{2a_1} \tag{3.7}$$

where ρ is the resistivity of the material used for contact. This apparent resistance essentially follows from the constriction of the electric field lines in the vicinity of the region of contact [25, 24]. In particular, as lines of

current are forced through a longer path to flow through the asperity, an increased contribution is perceived from the electrode resistivity, yielding a perceived contact resistance. Note that this effect is independent of resistivity perceived from the oxidation of electrodes.

We described above the fundamental equations for contact of Gaussian asperities. For two intersecting rough surfaces, it is now possible to approximate the effective contact area between each asperity, the applied force necessary to achieve such a contacting pattern, and the resulting contact resistance. In the next section, we will present the details of the computational Monte Carlo scheme used to compute electromechanical properties of metal contacts, and results obtained from this scheme on gold-gold micro-contacts.

3.3   Computational Approach and Results

Computation of the applied pressure and electrical resistance at a micro-contact is performed with the following scheme. Firstly, two rough surfaces are generated using the method described in chapter 2, based on a Gaussian autocorrelation function, and a separation z_c between the two surfaces. A binary map of the intersection between the two surfaces is then generated such that $MAP(\mathbf{r}) = 1$ if the surfaces intersect, and $MAP(\mathbf{r}) = 0$ otherwise (Fig. 3.4.a). The autocorrelation function of the map is then computed following the approach of section 2.4. As derived from the latter section, the auto-correlation of a map of distributed patterns has its first minimum at the equivalent radius of the largest pattern a_{max} (Fig. 3.4.c). The following approach effectively consists in converting the random map produced by the simulated intersection into an equivalent circle model, from which the contact force and contact resistance can easily be derived. Consequently, a new map is generated from a series $(a_{2,k})_k$ of circles of normally distributed radius such that

$$a_{2,k} = \max\left(a_{max}, \mathcal{N}(a_{max}, a_{max}/\sqrt{2})\right) \tag{3.8}$$

and for which the total contact area remains the same. As per equation 2.2, this approach consists in generating a surface map of exact same roughness rms Δ (Fig. 3.4.b). It is possible to interpolate the second order roughness statistics from each map from the best fit Gaussian parameters, as described in section 2.4. As such, we observe that the autocorellation length L of the equivalent circle model overshoots that of the original map by an average of 9.8% (under the contact of two surfaces with $\Delta = 13$ nm and $L = 60$
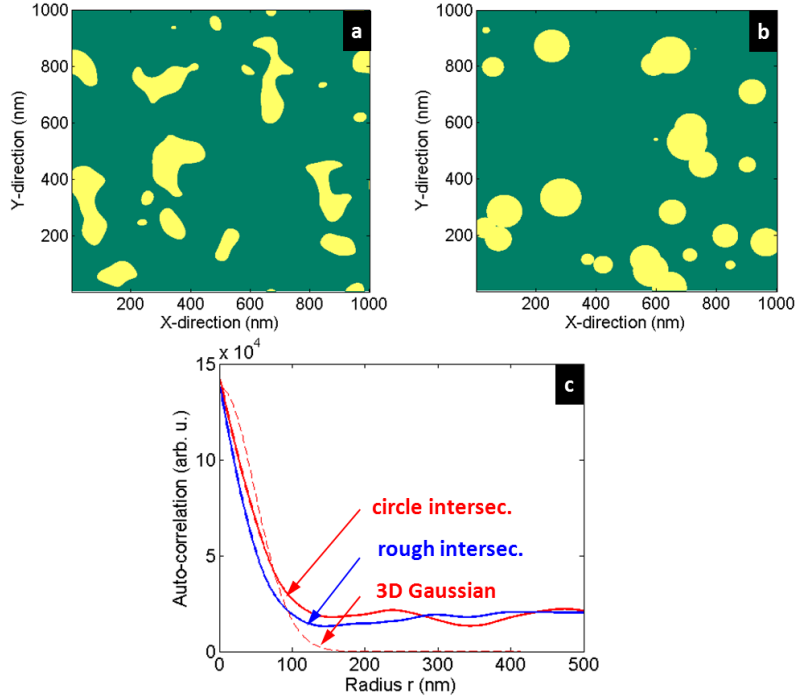
Figure 3.4: Computer model of a 1 μm $\times$ 1 μm gold-gold micro-contact. (a) Simulated map of the intersection of two surfaces generated from a Gaussian model with separation $z_c = 20$ nm. (b) Derived equivalent circle model of the intersection map. (c) Auto-correlation functions of models (a), (b) and the theoretical Gaussian model of the intersecting surfaces. $\Delta = 13$ nm and $L = 67$ nm throughout.

nm, with separation $z_c = 20$ nm). The second maximum of the autocorrelation function gives the average distance between contacting patterns d_{cc} (see chapter 2), and is generally in good agreement between the original map and the equivalent circle model. We only noticed minimal influence of these two quantities on the contact resistance and neccesary applied pressure computed from this model. Finally, for each asperity, the resistance r_k of the circular contact is computed from equation 3.7 , and the contacting force from equation 3.4. Subsequently, the total resistance is derived as the ensemble of asperity contact resistances in parallel, and the total contacting force as the sum of all asperity contacting forces.

In order to illustrate the method above, Fig. 3.5 shows the equivalence between the circular model and the effective contact map for different electrode separations using sputtered gold-gold contacts. It is noticeable that, in every case, we have by definition an exact equality between roughness rms Δ. In addition, it is remarkable that the fit for the autocorrelation length, that is in essence the average size of contact patterns, converges to

32
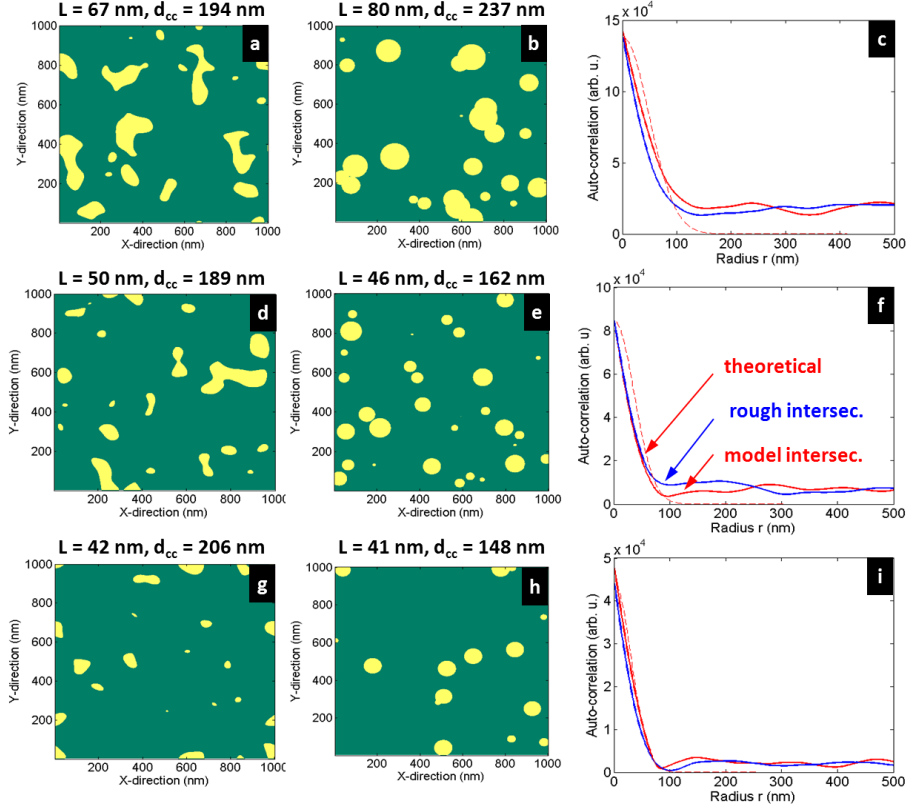
Figure 3.5: Effective contact map between two simulated sputtered gold surfaces ($\Delta = 13$ nm, $L = 60$ nm), equivalent circle model and associated auto-correlation functions. (a), (b), and (c) electrode separation $z_c = 20$ nm. (d), (e), and (f) $z_c = 25$ nm. (g), (h), and (i) $z_c = 30$ nm.

a stronger agreement at low contact pressures.

We use the following statistical Monte Carlo approach to investigate the optimal contact properties from the geometry and roughness of the electrodes. A statistical series of random surfaces are generated at a given separation z_c and fixed roughness parameters Δ and L. For each specimen, the equivalent circle model is computed, from which we derive the associated contact force and total electrical resistance (Fig. 3.6).

The simulation process consists, for each pair of contacting surfaces, in bringing the latter progressively in closer contact to each other, and determining the contact map and equivalent circle model, from which the contact force and electrical resistance are inferred. Importantly, at larger electrode separation, not all sets of electrodes are contacting (Fig. 3.6.d and 3.7.d). Therefore, the average electromechanical properties are collected on the ensemble of contacting surfaces, and must be put in perspective of the percentage of non-contacting surfaces. It is noticed that the percent of non-
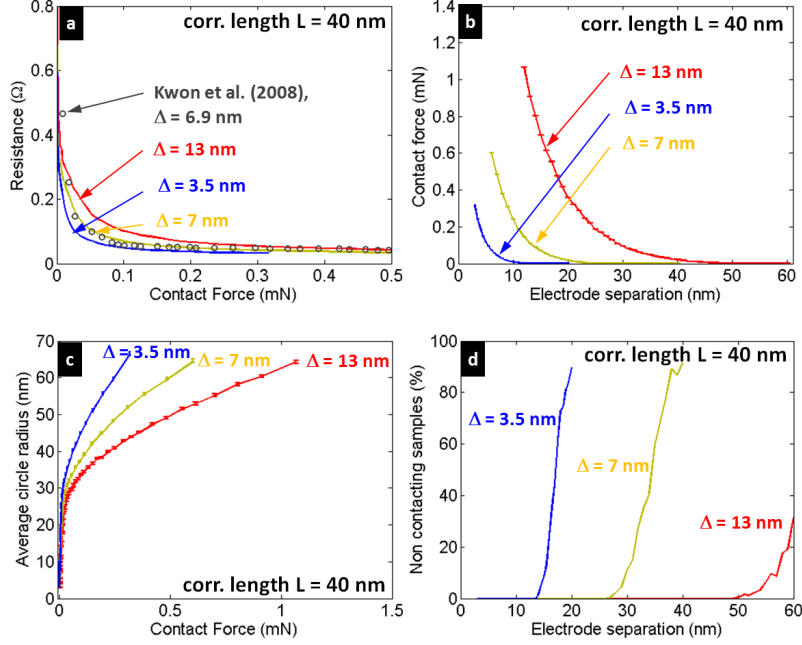
Figure 3.6: Electromechanical properties of rough gold-gold micro-contacts with varying roughness rms Δ, with $L = 40$ nm. (a) Contact resistance vs. applied mechanical pressure for sputtered gold-gold micro-contacts. (b) Necessary applied force to achieve a particular electrode separation for the same contacts. (c) Average radius of the equivalent circle model for the same surfaces. (d) Percentage of non-contacting samples during the Monte Carlo runs.

contacting surfaces is a unique function of z_c/Δ for the ensemble of Δ at a fixed autocorrelation length L. In particular, this rate increases at wide separation z_c and at lower Δ. There is a critical separation under which all electrodes are contacting. We report this critical separation for gold-gold surfaces as $z_{c,0} = 13$ nm, 26 nm, and 49 nm for $\Delta = 3.5$ nm, 7 nm and 13 nm respectively, and L fixed at 40 nm. These conditions represent different qualities of reported sputtered gold surfaces [23]. In addition, we show here that the proportion of non-contacting surfaces is higher at large autocorrelation length L. Indeed, L is in essence the periodicity of the asperity patterns. Rough electrodes with higher L will consequently contain lesser asperities at a fixed area, consequently increasing the probability of having none of these in contact at a given separation.

Next, we compute the average radius of the equivalent circles for varying combinations of (L, Δ). As the applied contact force increases, the equivalent radius increases, showing the expected growth of contact area (Fig. 3.6.c and 3.7.c). In particular, the latter increase of contact area is steeper
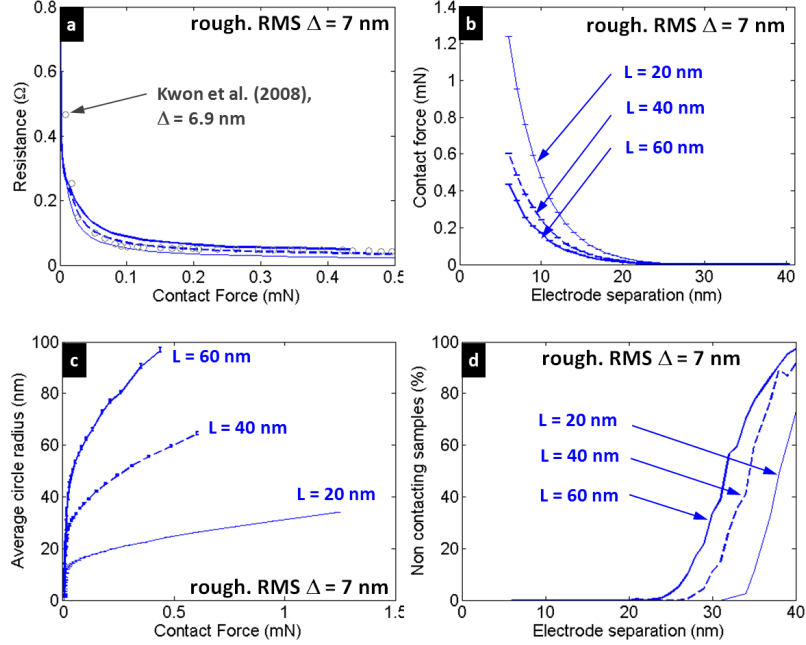
Figure 3.7: Electromechanical properties of rough gold-gold micro-contacts with varying roughness auto-correlation L, with $\Delta = 7$ nm. (a) Contact resistance vs applied mechanical pressure for sputtered gold-gold micro-contacts. (b) Necessary applied force to achieve a particular electrode separation for the same contacts. (c) Average radius of the equivalent circle model for the same surfaces. (d) Percentage of non-contacting samples during the Monte Carlo runs.

for smoother surfaces, at low Δ and large L. For an applied force of 0.5 mN over a 1 μm $\times$ 1 μm surface, the average radius of contact patterns strongly varies with L, from 22 nm at $L = 20$ nm to 98 nm at $L = 60$ nm. Higher contact radius directly translates into a reduced electrical resistance. In addition, we derive the applied pressure resulting from the simulated electrode separation (Fig. 3.6.b and 3.7.b). The necessary pressure to apply in order to bring rough electrodes to a given distance from one another is higher at low L and large Δ. Indeed, a greater energy is required to compress asperities of large magnitude (large roughness rms Δ). Similarly, roughness patterns of long wavelength result in less geometrical asperities brought in contact on a given surface. As a result, the necessary applied force to bring surfaces to a given separation z_c is lower at high autocorrelation length L. It is remarkable that the parameter Δ affects the contact force by one order of magnitude higher than L, making Δ the primary parameter to optimize when designing rough sputtered gold electrodes. For a given applied pressure of $F_c = 0.2$ mN, we report here a resulting separation of 3.5 nm, 11.9,

and 26.8 nm for sputtered gold-gold micro-contacts of $L = 40$ nm and $\Delta = 3.5$, 7 and 13 nm respectively.

Tested values of roughness rms Δ were selected after the reported values of Kwon et al. [23] on similar contacts, and we noticed a better experimental fit with $L = 40$ nm for the reported electromechanical properties on the latter contacts (Fig. 3.7.a). In particular, we compute the electrical resistance vs. applied pressure curve for experimental conditions analogous to observations of Kwon et al. with reported $\Delta = 6.9$ nm [23]. (Fig. 3.6.a and 3.7.a). Simulated results show good agreement with the experimental observation, with a strong fit for a simulated $\Delta = 7$ nm and $L = 40$ nm, resulting in a relative error of only 3.6%. The simulation shows that the contact resistance of gold-gold micro-contacts decreases at high contact pressures to uniformly level out around 80 mΩ for pressures above 0.3 mN. The parameter L lightly impacts the minimum achievable contact resistance. As such, rough contacts should be optimized for low L in order to achieve minimal contact resistance. The highest discrepancy is observed at low applied pressure below 20 μN. We expect this mismatch to be caused by a higher contribution of electrostatic force for barely contacting electrodes, which was not included in the present model.

3.4   Conclusions

In this chapter, we presented a novel, computationally efficient statistical simulation scheme in the scope of deriving essential eletromechanical behavior of MEMS micro-contacts based on the nano-scale rough geometry of metal electrodes. The scheme allows collection of statistical averages using a Monte Carlo approach based on a model derived from Holm's theory of metal contacts. In particular, the simulations are based on a geometrical description of the electrode surface using roughness parameters directly observable on experimental electrodes. As a result of a detailed analysis of geometrical roughness, our predictions of contact resistance curves show excellent agreement with experimental observations. In particular, inferring an equivalent circle model from the statistical roughness of the surface allows one to account with sufficient detail for the electrode mechanical and electrical properties at the contact while enabling fast computational schemes. As a result, the derived model can be used for fast computerized prototyping of N/MEMS micro-contacts, for optimized applications in the field of high frequency tunable filters and controllable switches. From the derived model,

we matched experimental observations on gold-gold micro-contacts, reporting minimal achievable contact resistance of 80 mΩ for applied pressures above 0.3 mN on 1 μm $\times$ 1 μm surfaces of sputtered gold with roughness rms $\Delta = 7$ nm and autocorrelation length $L = 60$ nm. As a result of fast computerized simulations, we were able to quickly analyze the effect of several contact design parameters. In particular, we noticed that the minimal achievable contact resistance will be mostly minimized by reduction of L at the surface. On the other hand, in order to decrease the necessary applied pressure to reach the minimal resistance, the main parameter to optimize for is Δ. Here, we reported values for gold-gold microcontacts. Nevertheless, the simulation framework derived here can be extended to other metals by injecting the material parameters in the framework. Hence, it is possible to simulate contacts between electrodes of platinum, silver, or composites to optimize for best usability in high frequency applications. We noticed most discrepency with experimental results at low applied pressure. In this regime, it is likely that the contribution of the electrostatic force between electrodes is the highest, which should be further included in the model. In addition, we did not account for the resitance due to the oxidation of the surface, which is expected to linearly add to the contact resistance with respect to the effective area of contact. Finally, contact ressitance is a critical factor in failure of N/MEMS micro-contacts through excessive heating in the ON state. However, it is equally important to understand other failure mechanisms, especially in the OFF state. In that respect, we propose, in the following chapter, to establish an additional computationally efficient simulation framework to study failure of the latter devices through electrical discharges.

CHAPTER 4

THE PERCOLATION APPROACH TO ELECTRO-THERMAL FAILURE IN M/NEMS CONTACTS

4.1 Breakdown of RF MEMS Switches by Electrical Discharge

Radio frequency micro-electromechanical systems (RF MEMS) have recently shown excellent performances that have led to their consideration as an alternative to solid-state switches. They indeed offer low on-state and high off-state impedance and a broad bandwidth for operation, which makes them good candidates for tunable filters with applications in wireless communication systems [2, 3]. However, their reliability is limited by various failure mechanisms, whose study is challenging due to the complex interaction between the processes of mechanical deformation, electrical transport, material exchange and heat transfer. As a result, the current state of the art of RF MEMS design is widely based on heuristics and has not yet seen the emergence of consistent fabrication patterns. In the previous chapter, we used our statistical knowledge of electrode roughness to understand the effective contacting resistance of such devices leading to their failure through overheating. In this chapter, we offer a novel approach to understanding electrical failure in the OFF state of micro-fabricated metal switches based on a percolation theory to account for the randomness of the medium between electrodes.

It was reported that during mechanical cycling, RF MEMS show an increase in contact resistance due to the formation of an insulating film at the metal surface [31]. Thus, several failure processes have been considered leading to film formation, including adhesion between the metal electrodes, melting, material transport, charge extraction from the interface, and arc formation [31, 45, 50]. Most recent models emphasize the considerable roughness of the metal surface given the microsopic scale of the devices. However, while conduction in the metal bulk or in the metal-metal contact is well determined by both ballistic and Monte Carlo approaches that show good agreement with experimental data [51, 52], electro-thermal phenomena

leading to arc formation have received little interest despite their importance in the mechanisms of RF MEMS failure.

Earlier research has shed light on similar metal-insulator-metal (MIM) structures, reporting brightly illuminated arcs in the oxide, whose radius increased with the bias voltage [53]. Similarly to recent approaches in the MEMS community, models of conduction in the insulator have been proposed based on the consideration of conducting filaments of normally distributed resistivity and radii [54, 55, 56]. More recently, thermodynamic descriptions of electrical arcs have been under scrutiny for microscopic MIM structures using air or vacuum as an insulator [57, 58]. Additionally, the increasing computational power available to researchers enabled the emergence of molecular dynamic models of charge and material exchange during electrical discharges [59]. While these models offer an accurate description of the various processes involved in the formation of electrical arcs, they cannot yet support a broad and dynamic picture of charge transfer in a metal switch device as a whole.

Hence, electrical discharges in metal-metal structures constitute a favorable ground for testing our multi-scale framework described in chapter 1. Indeed, such failure mechanisms arise from the successive thermally dependent ionization of adjacent small elements of gas. We introduce the high level model of our hierarchy by considering the gas medium between two electrodes as a network of uni-directional conducting elements. It is remarkable that, in the case of a gas, the 1D transport properties can be reasonably approximated by linear laws and do not need the inclusion of boundary effects for each small conducting channel. In this sense, the problem of gas discharge constitutes a simple ground to experiment with the high level model, in which small 1D systems of gas are stochastically integrated. From this approach, we will subsequently derive highly non-linear coupled electro-thermal behaviors of micro-systems based on the statistical knowledge of atomistic conduction phenomena. The simulated results will also provide the necessary basis to start exploring the effects of electrode geometry and roughness on the performances of RF MEMS when they are scaled down.

In this respect, percolation theory has already been used to describe electrical failure of conducting metal films [60, 61]. Percolation is a general probabilistic model to represent a random medium. A simple example of the use of percolation could involve, for instance, the calculation of the probability for a liquid to diffuse through a porous stone. In the scope of electrical engineering, percolation generally consists in modeling a dielectric

medium as a network of conducting elements (resistances, etc.) which can fail with a given probability. More advanced models included heat propagation and temperature dependence of the electrical properties of each element of the network, and have suggested that such schemes could be applied to electrical failure of insulating media [62]. Given the nature of arc failure in RF MEMS while keeping in mind the gaseous nature of the typical insulator under our scope, it is envisaged that percolation theory can lead to a successful description of electro-thermal failure of such devices. In the following sections, we report on how the stochastic modeling of electrical arcs in a metal-air-metal structure based on percolation theory may reproduce, even under simplistic assumptions, physically consistent breakdown phenomena in metal switches.

Finally, we would like to emphasize the general nature of our approach. While the problem within our scope in this chapter is the precise case of RF contacting metal switches, the same approach can be virtually applied to any type of non-linear conduction problem. Hence, discharges between silicon cantilevers or nanowhiskers can be studied in a similar fashion. The medium in between the electrodes need not be an inert gas, but may as well be made of a statistical solid state medium, such as carbon nanotube or nanowire networks. The range of possibilities is simply bound by creativity.

4.2   Physical Background on Gas Discharge

Figure 4.1 schematically represents the modes of operation of a MEMS switch. The roughness of the metal surface increases the contact resistance, and with it, the power losses in the contact. We are mostly interested, for now, in the breakdown of the device in the OFF state due to observed arc discharges.

Although air is, under normal atmospheric conditions, almost a perfect insulator, it is known that under certain applied electric fields and temperatures, it may carry high electrical and heat currents. The phenomenon of gas discharge has already been extensively studied (see for instance [63, 64]) for its various implications in capacitances reliability, vacuum tubes, plasma physics, and surface treatment. We wish here to provide the inexpert reader with the necessary elements of discharge physics in order to understand the model used in the following sections.

A gas can simplistically be seen as an ensemble of hard spheres representing neutral molecules, ions and electrons traveling in space under the influ-
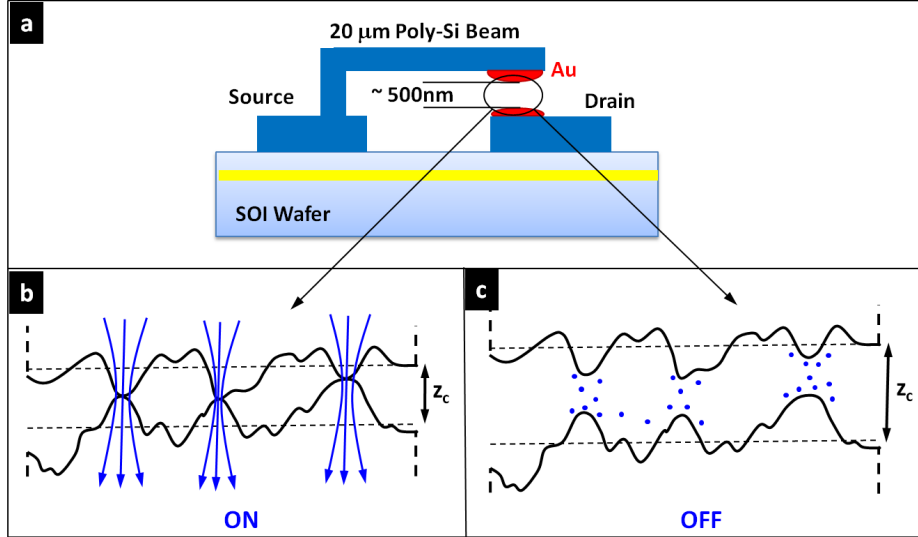
Figure 4.1: (a) Schematic representation of a MEMS switch. (b) Importance of surface roughness in the metal contact region.

ence of thermal agitation, applied forces, and internal interactions. Each particle may similarly scatter from different mechanisms, which accounts for the electrical and thermal resistance of an element of gas. Incidentally, each of these scattering mechanisms corresponds to an exchange of energy between the constituents of the gas.

One important process in a gas discharge arises from the collision of electrons and neutral atoms or molecules. If the electron transfers enough energy to the heavy particle to overcome its ionization energy V_i, the latter may free an additional electron from its outer shell. Hence, under sufficiently high electric fields or thermal agitation, one electron may create, through an avalanche-like process, an exponentially increasing number of electrons. Townsend [65] gives a statistical description of this process by assuming that n electrons traveling through an element of gas dx generate a differential in the number of electrons

$$dn = \alpha n dx \tag{4.1}$$

where α is called the first Townsend coefficient. The term α, the proportion of avalanche carriers generated by one electron per unit length, is dependent on the electric field and the particle mean free path which is, to the first order, inversely proportional to the pressure. Once a successful avalanche has started, electrons travel toward the anode while the ions travel toward the cathode. Energy exchange between those first charge carriers and the metal will feed the starting process and build up the arc of high current

41

between the electrodes. Hence, the time to breakdown is stochastically given by the statistical time for an electron to start a successful avalanche and the statistical time for a starting avalanche to start a successful feedback.

One may argue that, in order to start a successful avalanche, a certain number of electrons must initially be in the gas. At the scale of the device, those electrons are most likely to be injected from the metal contacts. Indeed, carriers in the metal contact see, at the surface, a potential barrier of approximately Φ, the work function of the metal. A certain portion of these carriers may nevertheless overcome this potential barrier and be dragged in the gas. This process may happen under high normal electric fields (field emission and tunneling) or high temperatures at the metal surface (thermionic emission).

Incidentally, the temperature of the gas has a strong impact on the proportion of electrons which can be found in the gas. Generally, a good approximation to the proportion of electrons X with respect to the total number of particles in the gas is given by the Saha equation [66]:

$$\frac{PX^2}{1 - X^2} = CT^{5/2}e^{-\frac{qV_i}{k_B T}} \tag{4.2}$$

where P is the pressure of the gas in atm, V_i is ionization energy of the gas molecules and $C = 3.16 \cdot 10^{-7}$ atm/K$^{5/2}$ is constant. Figure 4.2 represents the proportion of electrons for a pure gas of gold particles, air and a mixture of both.

Finally, for the short time involved in the arcing process, the effect of space charge cannot be neglected. Since electrons are considerably lighter than their positive counterparts, they indeed leave a positive column behind them while traveling to the anode. This column accounts for high fields in its vicinity, where additional avalanches will be favored.

4.3 Percolation Approach to Coupled Non-Linear Electro-Thermal Failure

Given the atomistic picture discussed in the previous section, running Monte Carlo type simulations, similarly to the case of semiconductors, is conceivable. However, due to the lack of periodicity of the gas structure and the high number of particles that must be accounted for, the problem becomes rapidly unsolvable. On the other hand, percolation theory [67] offers a computationally efficient platform to account for the stochastic failure of a random
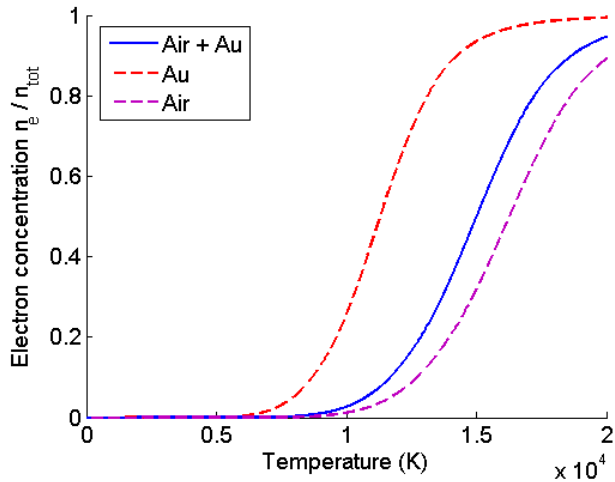
Figure 4.2: Proportion of electrons in a gas given by the Saha equation at atmospheric pressure. The mixture of gas corresponds to a proportion of 5% of Au particles.

medium at a microscopic level. It is important to note here that, because the arcing process involves strong interactions between the gas and metal properties, we seek a model to describe heat and electrical transport in both metal contacts and gas.

Following previous attempts from percolation theory [62], we adopt a random electrical network representation to describe the contact region, including the electrodes and the gas (see Fig. 4.3). Under this assumption the region is modeled by an $R \times S$ rectangle-lattice network of resistive elements, whose properties vary between gas and metal. Each element n is represented by a specific intensity-voltage (I-V) curve which, for a given point of operation in voltage and temperature, will determine its output current i_n and dynamic resistance r_n. Each element consequently represents an infinitesimal piece of metal or gas connected to its six closest neighbors. Although a detailed representation would require that the typical I-V curve for an element of air be computed, a linear approximation is used for the sake of simplicity. If one element connects two nodes of potential V_1 and V_2, our model classically predicts an output current of

$$i_n = \frac{V_2 - V_1}{r_n(\bar{T})} \tag{4.3}$$

The dependence of the dynamic resistance in $\bar{T}$ indicates that the electrical properties of an element are dependent on its own local temperature and the temperature of its closest neighbors which will diffuse in the lattice.

43
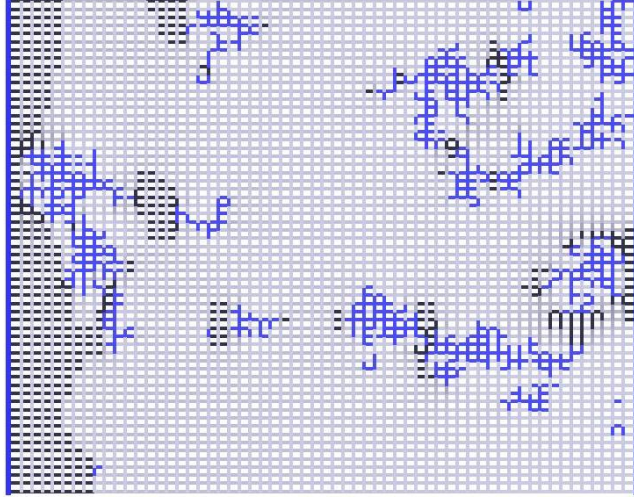
Figure 4.3: Network of 60×60 nodes approaching electrical breakdown. Different gray levels from light to dark correspond to increasing values of resistance from 1 to 2 MΩ. Blue elements represent conducting defects. The electrodes are limited to simple planar contacts.

The temperature dependence of the resistance on the local temperature is assumed linear and varies according to

$$r_n(T_n) = r_0 \left[1 + \alpha(T_n - T_0)\right] \tag{4.4}$$

where T_0 is the temperature of the medium and is assumed to be constant. We further assume that the system is initially in thermal equilibrium at a temperature T_0. The linear approximation of equation 4.4, with a thermal resistance coefficient $\alpha > 0$, is a natural first order assumption in a metal. In a gas, it accounts for the fact that electrons will exchange energy with the excited molecules of air and therefore see an increased resistivity at higher temperatures.

The instability in the gas is starting from the spontaneous creation of conducting defects. Subsequent ionization is reflected by the random appearance of conducting defects in the lattice, with a probability $W_n(T_n)$. This probability is a direct consequence of Saha's equation; therefore we consider an element of gas as conducting when its proportion of electrons reaches $X_{cond} = 0.2$.

Additionally, a screening factor is introduced to reflect the effect of space charge, which favors the creation of defects in the vicinity of already existing defects. Under this assumption elements surrounded by conducting defects

have a higher ionization probability.

$$W_n^{corr.} = \left[p_1 + (1 - p_1)\frac{N_{def}^{(n)}}{N_{neigh}^{(n)}} \right] W_n^{Saha} \tag{4.5}$$

where p_1 is a fixed probability which accounts for the probability of ionization given the number of surrounding defects $N_{def}^{(n)}$. $N_{neigh}^{(n)}$ is the fixed number of neighboring elements of the resistance n.

Finally, the temperature dependence of the system appears from Joule heating and assumes a partial thermalization of the element with its closest neighbors.

$$T_n = T_0 + A \left[r_n i_n^2 + \frac{B}{N_{neigh}^{(n)}} \sum_{k=1}^{N_{neigh}^{(n)}} \left(r_k i_k^2 - r_n i_n^2 \right) \right] \tag{4.6}$$

The constants A and $0 \leq B \leq 1$ are important determinants of the evolution of the system. A value of $B = 1$ assumes that the temperature of an element is only the result of Joule dissipation around it. $B = 0.75$ corresponds to a uniform heating of horizontal and vertical elements. It is also worth noticing that for the special case $A = 0$, under the condition $p_1 = 1$, the probability W_n is the same for all resistances and corresponds to the situation where the degradation is governed by standard percolation. The strictly positive value of p_1 reflects the possible spontaneous appearance of conducting defects under sufficient temperature and in the absence of surrounding elements of ionized plasma. Although the value of p_1 must remain small since a spontaneous ionization is unlikely to happen in the practical range of temperatures, it is necessary that it be non-null in order to ignite the process of electrical breakdown if a uniform and unperturbed starting lattice is assumed.

Further attention must be paid to initial and boundary conditions. A reasonable assumption is that the system starts at thermal and electrical equilibrium, all elements consequently having the same initial temperature T_0 and resistance r_0. The initial resistivity of an element is evidently dependent on the size of an elementary resistance. At a temperature of 300 K, we have the resistivity of air given by $\rho_{air} = 4 \cdot 10^{15} \Omega \cdot m^{-1}$.
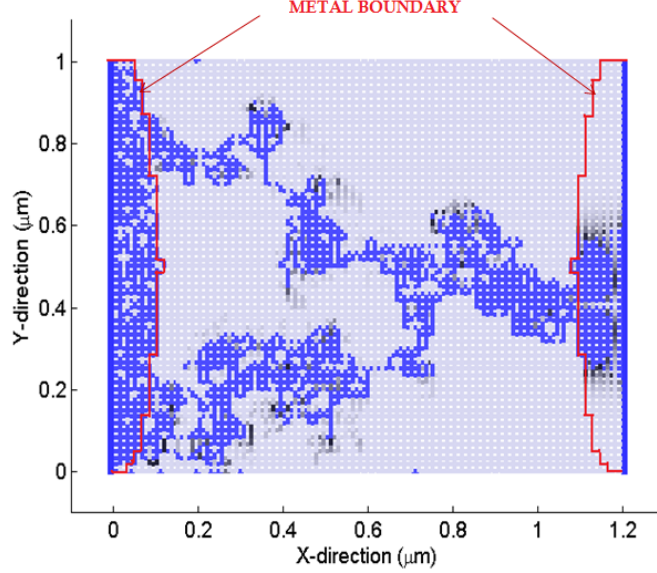
Figure 4.4: System at percolation. An electrical arc of high currents (levels of blue) connects the electrodes.

4.4 Results

Values of $A = 10^5$ K·W^{-1}, $B = 0.6$, $\alpha_{air} = 1.5 \cdot 10^4$K^{-1}, $\alpha_{Au} = 3.5 \cdot 10^{-3}$K^{-1}, and $T_0 = 300$ K are used. As expected, the conductive defects appear by an avalanche phenomena (Figs. 4.3, 4.4, and 4.5).

The resistive growth shows filamentary patterns (Fig. 4.3). As reported in [62], filaments of increased resistivity tend to appear perpendicular to the direction of the current flow. In contrast, conductive defects appear in the direction of the potential gradient. An electrical arc of high electrical current is observed between the electrodes at the time of electrical breakdown (Fig. 4.4).

Additionally, very high temperatures appear near the breakdown time, especially at the anode surface, accounting for the melting of the metal that has been experimentally observed.

The percolation threshold is precisely defined by observing the total current flowing in the grounded electrode. The breakdown is characterized by a drastic rise in the total current transmitted between electrodes. It has been noticed that, while the device average temperature remains close to 300 K before breakdown, it is drastically increased once the arc has appeared. High local temperature gradients, ranging from 300 K – 10,000 K, are observed as a result of the appearance of defects.

A full statistical description of the system is obtained from a statistical Monte Carlo run of the percolation model (Fig. 4.6). The statistical
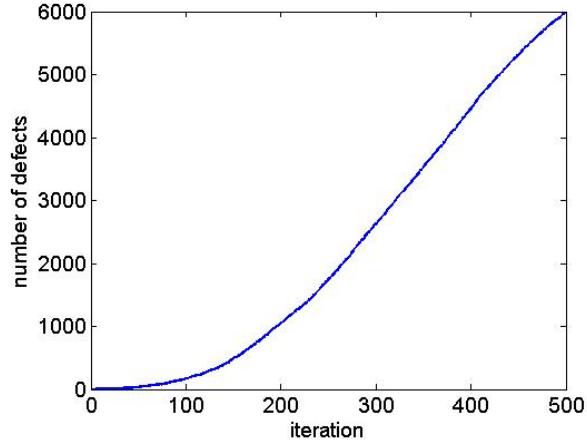
Figure 4.5: Evolution of the number of defects in the domain over time.
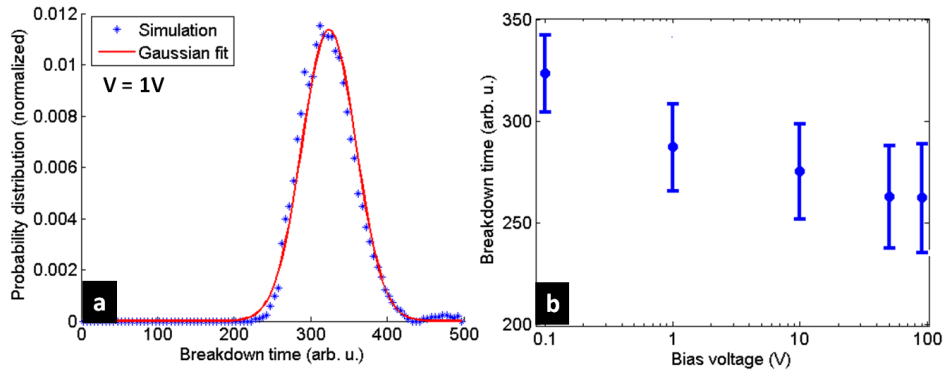
Figure 4.6: (a) Distribution and Gaussian fit of the breakdown times under a bias of 0.1 V. (b) Statistical breakdown times under increasing bias voltage and 1 μm gap.

breakdown time is, under small voltages, well interpolated by a Gaussian fit. Higher voltages favor the emergence of a second maximum, which is expected to reflect the difference between the two statistical times (first avalanche and feedback) mentioned previously. In addition, the results appear physically consistent, as the breakdown time is higher for small voltages while the variance of the system is reduced.

Ultimately, it was remarked that the same statistical time slightly decreases when the electrodes are separated from one another. This result is consistent with our model which does not account for field emission. Hence, wider gaps present a higher resistance, which leads to higher Joule dissipation and a faster heating of the gas. At the atomistic level, this is accounted for by the fact that, statistically, more electrons are initially in the gas to ignite the avalanche process.

4.5 Perspectives

In this chapter, we proposed a novel approach to the simulation of electro-thermal breakdown in micro-fabricated metal switches. Percolation theory was used to account for effects observed in elementary unidirectional elements of gas. The model shows promising results according to the appearance of electrical arcs in the cavity and the good approximation of the statistical distribution of breakdown times. Consequently, we recommend future investigation of electrical breakdown of RF MEMS based on percolation theory. In view of this, the results are expected to be confronted against experiments on real switches. Several refinements may additionally be provided to the model, in terms of diffusion of the ionized defects, non linearity of the current-voltage laws, and inclusion of surface emission from the metal interfaces.

In addition, this model revolves around the integration of a network of unidirectional conduction channels. The device engineer is surely not limited to the choice of an inert gas between electrodes. To exploit the properties of novel meta-materials, it was suggested among other options to use networks of carbon nanotubes to establish a percolative thermal and electrical contact between two electrodes [68, 69]. Our model remains general and we would like to envisage the case where structures of low dimensionality are used in a conductive network system. This requires the prior knowledge of the electrical and thermal behavior of such structures.

Lastly, we have not dealt yet with the conduction problem in beams and cantilevers themselves. Such structures, when scaled down, will be equivalent to nanowires. In this respect, we will focus in the next chapters on the thermal and electrical conduction processes in quasi-1D structures such as nanowires.

CHAPTER 5

ELECTRONIC TRANSPORT IN ROUGH SEMICONDUCTOR NANO-CHANNELS

5.1 Simulation of Electronic Transport in Nanostructures

In the multi-scale framework presented in the introduction (Fig. 1.1), it is required, in the second step, to develop a solid understanding of heat and electronic conduction at the nanometer scale. Our interest in developing efficient models of nanoscale transport is twofold. Firstly, as we are concerned with the large scale integration of conduction channels in percolative networks (see chapter 4) for novel NEMS designs, we must be able to predict the behavior of single elements at the quasi-atomistic scale. Secondly, in the scope of large scale integration the size of more conventional NEMS such as RF switches or cantilever devices needs to be constantly reduced. In this respect, we believe that models of nanoscale transport in silicon beams and cantilevers will be required to allow further technological advances in the near future.

Nevertheless, the problem of nanoscale electronic and thermal transport is not new, and has been addressed very extensively, for instance, in the area of semiconductor field effect transistors (FETs). In this field, the newest technological trends incorporate nanowire channels on the CMOS platform, such as FinFET, trigate or nanowire MOSFET [70, 71, 72, 73]. These structures, although pushing the scale of the devices to a quasi-1D limit, nevertheless suffer from the presence of a considerable surface-to-volume ratio which inevitably tends to degrade carrier mobility at small cross-sections.

The problem of transport in quasi-1D channels is indeed two-fold. On the one hand, the conductivity of devices near the 1D limit benefits from the high intrinsic directionality of such geometries. Consequently, it is profitable to exploit the physical properties of such systems in the mere hope of achieving better performance in terms of thermal and electrical conductivity. On the other hand, conducting nano-channels suffer from the increasing contribution of their interfaces. Indeed, at these scales, the natural geometrical fluctuations of the material surface are expected to strongly interact with
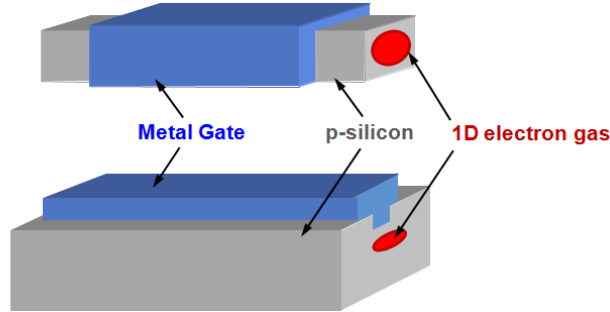
Figure 5.1: Physical models used to study electronic transport in quasi-1D silicon channels. Top: geometry confined 1D electron gas. Bottom: electrostatically confined electron gas.

the wave-like nature of the carriers of heat and electrical currents, in a way which is not unlike the diffraction of optical waves by a real surface.

In order to study these effects, we have chosen example devices operated in a FET way. Indeed, the generally better understanding of the underlying transport processes in FET structures will help grasp the electronic processes involved in quasi-1D transport. In addition, we will be able to easily compare the results of our approach to an extensive amount of available experimental data. Thus, in the approach presented in this chapter, we intend to theoretically predict the concurrent effects of directionality and surface roughness in electrically semiconducting nano-channels. In order to develop a thorough understanding of these processes, we have thought of two sensibly different devices, which both operate near the 1D limit (Fig. 5.1). The first model device represents the classical picture of the nanowire, where a 1D electron gas is confined by the geometry of the device. In order to study the relative effect of surface roughness, our second model device utilizes an applied gate bias to electrostatically confine the electron gas for 1D transport, hence reducing the contribution of interfacial processes.

In this scope, modeling a realistic nanoscale semiconductor device requires choosing which physical picture is most suitable to tackle the problem of interest. At the top of the available simulation hierarchy, one finds quantum models which accurately account for such effects as size quantization, quantum interference, and tunneling. At this level of abstraction, the many-body Schrödinger equation is the governing equation which fully describes the electronic structure of atoms, molecules, and solids. Solving the time-dependent Schrödinger equation for the many-particle system, however, is not an option for a realistic device, due to the large computational effort required to solve an order of 10^{23} coupled equations [74, 75, 76]. Alter-

native approaches to describing non-equilibrium correlated many-particle systems, such as non-equilibrium Green functions (NEGF), are furthermore limited by their inherent difficulty to accurately include scattering mechanisms. When size allows, we will present in chapter 8 a method to include boundary scattering rates in a Geeen function scheme for ultra-thin nanowires.

At a semi-classical level, the Boltzmann transport equation (BTE) is the governing equation which describes electron transport in semiconductors. Although solving the BTE is not a simple task, it is common, under certain conditions of quasi-equilibrium and long channel devices, to use approximate models derived from the moments of the BTE. These models include the hydrodynamic (HD) [77, 78] and the drift diffusion (DD) [79] approach. However, both DD and HD approaches prove inefficient in capturing such short channel effects as velocity overshoot.

As a good trade-off between the full quantum and the DD approaches, the semiconductor Monte Carlo simulation has been intensively used in the past decades to provide a fairly complete semi-classical picture of device physics [80, 81, 82] and successfully describe nonequilibrium effects such as hot carriers, which are common at short channel lengths [83, 84].

In this scope, a quantum corrected three-dimensional (3D) Monte Carlo simulation is a good TCAD candidate for the exploration of hybrid devices based on the semiconductor nano-channel platform. In the following subsections, such an approach is used to investigate the benefits of quasi-1D electronic transport in geometrical silicon nanowires and electrostatically confined silicon channels. The performance of the devices was assessed in the FET structure, and we observed enhanced performance of the confined silicon nano-channels. In this respect, the simulator was subsequently used to optimally size a novel MOS structure at the limit of lateral scalability. A quantum correction approach is adopted to account for size quantization in the narrow channels, which is accurate as long as the cross-section is not too small. For even narrower channels, explicit inclusion of the sub-band structure may be needed.

5.2 Monte Carlo Approach to Device Simulation

In the precise context of semiconductor device simulation, the Monte Carlo technique [85] is used to derive the properties of the device by tracking the movements of the carriers in space and time while they move in the crystal
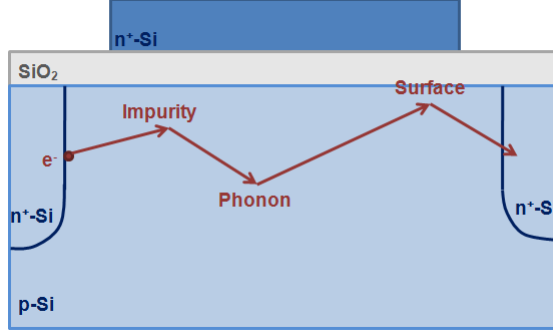
Figure 5.2: Schematic view of the trajectory of an electron during a Monte Carlo simulation of a n-MOSFET.

lattice according to a set of semi-classical rules. Then, a statistical model of the device is built by sampling information about the system. As the number of particles and duration of the simulation increase, the solution for the energy distribution in the device approaches the exact solution of the Boltzmann transport equation (BTE) [86]

$$\frac{\partial f}{\partial t} + \frac{\partial \mathbf{r}}{\partial t} \cdot \nabla_r f + \frac{d\mathbf{k}}{dt} \cdot \nabla_k f = \left.\frac{\partial f}{\partial t}\right|_{coll} \tag{5.1}$$

where $\mathbf{r}$ is the position in the device, $\mathbf{k}$ is the wave vector and $f(\mathbf{r}, \mathbf{k}, t)$ is the distribution function which, in essence, is related to the probability of finding carriers in the state $\mathbf{k}$ and at the position $\mathbf{r}$ at time t. The right-hand side accounts for the variation of the distribution in the infinitesimal region due to scattering and is generally derived quantum mechanically from Fermi's golden rule. Hence, the Monte Carlo approach yields a semiclassical picture which follows two assumptions:

1. The particles in the device are represented as point-like objects.

2. The motion of the particles is described as a series of free flights interrupted by scattering events (Fig. 5.2). During the free flights, the trajectories in real and $\mathbf{k}$-space are calculated from classical mechanics. The scattering mechanism and angle are stochastically determined from quantum mechanics.

The present work is based on the 3D Monte Carlo simulator of the University of Illinois, MOCA. The current version of the simulator originates from the two-dimensional work of A. Pacelli [87] in 1998, modified by B. Winstead [85] for quantum correction and brought to the 3D version by G. Khatawala.

52
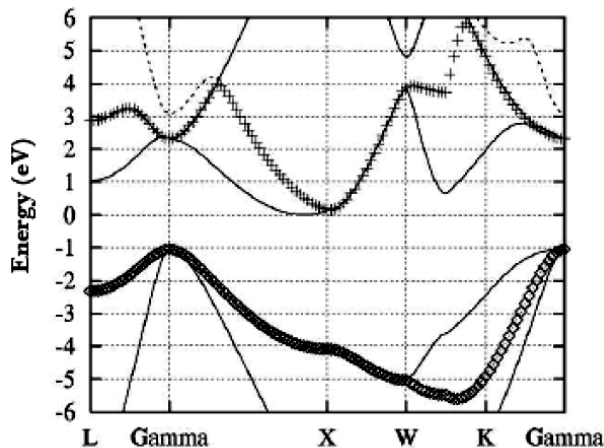
Figure 5.3: Example of band structure for Si used in MOCA. Figure courtesy of [88].

Here, the semiclassical equations describing the motion of electrons are

$$\frac{d\mathbf{r}}{dt} = \frac{1}{\hbar}\nabla_{\mathbf{k}}E(\mathbf{k}) \tag{5.2}$$

$$\frac{d\mathbf{k}}{dt} = \frac{q\mathbf{F}(\mathbf{r})}{\hbar} \tag{5.3}$$

where $\mathbf{F}$ is the electric field and $E(\mathbf{k})$ is the energy dispersion relation. Although a parabolic dispersion relation can often be assumed near equilibrium, the existence of hot carriers requires, in current applications, accounting for the anisotropy of the crystal lattice and using a full dispersion relation in the high energy tails. A full-band $E(\mathbf{k})$ relation can be obtained using the semi-empirical pseudopotential method [82]. An example of such full band structure for silicon is given in Fig. 5.3.

Finally, the electric field $\mathbf{F}(\mathbf{r})$ in the device depends on the electrostatic potential $u(\mathbf{r})$

$$\mathbf{F}(\mathbf{r}) = -\nabla u(\mathbf{r}) \tag{5.4}$$

The internal redistribution of charge carriers traveling in the device is handled by a self-consistent simulation. This approach couples the Monte Carlo procedure to Poisson's equation (eq. 5.5 below) in order to update the internal electrostatic potential.

$$\nabla(\epsilon\nabla u(\mathbf{r})) = -\rho(\mathbf{r}) \tag{5.5}$$

where ϵ is the anisotropic permittivity of the material and ρ is the charge density due to minority and majority carriers as well as ionized dopants.

Since solving Poisson's equation in a 3D self-consistent Monte Carlo requires about 90% of the total computational time [89], efficient numerical schemes must be employed. MOCA, in this scope, solves Poisson's equation with a Newton iteration based on finite differences, using a conjugate gradient technique and a line search to determine the descent direction and the step size [90].

Hence, two of the benefits of a semiclassical Monte Carlo approach in comparison to its DD, HD or quantum counterparts, are its capability to provide accurate quantum mechanical treatment of various distinct scattering mechanisms, and the absence of assumption about the form of carrier distribution in energy or **k**-space.

5.3 Scattering Mechanisms and Mobility

In the Monte Carlo approach, the duration of the free flight of a particle is stochastically determined from a set of scattering rates given for a variety of mechanisms. The present work accounts for acoustic and optical phonon scattering, impact ionization scattering, impurity scattering, surface roughness scattering and the carrier-carrier interaction by using the particle-particle–particle-mesh (P^{3}M) approach of Hockney and Eastwood [91]. Given the different scattering rates for a precise particle state, one may then randomly select the scattering process at the end of the free flight, as well as the scattering angle, through the use of a random number generator.

These scattering rates are very often derived using the Born approximation, in which a scattering event is merely a transition between two momentum states of the carrier involved. The quantum many-body problem arising from the interaction of a carrier with its surrounding environment (phonons, electrons, holes, plasmons, impurities, etc.) may indeed be reduced to a two-body problem using the quasiparticle approximation, which separates the carrier of interest from the rest of the crystal [80]. Within these approximations, Fermi's golden rule gives, to the first order, the transition probability per unit time from a state $|\mathbf{k}\rangle$ to a state $|\mathbf{k}'\rangle$

$$S(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} |\langle \mathbf{k}|H'|\mathbf{k}'\rangle|^2 \cdot \delta(E - E') \tag{5.6}$$

where H' is the perturbation Hamiltonian representing the collision and E and E' are, respectively, the initial and final energies of the system constituted of the carrier, the lattice and the electron gas. The Dirac δ-function stands for the conservation of energy. Under these circumstances, the scat-

tering rate $\tau_i^{-1}(\mathbf{k})$ for the scattering process i is simply the sum of the $S_i(\mathbf{k}, \mathbf{k}')$ over all possible terminal states $\mathbf{k}'$.

The mobility μ of particles in a crystal lattice is a statistical description of how fast a charged particle will travel in response to an applied electrical field. It is well defined in a direction of propagation x

$$\mu_x = \frac{\langle v_x \rangle}{\langle F_x \rangle} \tag{5.7}$$

where v_x and F_x are the average speed and field in the direction of propagation and the averages are, for instance, on the whole conduction channel of a device. Without entering into cumbersome mathematical derivation, let us note here that it is physically meaningful to assume that a high scattering rate, or in other words a greater number of collisional events encountered by a particle traveling from source to drain, will decrease the mobility in the channel of a device.

We would like to emphasize the special case of surface roughness scattering. The asperities of a rough semiconductor-oxide interface may indeed be treated as a perturbation of the potential near the boundary [14, 16]. Carriers may then, instead of specularly bouncing on the surface, scatter diffusively near the boundary, which accounts for momentum —and mobility— loss in the direction of propagation. Therefore, mobility enhancement can be achieved by reducing the surface of contact between the electron gas and the oxide.

5.4   Quantum Correction

As MOSFET devices are scaled down to the order of a couple tens of nanometers, large electric fields are created which reduce the inversion layer thickness to the order of a de Broglie wavelength. At this scale, the allowed energy levels for a carrier become discrete due to the wave nature of the carriers. Hence, in order to successfully describe recent devices, Monte Carlo simulations must account for the quantization in the direction of confinement, which is typically normal to the oxide surface. Consequently, a quantum corrected Monte Carlo algorithm must periodically solve the Schrödinger equation in planes perpendicular to the direction of propagation in order to account for quantization in the electron gas.

At the quantum mechanical level, the sudden change in potential at the Si-SiO$_2$ interface in essence leads to a repulsion of charge density from the
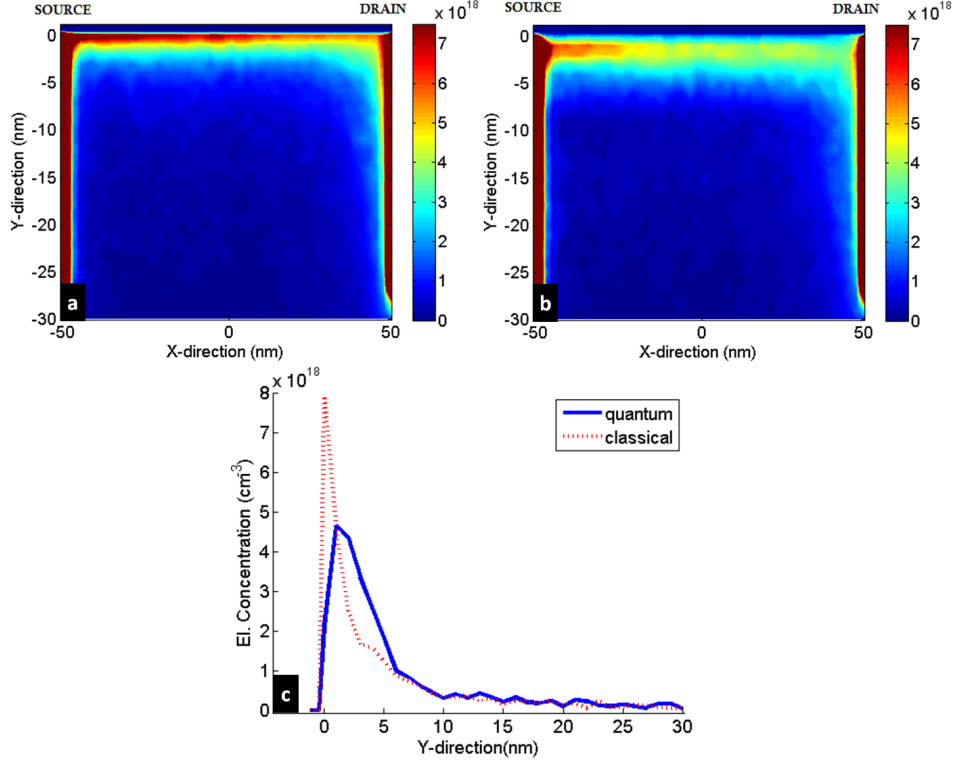
Figure 5.4: Comparison of the carrier distribution (cm^{-3}) in the 2DEG of a MOSFET: (a) without quantum correction, (b) with quantum correction, (c) perpendicular to the electron gas.

interface (Fig. 5.4). MOCA takes this into account by using a quantum correction potential which is linked to the quantum density as

$$V_{qc}(y, z) = -k_B T \log[n_q(y, z) - V_p(y, z) + V_0] \tag{5.8}$$

where $V_{qc}(y, z)$ is the quantum correction potential, $n_q(y, z)$ is the carrier density obtained from solution of the Schrodinger equation, $V_p(y, z)$ is the potential from solving the Poisson equation self-consistently, and V_0 is a reference potential point in the device where the quantum correction is fixed to zero.

5.5   Confined Silicon Nano-Channels

In order to incorporate nanowire structures on the CMOS technological plat-form while limiting the effects of surface roughness scattering, it has been suggested to scale down the width of the gate [92]. Such devices are ap-
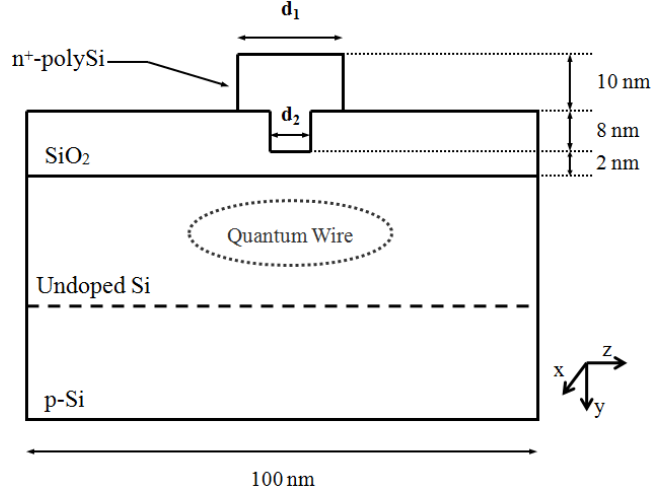
Figure 5.5: Cross section of a T-gate MOSFET perpendicular to the direction of propagation.

pealing in the sense that they achieve electrostatically a lateral confinement of the electron gas in a quasi-1D channel while limiting the Si-SiO$_2$ interface and remaining structurally close to a standard MOSFET. The promises of integrated nanowire-like transistors are manifold. Firstly, lateral downscaling of transistors allows an increase in the density of devices on a chip. More importantly, limiting the charge transport along one direction inherently offers the possibility to reach higher mobilities and subsequent higher on-current to off-current ratios for the next generation devices.

The quantum corrected 3D Monte Carlo simulator described in the previous sections was used to efficiently design quasi 1D nanowires, in which the confined conduction channel is created by etching an oxide trench, thus realizing a T-gate structure (Fig. 5.5). The simulated T-gate structure consists of a 5 nm metal-filled oxide trench (d_2) capped with a 20 nm metal gate (d_1). The silicon oxide thickness was fixed at 2 nm. The source and drain are n$^+$-doped with a concentration of 10^{20} cm^{-3} with the S/D junctions ending abruptly at the gate edge. The silicon substrate was lightly doped to 10^{16} cm^{-3}. Additionally, it has been considered to use an undoped epitaxial layer grown on the silicon substrate, in order to reduce impurity scattering in the conduction channel. It can be seen that, consistently with the coupled Schrödinger-Poisson solver of [92], the Monte Carlo simulation accounts for the confinement of carriers due to the potential created by the gate (Fig. 5.6).

The width of the confining potential is, as expected, strongly dependent on the width d_1 and, to the second order, on the width d_2. While better

57
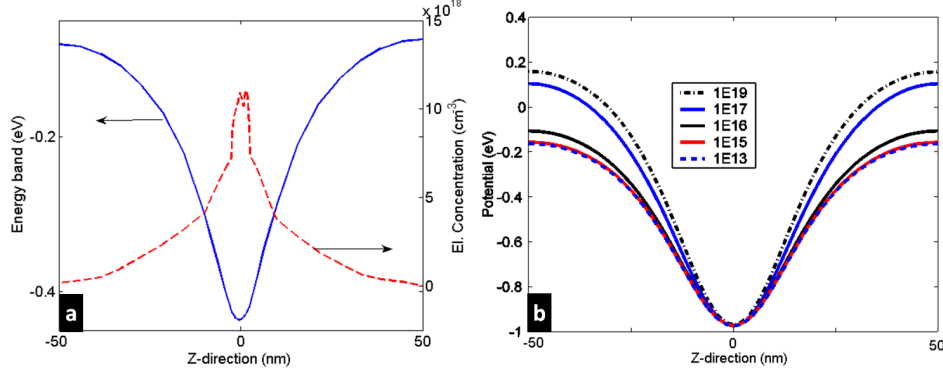
Figure 5.6: (a) Band structure and carrier concentration 2 nm below the oxide with $V_g = 1.5$ V and $V_d = 1$ V. (b) Effect of the carrier concentration under the same conditions, at 0.5 nm below the oxide and no epitaxial layer.

devices could be realized by extreme scaling of the gate width, we have kept these dimensions in the range of what is technically achievable. The choice of substrate doping results from a careful trade-off. On the one hand, as shown in Fig. 5.6, higher dopings reduce the effective width of the confinement. On the other hand, high impurity levels simultaneously increase the band bending Φ_F near the oxide as well as the charge in the depletion region, thus leading to higher threshold voltages. In this scope, better performances are achieved with a 20 nm epitaxial layer of undoped silicon on top of a 10^{16} cm^{-3} doped substrate. Figure 5.7 shows the output I-V characteristic for a T-gate in comparison to a standard planar MOSFET. The latter has a 10 nm thick oxide and 100 nm gate width, which corresponds to removing the etched tip of the former T-gate structure and fully metallizing the oxide. For consistency, the channel length, in both devices, is fixed at 100 nm and both devices are operated under a gate voltage of $V_G = 1.5$ V. The T-gate reaches saturation faster and achieves, at the device level, an increase of $\sim$ 10% in drain current at $V_d = 1$ V. Since all the carriers are concentrated in a channel of width $\sim$20 nm, the current density in the electron gas is increased by a factor of approximately 5. This effect, however, is mostly due to the reduction of the oxide depth to 2 nm at the tip of the T-gate, accounting for a local change in gate capacitance and a subsequent increase of saturation current.

We have assumed so far that the directionality and the lower surface scattering rate account for the higher carrier mobility in a T-gate MOSFET. In order to test this assumption, channel mobility and scattering rates are com-
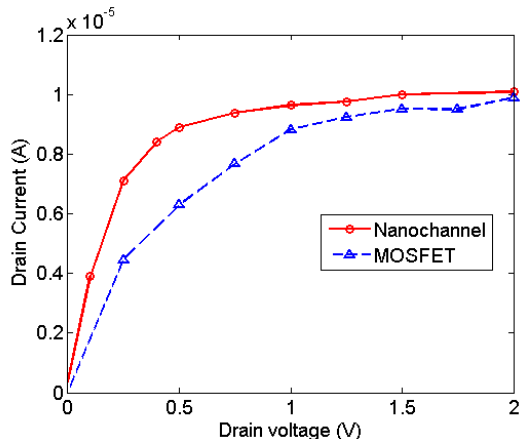
Figure 5.7: Comparison of drain current of a standard MOSFET and a T-gate MOSFET under $V_g = 1.5$ V. The standard MOSFET has a 10 nm thick oxide.

pared between a standard MOSFET, a wrap-around gate nanowire MOS-FET, and a T-gate. The wrap-around gate MOSFET [93, 4] is, in essence, a beam of doped silicon connecting source to drain, wrapped in an oxide and four metallic gates. In this way, the latter device reproduces the confinement of the T-gate MOSFET while keeping a high surface to volume ratio. For coherence with the previous device, we chose a nanowire length of 100 nm, an oxide thickness of 2 nm and a doping of 10^{16} cm^{-3}. In a T-gate, since the carriers are confined in a channel of effective width of $\sim$20 nm, a nanowire cross-section of 20×20 nm has been chosen. In order to study the effect of surface scattering, a scattering ratio is defined for a process i as a direct measure of the proportion of scattering events of type i among all the scattering events. Figure 5.8 shows that the contribution of surface scattering rises with the normal electric field, in accordance with the results of Yamakawa et al. [94]. The slight drop in surface scattering ratio at higher gate voltages of the wrap-around nanowire is expected to be a result of volume inversion. The T-gate is subject to a lesser contribution of surface scattering in comparison to the wrap-around gate, which improves the mobility in the electrostatically confined channel at higher drain voltages (Fig. 5.8). Additionally, it is seen that the channel mobility in both of the quasi-1D structures is increased, mostly because the carriers are forced to flow in the direction of propagation, thus reducing the number of available states after a scattering event takes place.

Figure 5.9 represents a local snapshot of the scattering rates in a cross section of a gate all around (GAA) nanowire device. It can be seen that
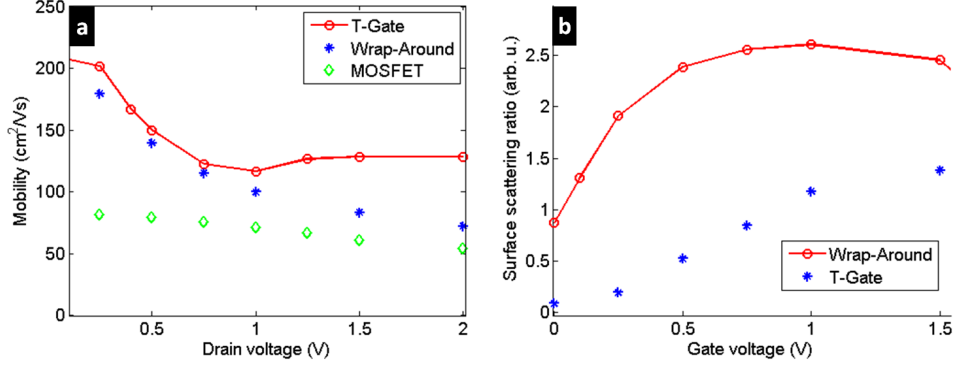
Figure 5.8: (a) Comparison of mobility for different type of MOSFETs under $V_g = 1.5$ V. (b) Comparison of the surface scattering ratio of a T-gate and a wrap-around gate nanowire MOSFETs under $V_d = 1$ V.

our model accounts well for the fact that surface roughness scattering is localized near the Si-SiO$_2$ interfaces. Given the larger surfaces of the GAA in comparison to a T-gate, it is not surprising that surface scattering processes have a stronger effect on conduction in GAA. More importantly, a series of cross sections along the conduction direction reveals that surface scattering rates are higher near the drain where the electron velocities are typically the highest in the channel. This reveals the energy dependence of the surface roughness scattering rate. Indeed, similarly to the optical theory, electrons with high wavelength have a smaller probability of interacting with the small asperities at the silicon interface.

Finally, we wish to determine how the surface scattering rate contributes to the conduction when the nanowire diameter is scaled down (Fig. 5.10). For this purpose, we have tested two GAA of diameter 20 nm and 8 nm under an applied drain voltage of 1 V and varying gate voltages. It is remarkable that at low gate biases, the relative influence of surface roughness scattering rate is 3 to 5 times more pronounced in the 8 nm nanowire. At higher applied gate biases, this difference tends to be reduced, as other scattering processes such as electron-phonon will be more influential in the conduction process. Hence this experiment raises two new challenges. Firstly, it is equally important to have understanding of the transmission of phonons in small nanowires. Secondly, the strong effects of sub-bands below 10 nm must be taken into consideration, which is not the case in the example cited above.

60
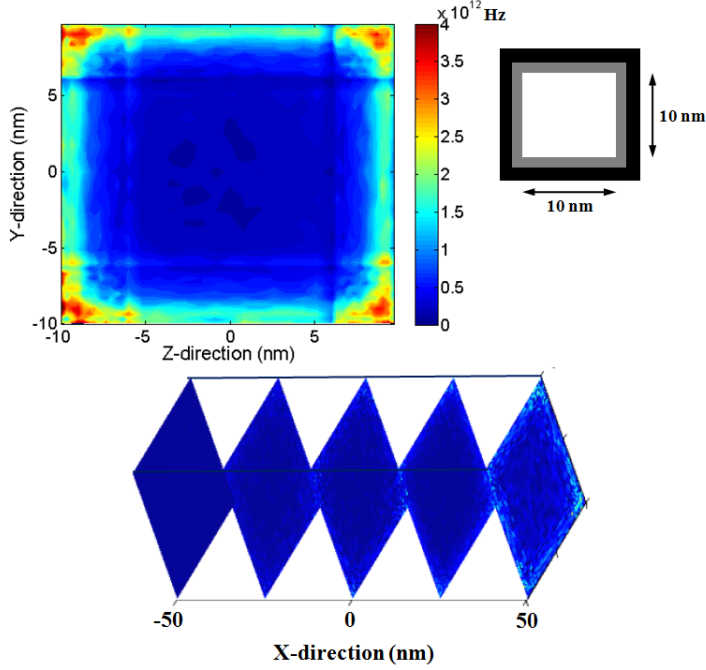
Figure 5.9: Cross section of the surface scattering rate in a geometrically confined nanochannel. The cross section is 20 × 20 nm, under an applied gate bias of $V_G = 1.5$ V and $V_d = 1$ V. Bottom: surface scattering rate along the conduction channel from source (right) to drain (left).

5.6   Conclusions

In this section, we explored the electronic transport in quasi-1D silicon nanochannels using a quantum corrected full-band 3D semiconductor Monte Carlo approach. Based on a study of scattering mechanisms in FET operation of quasi-1D nano-channels, we introduced an innovative way to achieve unidirectional electronic transport through electrostatic confinement while avoiding the spurious contribution of surface roughness scattering. This model device allowed a better understanding of electrical conduction in nanowires, and especially the impact of the geometry of these devices on their electrical performances. In this respect, by limiting the surface roughness scattering rate, the T-gate structure shows an increase in current density by a factor of five in comparison to an equivalent planar MOSFET. In addition, electrostatically confined devices show an increase in mobility in comparison to a planar MOSFET or a geometrically confined nanowire by a factor up to 2.5 at a drain voltage of $V_d = 2$ V across a 100 nm channel. In addition, we examine the physical processes involved in designing beams and wires with increased electrical performances, the latter being achieved through a careful balancing of all scattering mechanisms through confine-
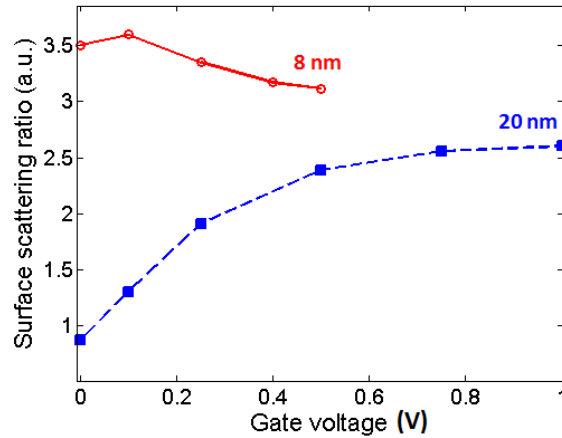
Figure 5.10: Ratio of the surface scattering rate among the other scattering mechanisms for geometrical nanowires of 8 nm and 20 nm diameter.

ment width, doping, and applied electric fields. Importantly, this exploration of design paramters was conducted purely through calibrated simulations, ensuring consistency of the physical approach without incurring the cost of experimental prototyping.

Nevertheless, studying nanowires below 10 nm will undoubtedly involve the inclusion of finer quantum mechanical effects. This will be the object of chapter 7. Equally important is the understanding of heat conduction in these nanowires, which will be the object of chapter 6. As a result, we will aim at establishing a comprehensive model of coupled electro-thermal conduction in quasi-1D nano-channels, which will eventually be integrated in the higher level of our multi-scale simulation hierarchy for the design of innovative integrated NEMS based on the enhanced performance of nanoscale devices.

CHAPTER 6

HEAT TRANSPORT IN ROUGH SEMICONDUCTOR NANO-CHANNELS

6.1   Thermal Properties of Semiconductor Nanowires

In chapter 5, we emphasized electronic transport in quasi-1D silicon nano-channels. In this respect, it was shown that the electrical performance of such conducting channels was mainly limited by the concurrent effects of electron-phonon and surface roughness scattering. In this scope, we discussed the particular case of silicon nanowires, which, over the past 10 years, have drawn much attention for their potential applications in field effect transistors [4], interconnects [5], heterostructures [6], and most recently, thermoelectric applications [7, 8, 9, 10, 11]. Within this area of research, electron-surface interaction was studied early on due to its major impact on conduction in MOSFET channels. Incidentally, the necessity to engineer thermally efficient devices only appeared in the early 2000s, and naturally drew attention to the phonon-surface interaction.

Thermal transport indeed plays a critical role in thermal management of electronic devices as their performance and reliability is a strong function of temperature [95]. More importantly, in the current economic context, scientific and technological innovations are undoubtedly needed in order to address the challenges of energy management. Efficiently controlling energy losses in devices principally involves careful engineering of thermal transport, whether it is in the form of heating, cooling, or waste heat. Such technological breakthrough remains hypothetical, though, with the scientific momentum involved in the field, the answer may be found in integrated nano-systems. In particular, being able to independently control electrical and thermal conductivity of nanostructures through geometry, strain, or doping is extremely appealing for novel applications of thermoelectrics and energy transport. For instance, earlier data showed that reducing the nanowire diameter below 100 nm leads to a drastic reduction in their electronic and thermal conductivity [5, 9, 11]. More puzzling are the recent experimental results of Hochbaum and Boukai [7, 8] which show that inten-

tionally etched rough edges reduce the thermal conductivity of crystalline silicon nanowires by a factor of about 100, to nearly the value of amorphous silicon. Several efforts have previously been made toward an accurate understanding of phonon-surface scattering [10, 11, 6]; however, no study accounts for such experimental observations in very rough wires with diameter below 50 nm. Similarly, no model provides guidance on how thermal conductivity of such nanowires scales with surface roughness.

There are two dominant means of heat conduction in solids: electrons and lattice vibrations (acoustic and optical waves). In good electrical conductors, usually metals such as Cu and Al, electrons are the dominant heat carriers. However, in the case of semiconductors and insulators, lattice vibrations are generally the main vehicle of thermal transport. Similarly to electrons, quanta of lattice vibration may be represented by quasi-particles, called phonons. Therefore, following the ideas developed in chapter 5 on electron-surface scattering, we introduce in this chapter a comprehensive approach to phonon-surface scattering in thin nanowires based on a perturbative treatment of interface roughness. In this respect, we will derive a new matrix element for phonon-surface scattering which is directly related to a parametric description of the roughness of the surface. Based on this perturbative approach, the repercussions of temperature change and surface quality on the phonon-surface scattering rate will be evaluated in silicon nanowires of diameters below 100 nm. Resulting theoretical predictions of nanowire thermal conductivity show unprecedented agreement with experimental values below 50 nm, where the effect of surface roughness is the strongest.

When the characteristic dimensions of asperities at a rough surface come to the order of the phonon or electron wavelength (5-30 Å), it is indeed expected that the surface scattering rate will be altered to reflect the effect of the interface roughness. Such mechanisms are crucial to the understanding of electron transport in transistor inversion layers, where accurate models have been developed based on perturbation theory [12, 13] (see chapter 5). While such formalism exists in the case of bulk phonon transport [96], current models of phonon-surface scattering in nanowires are based on simplified assumptions, most of them using the probability of diffuse scattering as a fitting parameter [5, 10] (see chapter 1 for the method to extract roughness parameters consistent with the following model). Yet, the latter probability can be directly related to physical properties of the interface, which, among other options, may be experimentally observed by means of transmission electron microscopy (TEM). Besides, it seems relevant that the effect of

rough surfaces should be stronger in thin nanowires and should vary with the frequency of incident phonons. We will describe this effect in detail for semiconductor wires and sheets in the following sections.

6.2   Thermoelectric Properties of Nanowires

In the following sections, we model for the first time the remarkably strong effect of surface roughness on the thermal conductivity of Si, Ge, and GaAs nanowires, and graphene nano-ribbons in 20-400 K temperature range, and with transverse dimension scaling from 22 nm to a few microns. In order to support our models below, we derive a novel theoretical expression for phonon - surface roughness scattering rate. This model accurately reproduces experimental observations on rough Si nanowires, which we first reported in [19]. Within this scope, both Chowdury et al. (2009) [97] and Majumdar et al. (2009) [98] foresaw thermoelectric nanowires as a feasible and efficient engineering solution for site-specific on-demand cooling solutions. However, previous experimental devices [99, 100, 97, 98, 101, 102] based on classical high efficiency thermoelectric materials such as Bi_2Te_3 and Sb_2Te_3 faced potential challenges in their integration onto conventional microscale semiconductor platforms. Though not competing with the high thermoelectric performance $ZT > 2$ of these devices, showing that nano-engineered Si, Ge or GaAs nanowires achieve thermoelectric figures of merit close to $ZT = 1$ is nevertheless of considerable importance, given their natural and convenient compatibility with such technologies such as SOI integrated circuits, Ge based high current FETs and strained Si FETs, or the commonly available GaAs based optical applications. In particular, earlier work published in 2008 [7] experimentally proved that efficient thermoelectric figures of merit in rough nanowires were obtained mostly by reduction of thermal conductivity.

Thermoelectricity results from the concurrent transport of heat and electricity in a device that particularly allows heat to be conducted in parallel, while restricting electricity to be conducted in series. Figure 6.1 depicts a general thermoelectric device, where junctions between materials A and B are maintained at temperatures T_1 and T_2. Under an applied temperature gradient between contacts, diffusion of charge carriers occurs, and, in the presence of an electrical short in one of the materials, an electrical potential ΔV builds up such that

$$\Delta V = S_{Seebeck}\Delta T \qquad (6.1)$$
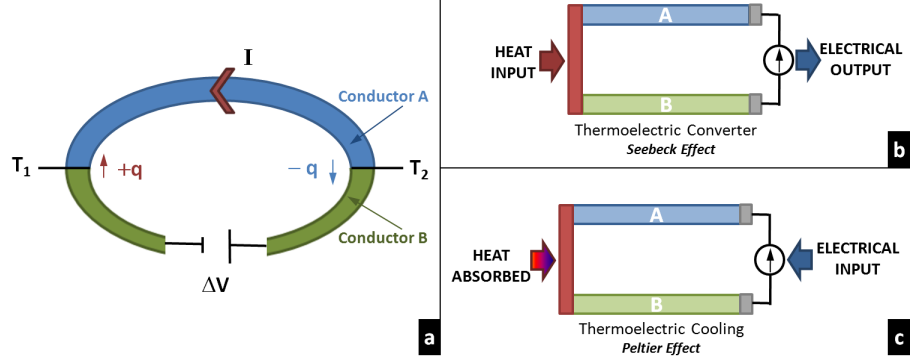
Figure 6.1: (a) Model of the thermoelectric junction, Seebeck, Peltier, and Thomson effects. (b) Model of a thermoelectric converter. (c) Model of a thermoelectric cooler.

This thermoelectric energy conversion is generally known as the Seebeck effect, with S the Seebeck coefficient. Similarly, an applied electrical input will cause the charge carriers to drift and carry heat along their drift path. As a result, heat transfer happens at the contacts, with a transfer rate $+\phi_q$ and $-\phi_q$ at each contact, such that

$$I = \Pi_{Peltier}\phi_q \qquad (6.2)$$

This thermoelectric cooling phenomenon is generally known as the Peltier effect, with Π the Peltier coefficient. The third thermoelectric effect is often of lesser order in most thermoelectrics and occurs when a small temperature differential is applied at the ends of a thermoelectric rod, along with an applied electrical current I. In this case, the rate of heat transfer ϕ_q is given by Thomson's formula

$$\phi_q = \beta I \Delta T \qquad (6.3)$$

where β is Thomson's constant. Such thermoelectric behaviors have been used in the typical configurations illustrated by Fig. 6.1.(b) and 6.1.(c) for thermoelectric conversion of waste heat, or on-site cooling. In this case, the maximum coefficient of performance of the energy conversion is given by [103]

$$\eta_{max} = \frac{T_2}{T_1 - T_2} \frac{\sqrt{1 + ZT} - T_1/T_2}{\sqrt{1 + ZT} + 1} \qquad (6.4)$$

where the product ZT is the figure of merit, and T is the absolute temperature of the thermoelectric device. Hence, optimal performance of the energy conversion is reached by physically increasing the coefficient ZT. The physical expression of ZT results from the thermodynamical laws of drift and

66

diffusion of heat and charge carriers. Most thermoeletrics textbooks include an ab initio derivation of the expression of the coefficient of merit (see for instance Rowe [103]), which physically depends on the ratio between electrical conductivity σ and heat conductivity κ such that

$$ZT = S^2 \frac{\sigma}{\kappa} T \qquad (6.5)$$

Recent efforts in synthesizing thermoelectric materials focused on reducing thermal conductivity. Mainly, this design approach consists in fabricating a *phonon glass – electron crystal*, that is, a material which heats like a glass (low κ), while keeping the excellent electronic transport properties of a crystal (σ). In that respect, nanostructured semiconductor materials present two principal features of interest: they allow decoupling of heat and electrical conduction through phonon decays at their surfaces, and they would find immediate application in micro- and nano-electronics. Thus, nanowires are a privileged candidate, as their shape is the most suited to conventional device geometries represented in Fig. 6.1. With this in mind, we propose in this chapter a study of means of achieving extremely low thermal conductivity in nano-engineered rough Si, Ge and GaAs nanowires, which we thought would draw optimal benefits from the features mentioned above.

While experimental values of Seebeck coefficients of rough Ge and GaAs nanowires are not available to our knowledge, we found from commonly cited sources [104, 105, 106] the bulk values of $S = 680$ V/K, 200 V/K, and 600 V/K for GaAs, Ge, and Si respectively. In the case of Silicon, the experimental observations of Hochbaum et al. [7] clearly show that enhancement of the thermoelectric figure of merit ZT is due to a strong reduction by a factor 10 of the thermal conductivity, while the factor S^2 remains constant under the effect of surface roughness at 300 K in comparison to bulk values. We expect GaAs and Ge to behave quantitatively the same, and predict in the following sections that surface roughness will reduce GaAs and Ge thermal conductivity by factors of 7 and 60 respectively, with an expected increase of thermoelectric figure of merit in the same proportions. In addition, earlier modeling by N. Mingo [107] of the variation of the power factor S^2 and thermal conductivity with the width of smooth nanowires, based on solutions to the Boltzmann transport equation, revealed that improvement of figure of merit in thermoelectric nanowires is due to a reduction of thermal conductivity rather than the power factor. Indeed, improvement of the power factor is theoretically limited since the acoustic phonon-electron

scattering drastically increases when the width of the nanowire is reduced [107, 108]. In particular, models of GaAs nanowires predict a power factor of the order of 0.015 W/m/K^2 for diameters D in the 30-50 nm range [107]. The thermal conductivity of 4.5 W/m/K predicted in the case of rough 56 nm GaAs nanowires in section 6.4 leads to expectations of the figure of merit of approximately $ZT \approx 1$. We would like to point out that the main advantage of using surface roughness as a heat resistive process originates from the fact that charge carriers are naturally pushed away from interfaces and see less contribution from surface roughness than heat carriers, which allows decoupling of thermal conductivity from electrical conductivity. As we consider here intrinsic materials, the effect of dopants is not accounted for, though we include the effect of isotopes on the power factor, which similarly reduce thermal conductivity while leaving electrical mobility unchanged.

6.3 Model of Phonon-Surface Roughness Scattering in Thin Si Nanowires

In this section, we will develop our new model for a frequency dependent treatment of phonon-surface roughness scattering. Based on this model, we will reflect extraordinarily low thermal conductivity of silicon nanowires, in accordance with recent experimental observations [7]. To our knowledge, Si is the only material for which experimental data on rough nanowires is available thus far. Consequently, in this section, we will validate our model on Si nanowires. In a thin nanowire, variations of the confinement width perpendicular to the propagation direction influence phonon transport by perturbing the Hamiltonian of the system (Fig. 6.2 (a) inset). Here, it is assumed that boundary scattering is mainly an elastic process and no phonon is emitted into the surrounding environment. This condition reflects the case where the nanowire is wrapped in a medium of considerably different thermal conductivity, as it is for Si nanowires in SiO$_2$ or vacuum. We focus on modeling phonon transport in such Si nanowires, and of diameter below 115 nm comparable to experimentally available data. We introduce a new type of phonon scattering originating from the roughness of the nanowire surface. In essence, this scattering mechanism accounts for the fact that phonons "see" a rough nanowire as a series of constrictions along their propagation direction. In order to accurately model this effect at the nanometer scale, perturbation theory is used to derive the transition probability per unit time of an incident phonon of momentum $\mathbf{k}$ and energy $\hbar\omega$ to a new
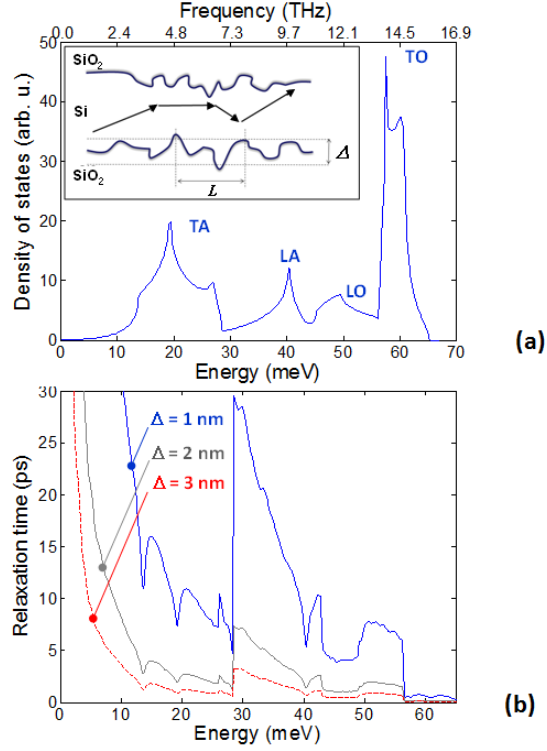
Figure 6.2: (a) Phonon density of states for silicon computed from a full phonon dispersion relation. (b) Phonon relaxation time due to interface roughness for rms $\Delta = 1$ nm, 2 nm, and 3 nm, $T = 300$ K, and a cross section equivalent to a circle of 115 nm diameter. L is fixed to 6 nm.

state of momentum $\mathbf{k}'$ and energy $\hbar\omega'$ due to the perturbed Hamiltonian H' [109]:

$$
\begin{aligned}
P(\mathbf{k},\mathbf{k}') &= 2|\langle\mathbf{k}|H'|\mathbf{k}'\rangle|^2 \frac{d}{dt}\left\{\frac{1-\cos[(\omega'-\omega)t/\hbar]}{(\omega'-\omega)^2}\right\} \\
&= \frac{2\pi}{\hbar}|\langle\mathbf{k}|H'|\mathbf{k}'\rangle|^2\theta(\omega'-\omega)
\end{aligned} \tag{6.6}
$$

Under a sufficiently long time t in comparison to the energy relaxation time, $\theta(\omega)$ reduces to the Dirac delta function. The roughness of the interface is considered as a space varying dilation $\Delta(\mathbf{r})$ of the wire. This alters frequencies in a plane perpendicular to the propagation direction in such a way that $\omega(\mathbf{k}) = \omega_0(\mathbf{k})[1-\gamma\Delta]$ where γ is a fitting constant determined from the thermal expansion of the material, and $\omega_0(\mathbf{k})$ is the phonon dispersion of the unperturbed Hamiltonian. Following the derivation of Klemens [96],

the matrix element for a perturbation due to a space varying dilation is

$$|\langle \mathbf{k}|H'|\mathbf{k}'\rangle|^2 = \frac{4\gamma^2}{3V_{ol}}\omega'^2(\langle n'\rangle + 1)\Delta(\mathbf{k} - \mathbf{k}') \tag{6.7}$$

where V_{ol} is the volume of the device, and $\Delta(\mathbf{q})$ is the Fourier transform of the spatial perturbation

$$\Delta(\mathbf{q}) = \int \Delta(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}\,d\mathbf{r} \tag{6.8}$$

Additionally, the occupation number is given by the Bose-Einstein distribution and includes the temperature dependence of the scattering process $\langle n\rangle = \left(e^{\hbar\omega/k_BT} - 1\right)^{-1}$.

Following the statistical models derived in chapter 2, and as shown by Goodnick et al. [15], the autocovariance function of Si surface roughness is "roughly" fit by either a Gaussian function or a Lorentzian function, which, by the Wiener-Khinchin theorem, respectively yield a power spectrum of

$$\Delta_G(\mathbf{q}) = \pi\Delta^2 L^2 e^{-q^2L^2/4} \tag{6.9}$$

$$\Delta_L(\mathbf{q}) = \frac{\pi\Delta^2 L^2}{(1 + q^2L^2/2)^{3/2}} \tag{6.10}$$

where Δ is the root-mean-square (rms) value of the roughness fluctuations and L is the autocovariance length, which, in essence, is related to the mean distance between roughness peaks at the Si-SiO$_2$ interface (see Fig. 6.2 (a) inset). In practice, these process-dependent parameters are experimentally set by the quality of the surface. It is worth noting here the frequency dependence of the matrix element. Due to the ω'^2 term in equation 6.7, low frequency phonons see little contribution from the surface perturbation. On the other hand, the power spectrum of equation 6.10 favors scattering processes in a certain frequency range. Indeed, as illustrated by Fig. 6.3, increasing correlation length L (e.g. smoother surface), favors scattering processes of the specular type. It is also worth noticing that the Lorentzian approximation, which is believed to be more accurate for the Si-Si0$_2$ interface, similarly favors specular scattering over the Gaussian model. As a result, for comparable scattering rates, we can expect the Lorentzian model to exhibit higher roughness rms Δ than the Gaussian model.

Incidentally, the phonon scattering rate from a branch i to a branch j is given by

$$\tau_{i,j}^{-1}(E) = \int P_{i,j}(\mathbf{k}, \mathbf{k}')\,d\mathbf{k}' \tag{6.11}$$
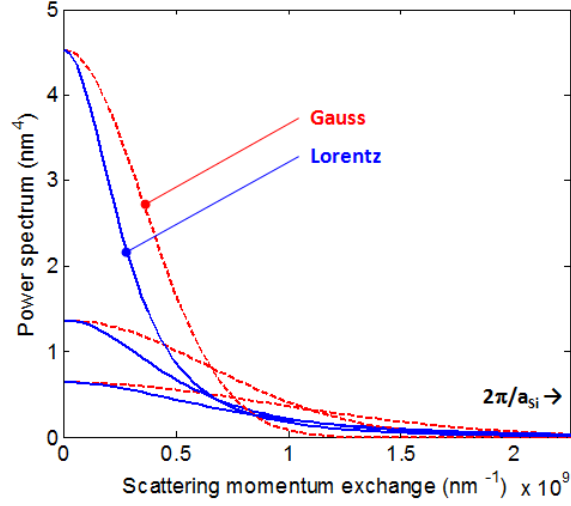
Figure 6.3: Power spectrum of the silicon surface roughness as a function of the surface scattering momentum exchange $\mathbf{q}$. The surface rms Δ is of 3 Å. Values of the correlation length are, from bottom to top at the origin, $L = 1.5$ nm, $L = 2.2$ nm, $L = 4.4$ nm.

The complexity of the volume integration over $\mathbf{k}'$-space can be relieved by reducing the expression to a surface integral as shown in [110]:

$$\tau_{i,j}^{-1}(E) = \frac{2\pi}{\hbar N_i(E)} \int_{E'=E_i} \frac{|\langle \mathbf{k}|H'|\mathbf{k}'\rangle|^2}{\nabla_{\mathbf{k}'} E'(\mathbf{k}')} \, dS \qquad (6.12)$$

where $N_i(E)$ is the phonon density of states in the i_{th} branch, and $E'(\mathbf{k}')$ goes along the j_{th} branch. The total scattering rate $\tau_i(E)$ starting in branch i is the sum over all branches j of the $\tau_{i,j}(E)$.

A Gilat-Raubenheimer scheme [111] is used to compute surface integrals, which constitutes the optimal trade-off between accuracy and computational efficiency. In order to carefully account for frequency dependence, a full phonon dispersion is used, which is obtained from an adiabatic bond charge model and tabulated for look-up [112, 113]. The Gilat-Raubenheimer (GR) method is similarly applied to compute the phonon density of states based on the dispersion relation mentioned above. In essence, the GR scheme computes surface integrals such as the one encountered in equation 6.12 by dividing the first Brillouin zone (FBZ) in small cubes. In each cube, the energy gradient $\nabla_{\mathbf{k}'} E'(\mathbf{k}')$ is precomputed, and the surface element $dS = S(E)$ is approximated to the first order by a plane representing the intersection of the energy isosurface (see Fig. 6.4) with the elementary cubes. The contribution to the integral expression of each cube is subsequently added.
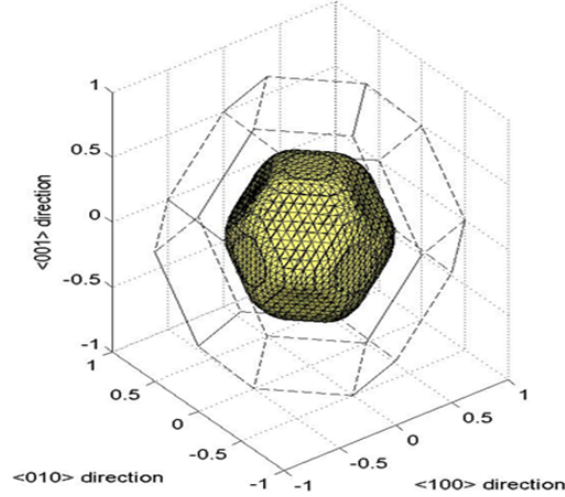
Figure 6.4: Energy isosurface at 29 meV computed in the first Brillouin zone for the longitudinal acoustic phonon modes of Si used in the Gilat-Raubenheimer scheme. The isosurface is directly obtained from the full-phonon dispersion relation. Picture courtesy of [113].

The energy isosurface is directly obtained from the full phonon dispersion relation.

In this scope, modeled silicon nanowires have a square cross section whose area is equivalent to a circular cross section of diameter $D < 115$ nm, while the nanowire length is arbitrarily fixed to 2 μm. In this respect, the Gilat-Raubenheimer scheme divides the first Brillouin zone in a lattice of $40 \times 40 \times 40$ cubes. Transitions among all acoustic and optical branches are considered.

The scattering rates are first computed for nanowires of equivalent diameter $D = 115$ nm at $T = 300$ K. For a fixed correlation length $L_0 = 60$ Å, phonon lifetime is calculated for increasing Δ in Fig. 6.2(b). At this point, let us emphasize again that no reliable data on the roughness of small wires is available yet. Indeed, three main reasons lead to a difficult experimental observation of nanowire surface roughness. First, small nanowire growth and etching techniques cause the nanowire width and roughness to vary along the length of the device. Second, experimental observation techniques at the scale of the thinnest nanowires lead to imprecise measurements and statistical averages on the small nanowire surfaces. Finally, the definition itself of surface roughness becomes unclear when the nanowire edges tend to be curved. Various theoretical and experimental studies have nevertheless reported roughness rms ranging from 3 Å to 5 nm in the case of extremely rough nanowires [7, 114]. In this study we consider effective values for Δ in

this range, and use the 3 Å value for "smooth" nanowires, while using an average of 3 nm for "rough" nanowires. Under the conditions cited above, the average phonon lifetime due to surface roughness alone is approximately 15 ps, and decreases with higher roughness rms values as Δ^{-2}. As deduced from equation 6.10, a long correlation length L favors scattering processes close to the specular type. While in the strong roughness limit where $\Delta/L > 1$ the Gaussian approximation may be put at fault, the effect of the L^2 term tends to average out the contribution of L. We noticed only little deviation of the predicted thermal conductivity in the strong roughness limit, and consistently used a value of $L = 6$ nm, which is estimated from the TEM images of Hochbaum et al. [7] and provides a best fit in our case.

In our perturbative approach, an additional thermal variation of the phonon-surface scattering rate appears, in essence, from the frequency dependent model which was adopted. Thus, the temperature delimits the occupation of each frequency range as predicted from equation 6.7. Hence, the effect of the occupation of lower energy branches at low temperature is retrieved in our model of surface roughness scattering. Subsequently, it is possible to determine the thermal conductivity for nanowires of different cross sections [115]. The contribution to the thermal conductivity of phonon branch i is

$$\kappa_i(T) = \frac{1}{3} \int_i E N(E) \frac{d\langle n \rangle}{dT} v_i(E) \tau_i^{tot}(E) \, dE \qquad (6.13)$$

where v_i is the velocity of sound which is dependent on the direction of propagation, here assumed to be in the $< 001 >$ direction. In order to reproduce the measured physical behavior of the nanowire in the 10-350 K temperature range, Umklapp, normal, impurity, boundary, and surface roughness scattering mechanisms have been considered in the derivation of the branch-specific scattering time, as summarized in Table 6.1. Umklapp scattering in transverse acoustic branches is efficiently described with the law derived in [116], which has shown good agreement with Si nanowire experiments. Additionally, in the temperature range considered, normal scattering in longitudinal acoustic branches is accounted for according to the derivation of Holland [117]. A $C\omega^4$ law is often used in the literature for impurity scattering, and we found that a constant $C = 8 \times 10^{-45}$ s/K^4 fits experimental data more accurately than the value of 1.05×10^{-44} s/K^4 resulting from the derivation of [117]. In the case of nanowires etched from a 10 Ω-cm wafer [7], the low doping concentration has almost no impact on the value of C. It is also worth noting that the boundary scattering rate, which depends on the sound velocity in a given branch, is effectively frequency

Table 6.1: Summary of the scattering processes considered in the derivation of the thermal conductivity.

Mechanism	Analytical Model	Constants
Phonon-Phonon:		
Longitudinal [117]	$B_L T^3 \omega^2$	B_L=2×10^{-24} s/K^3
Transverse [116]	$B_T T e^{-\theta_1/T} \omega^2$	B_T=1.73×10^{-19} s/K
		$\theta_1 = 137.3$ K
Impurity [117]	$C\omega^4$	C=8×10^{-45} s/K^4
Boundary	$v_i D^{-1}\sqrt{1+4\omega c_i/v_i}$	v_L=9.01×10^3 m/s
[115, 118]		v_T=5.23×10^3 m/s
		c_L=-2×10^{-7} m/s^2
		c_T=-2.26×10^{-7} m/s^2
Surf. Roughness	cf. Eq. 6.12	Δ= 3 - 50 Å
		$L = 60$ Å

dependent. As a result, the present work uses an analytical expression for the sound velocity in acoustic branches, which is derived from [118] for its strong fit with bulk Si data in the $< 001 >$ direction.

Figure 6.5 (a) and 6.5 (b) compare thermal conductivity computed from the model presented above vs. experimental data from [9] and [7], for nanowire ranging from 115 nm to 22 nm diameter. As nothing guarantees that all nanowires have similar roughness parameters, we considered smooth nanowires with 1 Å $< \Delta <$ 3 Å, and rough nanowires with 3 nm $< \Delta <$ 3.25 nm. A good fit is found for smooth nanowires grown by vapor-liquid-solid (VLS) mechanism with diameter above 37 nm. Similarly, the model reproduces the drastic decrease in thermal conductivity for rough electroless-etching (EE) nanowire presented in [7]. Higher discrepancy is found for the 22 nm smooth VLS nanowire, for which the sensitivity to surface roughness is expected to be higher. Besides, the perturbative approach remains valid as long as perturbations remain small in comparison to the total phonon energy. For low temperature phonons and nanowires of diameter below 20 nm, explicit quantum treatment may be required. Since surface roughness scattering has little impact on low energy phonons, additional low temperature discrepancy is attributed to impurity and classical boundary scattering.

It is important to point out that our approach based on perturbation theory introduces a dependence $(D/\Delta)^2$ of the nanowire thermal conductivity, in contrast to the typical linear scaling with D used in previous descriptions. In this scope, Fig. 6.6 (a) and 6.6 (b) show how the thermal conductivity is lowered by concurrent effects of small diameter and rough surfaces. In par-
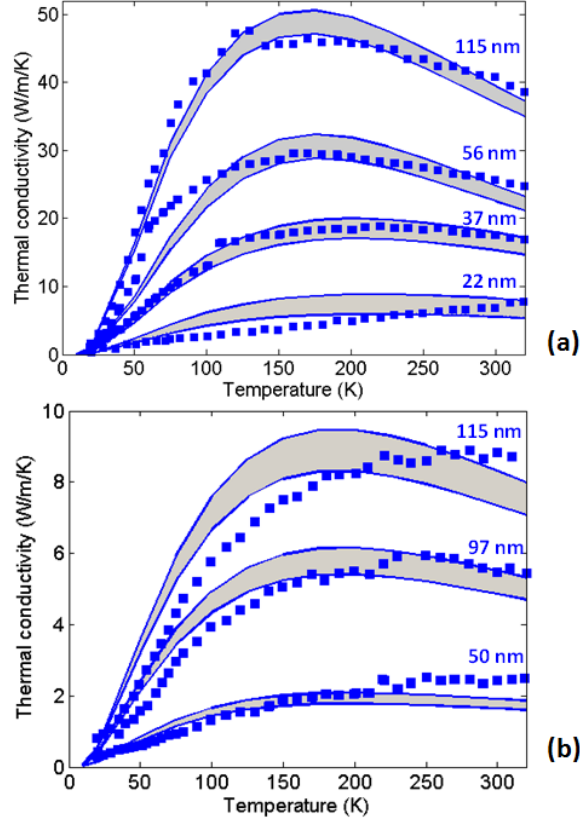
Figure 6.5: (a) Thermal conductivity of smooth VLS Si nanowire. Shaded areas are theoretical predictions with roughness rms. $\Delta = 1$ - 3 Å, blue squares are taken from [9]. (b) Thermal conductivity of rough EE Si nanowire ($\Delta = 3$ - 3.25 nm). Squares are taken from [7]. Simulation and experimental data are compared at similar cross sections. $L = 6$ nm.

ticular, there exists a critical diameter below which the roughness-limited thermal conductivity of the nanowire noticeably deviates from the classical linear approximation. The fact that this critical diameter increases with higher roughness rms Δ supports, in turn, the expectation that heat conduction at small nanowire scales is strongly limited by their surface roughness. This assumption is further justified in Fig. 6.7, where it is observed that the total contribution of surface roughness to limiting the thermal conductivity is increased about 7 times from the 115 nm to the 22 nm case. Finally, the temperature dependence that arises from the roughness scattering rate is also retrieved. Thus, at low temperatures, low frequency phonons see little effect from the perturbations at the silicon interface.

In addition, it is worth pointing out that the previous model was based on a Gaussian approximation to the surface roughness power spectrum $\Delta(\mathbf{q})$,
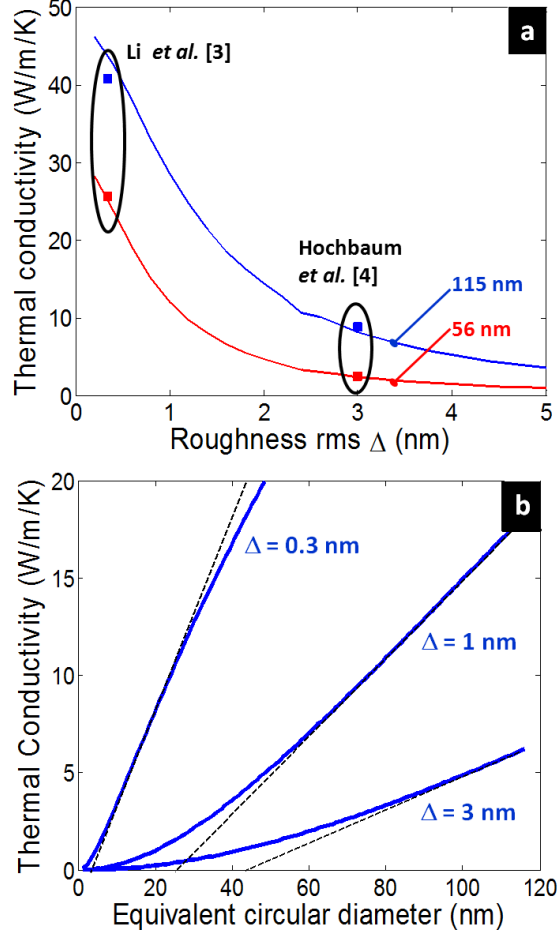
Figure 6.6: (a) Predicted effect of roughness rms on the thermal conductivity of 115 nm and 56 nm nanowire at $T = 300$ K, $L = 6$ nm. Simulation and experimental data are compared at similar cross sections. (b) Effect of nanowire equivalent circular diameter on thermal conductivity ($L = 6$ nm).

and phonon-phonon and phonon-impurity scattering parameters borrowed from the literature. While this model accounts for the physics of the problem with sufficient accuracy to reproduce very well the results observed experimentally, some of the parameters used may falsely represent the behavior of our precise case. Indeed, variation of the confinement width will also have an impact, albeit to a higher order, on phonon-phonon and phonon-impurity interactions. More importantly, the small scales involved also impose boundary conditions on the transverse energy modes which do introduce, to a minor extent, band-folding and the appearance of sub-bands in the phonon dispersion relation. All these effects, which arise from the quasi-1D nature
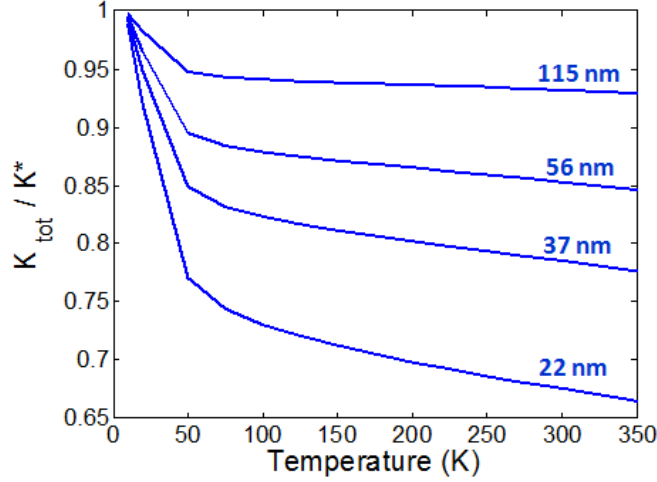
Figure 6.7: Proportional contribution of the surface roughness scattering term to the thermal conductivity of VLS nanowires ($\Delta = 3$ Å, $L = 6$ nm). The reference value K^* is the thermal conductivity of a perfectly smooth wire ($\Delta = 0$ nm).

of the devices in our scope, will slightly alter the parameters used in the previous model, which, in the literature, were fitted on different cases.

Hence, considering that the nanowires which we examined remain intrinsically 3D in nature, and that the effect of sub-banding remains minor, it is meaningful to address such adjustments in our model by finding a best fit to the parameters used instead of including directly the sub-bands. Based on the experimental data $\kappa_{exp}(T, D_i)$, and the model data $\kappa_M(T, D_i, C, B_T, \theta_1, \Delta)$, the parameters of the model are adjusted by a least square method to find what set of parameters achieves the optimal accuracy by minimizing the objective function

$$f(C, B_T, \theta_1, \Delta) = \sum_{T, D_i} \left| \frac{\kappa_{exp}(T, D_i) - \kappa_M(T, D_i, C, B_T, \theta_1, \Delta)}{\kappa_{exp}(T, D_i)} \right|^2 \quad (6.14)$$

The minimum of the least square objective function is found by a simplex method [119]. Since the experimental data on EE rough nanowires is believed to be more sensitive, the protocol used first explores the 4 parameter space for only smooth VLS nanowires to find the optimal C, B_T, and θ_1 and "smooth" Δ. Only then do we determine the optimal "rough" Δ by exploring the 1 parameter space for EE nanowires and assuming that C, B_T and $theta_1$ remain unchanged. Finally, it must be observed that the simplex method will only find a local minimum to the objective function. However, by providing the algorithm with a starting point based on the ob-
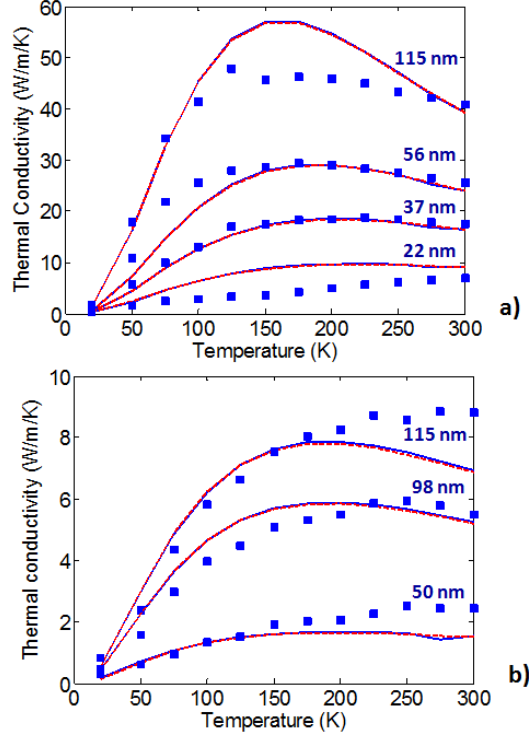
Figure 6.8: Best fit of the model on the experimental data found by a least square approach. (a) Smooth VLS nanowire, (b) rough EE nanowire. The respective parameters are given in Table 6.2. Solid line: Lorentz approximation, dashed line: Gauss approximation.

servations made in table 6.1, and given the form of equation 6.14, we can legitimately assume that the point of convergence is effectively the best fit. Figure 6.8 shows the result of such a best fit for both smooth and rough nanowires. A discrepancy still remains on the 22 nm case, where the effect of sub-bands is expected to be the strongest. Better agreement is found for the EE wires. The Lorentzian approximation provides slightly better accuracy in this case. Table 6.2 summarizes the fitted parameters. As a final remark, we would like to mention that we introduced a new method for the surface engineer to extract the shape factors (L, Δ) of silicon nanowires by means of a non-destructive and thermal test.

6.4   Thermal Conductivity of Rough Ge and GaAs Nanowires

In the previous section, we derived a new model of phonon - surface roughness interaction which accurately accounts for experimental observations of

Table 6.2: Summary of the fitted parameters for estimation of scattering rates in nanowires.

Model Parameter	Best Fit (Lorentz)	Best Fit (Gauss)
C (s/K^4)	1.748×10^{-45}	1.68×10^{-45}
B_T (s/K)	5.58×10^{-19}	5.70×10^{-19}
θ_1 (K)	156.51	157.72
Δ_{VLS} (m)	4.11×10^{-10}	3.45×10^{-10}
Δ_{EE} (m)	3.60×10^{-9}	3.422×10^{-9}

particularly low thermal conductivity in silicon nanowires. As discussed in section 6.2, designing a phonon glass electron crystal is of tremendous importance for the emergence of efficient thermoelectric devices. In this scope, silicon nanowires naturally find applications in on-site cooling or energy conversion for silicon based integrated circuits and micro-processors. Nevertheless, the theory developed above is general to any semiconductor crystal, as long as one is able to measure or compute the material specific physical constants of the model.

In this section, we model the thermal conductivity of rough Ge and GaAs semiconductor nanowires and compare it to the results on Si that we derived above. Here also, we base our derivation on the material specific full phonon dispersion relations. Under this assumption, the effect of nanowire surface roughness on thermal conductivity is derived from perturbation theory, and a phonon-surface roughness scattering rate is derived for Ge and GaAs nanowires with diameter $D < 200$ nm. We show that for those materials, in thin wires with large root-mean-square (RMS) roughness Δ, thermal conductivity is limited by surface asperities and varies quadratically as $(D/\Delta)^2$. At room temperature, this model previously agreed with experimental observations of thermal conductivity down to 2 W/m/K in rough 56 nm Si nanowires with $\Delta = 3$ nm. Based on the latter assumptions, the identical physical model predicts remarkably low thermal conductivity in Ge and GaAs nanowires of respectively 0.1 W/m/K and 0.4 W/m/K at similar roughness and diameter.

In semiconductor nanowires, scattering of carriers at the surface limits transport in devices used for FET, heterostructures, and interconnects [4, 5, 6, 108, 9, 7, 120, 107], owing to their high surface-to-volume ratio. In the section above, we discussed recent experimental and theoretical work

[7, 19, 20] which suggested that potentially high thermoelectric figures of merit ZT could be achieved by scattering phonons more effectively than electrons at the interfaces of rough nanowires. In particular, thermal conductivity in Si nanowires with etched rough edges was experimentally reduced by a factor of about 100 in comparison to bulk crystalline Si to nearly the value of amorphous Si, yielding a subsequent $ZT \approx 0.6$ at room temperature [7]. Such nanostructured semiconductor materials present the advantage to enable energy efficient site-specific and on-demand energy cooling solutions [99, 100], and overcome the difficulties of integration into traditional microscale electronic devices [98, 97]. In this respect, Ge or GaAs rough nanowires could provide convenient on-site energy conversion solutions for high current FET and optical applications based on Ge, GaAS or strained Si on Ge technologies. In comparison to Si, Ge and GaAs have a naturally lower bulk thermal conductivity of ~59 W/m/K and ~42 W/m/K respectively, and higher bulk electrical conductivity [121, 122, 123], and similar Seebeck coefficients, hence suggesting that rough Ge and GaAs nanowires should have thermoelectric efficiency comparable to that of Si nanowires. In previous Bi_2Te_3/Sb_2Te_3 superlattice solutions, increases in ZT mainly originated from a reduction of the lattice thermal conductivity [101, 102]. In semiconductor nanowires, charge carriers are naturally pushed away from the surface. Consequently, specular surface scattering is expected to reduce thermal conductivity in greater proportion than electrical conductivity and Seebeck coefficient. Hence, in rough Si nanowires, it was experimentally shown that the thermoelectric power factor remains high despite reduction of thermal conductivity [7]. In Ge and GaAs rough nanowires, under the assumption that the power factor remains similarly unchanged by surface asperities, low thermal conductivity could have interesting repercussions for enhanced thermoelectric efficiency. In order to explore possible designs of efficient embedded on-site thermoelectric nanowire cooler or generators, this section computationally addresses options in nanowire surface nanostructuring in order to optimize trade-offs between different scattering mechanisms in Si, Ge, and GaAs, and subsequently achieve low thermal conductivity.

As confirmed in experiments on Si nanowires, it is expected that vibrations of the crystal lattice interfere with the spatial fluctuations of the nanowire boundaries. Nevertheless, most of the work in phonon transport only accounts for surface roughness scattering by a constant fitting parameter which reflects an average probability of diffuse scattering at the nanowire surface. We recently proposed [19] a perturbative approach to phonon – surface roughness scattering, where a matrix element is derived for such interac-

tions. This model has shown excellent agreement to experimental observations on rough Si nanowires with diameter below 50 nm where the effect of surface roughness is the strongest. Here, we use our modeling platform to investigate the case of GaAs and Ge nanowires. The roughness of a surface is generally quantified by its RMS height Δ and auto-correlation length L. The present model accounts for the variation of nanowire thermal conductivity with the experimentally observable Δ, L, and the phonon angular frequency ω; this approach accurately predicts phonon-surface roughness scattering to contribute to the reduction of thermal conductivity by a factor $(D/\Delta)^2\omega^{-2}$ where D is the nanowire diameter. In addition, the quality of thermoelectric devices is measured by the dimensionless figure of merit $ZT = (S^2\sigma/\kappa)T$, where S is the Seebeck coefficient, T the temperature, σ the electrical conductivity and κ the thermal conductivity. In this respect, a standard approach to achieve performance is to design a "phonon glass - electron crystal," with high electron mobility and poor thermal conductivity. Among thermally resistive mechanisms in a semiconductor crystal, scattering of phonons by isotopes [121] or in three-phonon anharmonic decays [96] is also of considerable importance.

Our model assumes that no phonon is emitted to the surrounding environment. This condition reflects the case where nanowires are wrapped in a medium of considerably different κ, as is the case for semiconductor nanowires wrapped in oxide or suspended in vacuum. Despite neglecting effects of natural oxide layers present in the case of a suspended nanowire, the assumption above yielded accurate predictions in previous Si models [19]. In addition, considering higher order phonon decays in such peripheral oxide layers of poor κ [124] will only serve to reduce the exceptionally low Ge and GaAs nanowire thermal conductivities reported here. While propagating along wires of diameter $D < 500$ nm, phonons scatter from a series of constrictions along the transport direction which result from the surface asperities and are reflected by a perturbed Hamiltonian H' of the system. The interface roughness is treated as a space varying dilatation $\delta(r)$ of the wire, whose auto-correlation function is assumed to be exponential $\Delta(r) = \Delta\exp(-r/L)$. Following the derivation of Klemens [96], the matrix element for a perturbation due to a space varying dilation reflects the transitions of phonons from a state of momentum $\mathbf{k}$ to $\mathbf{k}'$ such as

$$|\langle\mathbf{k}|H'|\mathbf{k}\rangle|^2 = \frac{4\gamma^2}{3V_{ol}}\omega^2(\langle n'\rangle + 1)\Delta(\mathbf{k} - \mathbf{k}') \qquad (6.15)$$

where V_{ol} is the volume of the device, γ is Grüneisen's parameter, ω is the
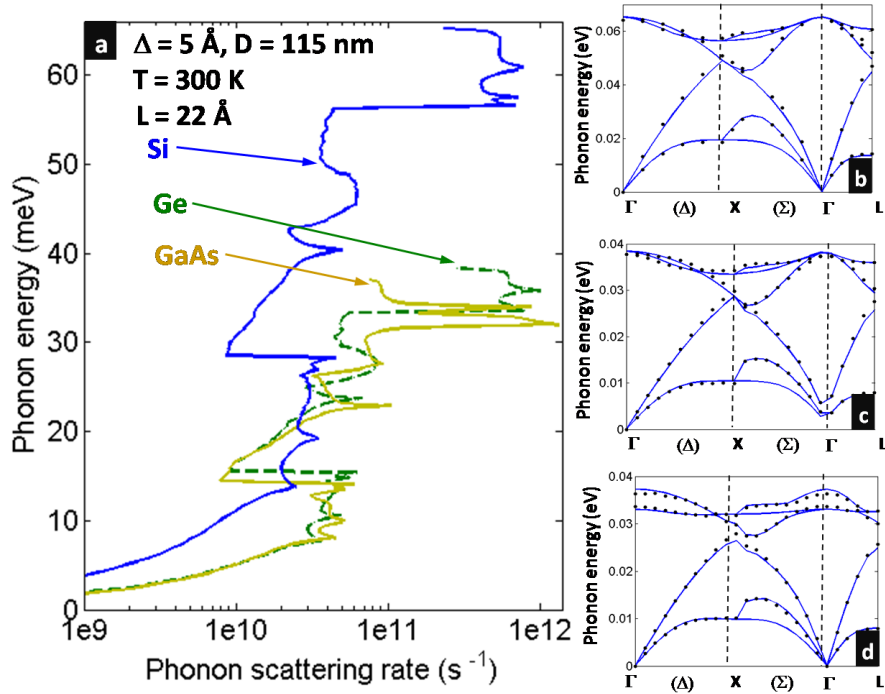
Figure 6.9: (a) Phonon – surface roughness scattering rates computed at room temperature in Si, Ge, GaAs. (b),(c),(d) Computed full phonon branches in the FBZ of Si, Ge, and GaAs respectively. Dots are extracted from experimental observations [125, 126] for comparison.

angular frequency of the incident phonon, and $\Delta(\mathbf{q}) = \pi\Delta^2 L^2 \exp(-L^2\mathbf{q}^2/4)$ is the roughness power spectrum associated to $\Delta(r)$. Close to equilibrium, the occupation of the destination phonon branch $\langle n' \rangle$ is given by the Bose-Einstein distribution and reflects the temperature dependence of the model $\langle n' \rangle = (\exp(\hbar\omega/k_B T) - 1)^{-1}$. The probability of transitions from $\mathbf{k}$ to $\mathbf{k}'$ due to surface roughness is directly proportional to the matrix element derived above. As a consequence of the ω^2 factor, low frequency phonons experience little thermal resistance from surface asperities. In addition, the dependence on the autocorrelation length L, which represents the average width of roughness peaks, shows that smoothly varying surfaces favor scattering events of the specular type.

Similarly to Si, we use a Gilat-Raubenheimer (GR) [111, 112] scheme to compute the surface roughness scattering rate integrals in Ge and GaAs (see Fig. 6.9.a)

$$\tau_{i,j}^{-1}(E) = \frac{2\pi}{\hbar N_i(E)} \int_{E'=E_i} \frac{|\langle \mathbf{k}|H'|\mathbf{k}'\rangle|^2}{\nabla_{\mathbf{k}'} E'(\mathbf{k}')} \, dS \tag{6.16}$$

where $\tau_{i,j}^{-1}$ is the transition rate from phonon branch i to j due to surface

roughness and $E'(\mathbf{k})$ is the phonon dispersion relation on the destination branch j. In order to account for the frequency dependence of phonon-surface roughness interference process, full dispersion relations are used, which are obtained from an adiabatic bond charge model, and approximate the dispersion relations observed in Si, Ge and GaAs [125, 126] with a 1-2.5% error (Fig. 6.9.b, c, d). The GR scheme approximates surface integrals on the first Brillouin zone (FBZ) discretized over $40 \times 40 \times 40$ cubes, which achieves sufficient accuracy in the 10-350 K range [113]. While the Gaussian approximation may be put at fault in the strong roughness limit where $\Delta/L > 1$, the effect of the L^2 term tends to average out the contribution of L. We noticed only little deviation of the predicted thermal conductivity in the strong roughness limit, and consistently used a value of $L = 22$ Å which provides a good fit in all cases.

The latter surface roughness scattering rate is added to the total scattering rate by Mathiessen's rule. Subsequently, we compute thermal conductivity in Holland's formalism [127]

$$\kappa = \frac{\hbar^2}{6\pi^2 k_B T^2} \sum_i \int \omega_i^4(q) \tau_i^{tot}(q) \frac{\exp(\hbar\omega_i/k_B T)}{(\exp(\hbar\omega_i/k_B T) - 1)^2} \, dq \qquad (6.17)$$

where i goes over all incident transverse and longitudinal branches, and the wave vector q models conduction in the nanowire along the Γ-(Δ)-X direction. For additional computational speed, an analytic expression is used for the phonon branches in the propagation direction where

$$\omega(q) = \omega_0 + v_\lambda q + c_\lambda q^2 \qquad (6.18)$$

The parabolic parameters are obtained in Ge and Si from a best fit on experimental observation [125, 126] as summarized in Table 6.1 and Fig. 6.10.

In addition to surface roughness scattering, computations of κ include boundary, isotope, 3 phonon Umklapp, and normal decay processes. Parametric laws are used for the latter rates, which are fitted to best reflect experimental bulk κ values [127, 121] as described in Table 6.3 and represented in Fig. 6.11 in the 10-400 K temperature range. In the case of Si, the parameters of Martin et al. [19] are used for their good fit with experimental observations on Si nanowire. Bulk boundary scattering is fixed to the value reported from experimental observations [127, 121]. The impurity scattering rate varies as $A\omega^4$ where A reflects isotope concentrations. A is not fitted, as we chose isotope concentrations of group IV semiconductors to match

Table 6.3: Summary of the scattering processes and associated parameters used in the derivation of Ge and GaAs thermal conductivity. Si properties are listed in Martin et al. [19].

|	GaAs[127]	Ge[121]
Phonon-Phonon:		
Umklapp	$B_\lambda T^3 (\hbar\omega/k_B T)^2 \exp(-\theta_\lambda/T)$	$B_\lambda T^3 (\hbar\omega/k_B T)^2 \exp(-\theta_\lambda/T)$
Normal: Transverse	$D_t T^5 (\hbar\omega/k_B T)$	$D_t T^5 (\hbar\omega/k_B T)$
Normal: Longitudinal	$D_l T^5 (\hbar\omega/k_B T)^2$	$D_l T^5 (\hbar\omega/k_B T)^2$
Phonon disp. <100>	$\omega(q) = \omega_0 + v_\lambda q + c_\lambda q^2$	$\omega(q) = \omega_0 + v_\lambda q + c_\lambda q^2$
B_t (K^{-3}/s)	8.232×10^3	6.119×10^3
θ_t (K)	40.791	58.06
B_l (K^{-3}/s)	3.268×10^{-3}	2.80×10^{-3}
θ_l (K)	22.764	221.337
D_t (K^{-4}/s)	4.17×10^{-4}	9.26×10^{-4}
D_l (K^{-3}/s)	0.4814	42.38
A (s^3)	0.466×10^{-44}	1.78×10^{-44}
v_{ta}, v_{la} (m/s)	3566.9, 5320.6	3616.7, 5342
v_{to}, v_{lo} (m/s)	-50.84, 155.82	-960.41, 294.97
c_{ta}, c_{la} (m^2/s)	-2.017×10^{-7}, -1.314×10^{-7}	-2.09×10^{-7}, -1.10×10^{-7}
c_{to}, c_{lo} (m^2/s)	2.29×10^{-8}, -9.84×10^{-8}	4.134×10^{-8}, -1.24×10^{-7}
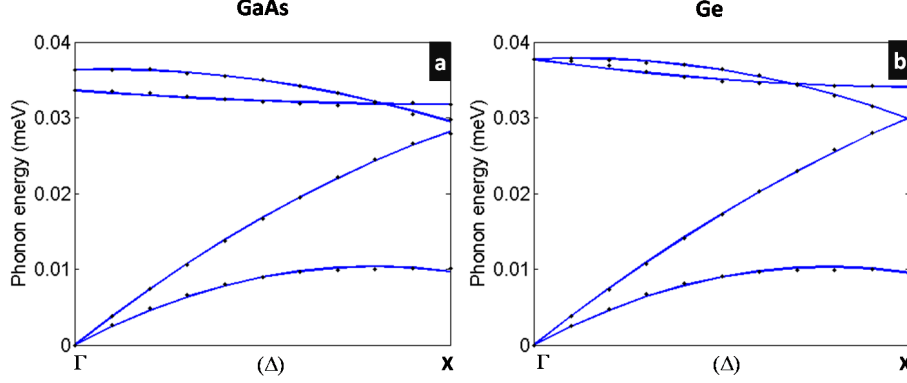
Figure 6.10: Approximation of the $< 100 >$ phonon branch in Ge and GaAs using the parameters of Table 6.3. The error is 1% in the case of Ge and 2.5% in the case of GeAs. Both analytic bands are obtained from a least square minimizing algorithm.

their natural occurrence, with 92.2 % of ^{28}Si, and noticeably only 36.5 % of ^{74}Ge [121]. As a result of its heavier atomic mass and high natural isotope concentration, Ge shows lower κ throughout all simulations (Fig. 6.12 and 6.13).

Subsequently, κ is computed for Ge and GaAs nanowires of various diameters and Δ (6.11). Although no reliable data on the roughness of small wires is available yet, recent studies have reported roughness RMS ranging from 3 Å for smooth vapor-liquid solid (VLS) grown nanowire, to 5 nm in the case of extremely rough etched nanowires [7]. It is noticeable that the effect of reduced dimensions is stronger in rough nanowires in comparison to the ideally smooth case. In Fig. 6.12 and Fig. 6.13, nanowire thermal conductivity is predicted to decrease in rough wires (high RMS Δ). The $(D/\Delta)^2$ behavior is apparent in thin nanowires of diameter $D < 115$ nm, showing that phonon–surface roughness scattering is the dominant resistive process in this case. For intermediate diameters ($D/\Delta \approx 100$ for Ge and GaAs, 50 for Si), the linear dependence in the diameter suggests that phonons dominantly decay from classical boundary scattering $\tau_B^{-1}(q) = v_s/D$, with v_s the average speed of sound in each branch. For wider D, κ converges to its bulk value.

Low nanowire thermal conductivity results from the combination of all concurrent interference processes mentioned above: isotope, impurity, Umklapp, and surface. While impurities tend to also reduce mobility, isotope scattering limits thermal conductivity and keeps electrical conductivity unchanged. At temperatures of 300 K and above, where the conduction is primarily limited by Umklapp scattering, the most sensitivity to isotope
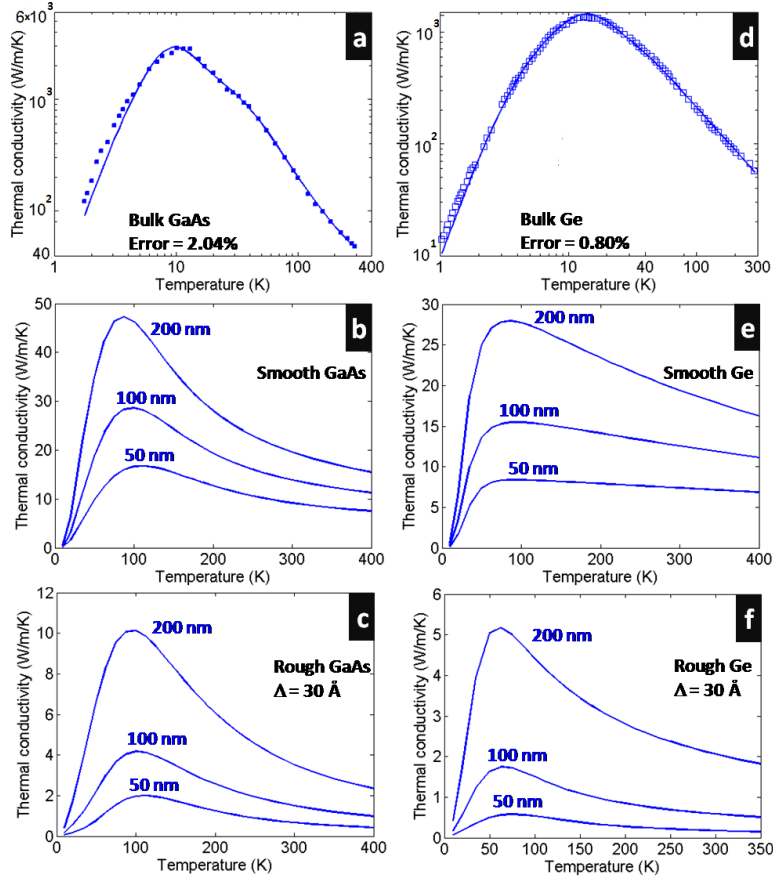
Figure 6.11: Parametric fit of the scattering rates on experimental bulk thermal conductivity (squares) of **(a)** GaAs [127] and (d) Ge [121]. Diameter dependence of nanowire thermal conductivity in (b) ideally smooth and (c) rough ($\Delta = 30$ Å, $L = 22$ Å) GaAs nanowire, and (e) ideally smooth and (f) rough ($\Delta = 30$ Å, $L = 22$ Å) Ge nanowire.
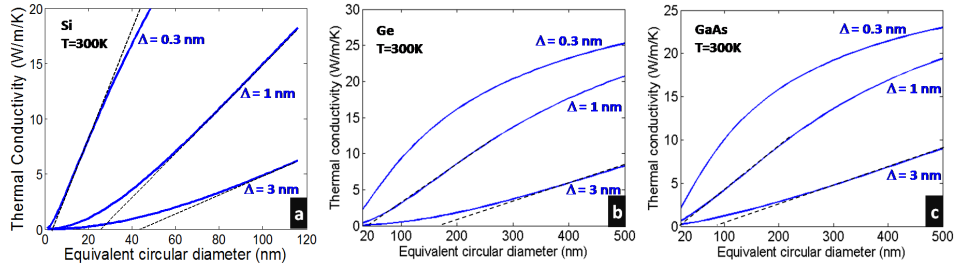
Figure 6.12: Effect of nanowire diameter on thermal conductivity for varying roughness RMS in (a) Si (after Martin et al. [19]), (b) Ge, and (c) GaAs. $L = 22$ Å throughout.

and surface scattering is obtained in nanowires fabricated from materials that have naturally low bulk Umklapp scattering rates. In particular, the effect of artificially high Δ and low D decreases thermal conductivity below

86

Figure 6.13: Effect of roughness RMS on thermal conductivity of Si, Ge and GaAs nanowire with diameters of 115 nm (upper curve) and 56 nm (lower curve). $L = 22$ Å throughout.

1 W/m/K in Ge and GaAs nanowires of diameter $D = 56$ nm. At this scale, fabrication of rough etched nanowires is feasible [7], and the perturbative approach holds validity as th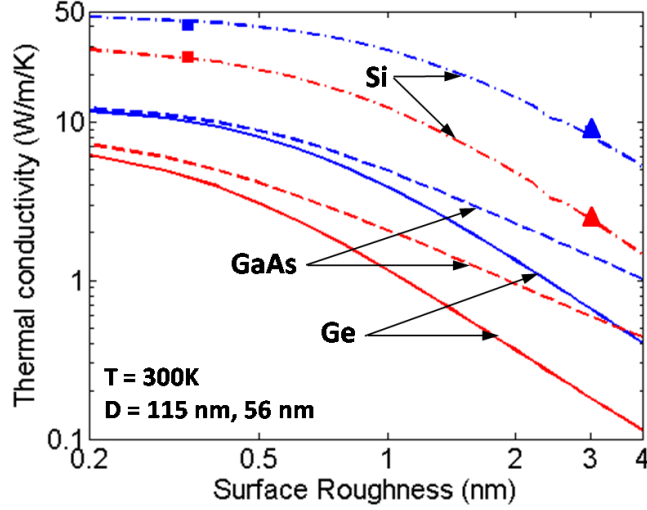e ratio Δ/D is only of a few percent, showing good agreement with experimental results (Fig. 6.13). Noticeably, the best fit for the natural Ge sample shows strong isotope and weak Umklapp scattering. As a result, Ge experiences the starkest influence from surface roughness scattering at high temperatures, with an extraordinarily low $\kappa \approx 0.1$ W/m/K at $D = 56$ nm and $\Delta = 4$ nm. GaAs, with respect to all phonon decay processes, experiences the strongest Umklapp scattering rate, and the effect of surface roughness remains moderate in comparison to Si and Ge (Fig. 6.13). In the case of silicon, the experimental observations of Hochbaum et al. [7] clearly show that enhancement of the thermoelectric figure of merit $ZT = S^2\sigma/\kappa T$ is due to a strong reduction by a factor 100 of the thermal conductivity, while the factor S^2 remains constant under the effect of surface roughness at 300 K in comparison to bulk values. We expect GaAs and Ge to behave quantitatively the same, and predict in this manuscript that surface roughness will reduce GaAs and Ge thermal conductivity by factors of 7 and 60 respectively (FIg. 6.13), with an expected increase of thermoelectric figure of merit in the same proportions. In addition, earlier modeling [107] of the variation of the power factor $S^2\sigma$ and thermal conductivity with the diameter of smooth nanowires, based on solutions to the Boltzmann transport equation, revealed that improvement of ZT in thermoelectric nanowires is due to a reduction of thermal conductivity rather than

the power factor. Indeed, improvement of the power factor is theoretically limited since the acoustic phonon-electron scattering drastically increases when the diameter of the nanowire is reduced [107, 108]. Nevertheless, in order to obtain accurate expectations for the figure of merit, the effect of surface roughness needs to be included into estimates of the power factor σS^2. The authors are currently unaware of such models and therefore recommend further investigation in this direction. Our model eventually predicts weak interference between low energy phonons and surface asperities. Consequently, we notice very little variation of thermal conductivity with respect to the roughness RMS at temperatures below 100 K.

6.5   Thermal Conductivity of Rough Graphene Nano-Ribbons

In previous sections, we demonstrated a remarkably strong effect of surface roughness on the thermal conductivity of thin semiconductor nanowires. Based on a full phonon dispersion relation we introduced a novel frequency-dependent model of boundary scattering for phonons. The resulting simulated thermal conductivity of nanowires of diameter $D < 200$ nm shows excellent agreement with recent experimental work. In particular, at low nanowire diameters, we predicted a strong deviation of the roughness-limited thermal conductivity from the linear diameter dependence ($\sim D$) to a scaling as $(D/\Delta)^2$, where Δ is the rms surface roughness.

The extension to nanowires of new materials solely implies a prior knowledge of the materials phonon dispersion relation. Consistently with our exploration of nano-engineered 2D ribbon structures, we are interested, in this section, in extending the simulation framework described above to rough graphene nano-ribbons. Recently, thermal transport in graphene captured tremendous attention from the scientific community for its extraordinarily high electrical conductivity [128, 129] and thermal conductivity [130, 131, 132, 30]. For this very reason, graphene ribbons have been kept from being considered for thermoelectric applications despite good electrical performance and possibility of gating. In this respect, we propose in this section a study of the effect of edge roughness on graphene thermal conductivity.

In addition, graphene presents the theoretical advantage of being based on a two-dimensional crystal lattice. In the models of 3D crystals, deriving scattering rates resulting from three phonon processes is in general com-

putationally intensive. In particular, calculations of anharmonic decays in zinc-blende lattices from the dispersion relation is a tedious task that eventually requires ab-initio calculations. This represents a consequent investment for little gain in comparison to the fitted analytic expressions we used above. Nevertheless, most of this complexity is alleviated in the case of a 2D lattice, and we will take advantage of this particularity here in order to derive a general framework to compute all phonon scattering rates from the phonon dispersion curves.

In this scope, computing graphene thermal conductivity requires key modifications to the nanowire model presented earlier. First, the roughness model must be adjusted to account for rough 1D edges instead of 2D surfaces. The statistical model for rough edges has been derived in chapter 2. Secondly, thermal conductivity computations must be similarly adjusted to account for the 2D nature of the phonon gas. Finally, we will provide details on the computation of phonon scattering rates from dispersion relation in the 2D graphene hexagonal crystal lattice.

In the direction of propagation, assuming a discrete distribution of phonon momentum states q along branches s, thermal conductivity is, by definition

$$\kappa(T) = \frac{1}{V} \sum_{s,q} \hbar\omega(q)v^2(q)\tau_{tot}(q)\frac{\partial n}{\partial T} \qquad (6.19)$$

First, it must be noted that thermal conductivity, for unit consistency, calls for a volume V of the propagative sample in m^3. In the case of graphene, the volume is artificially reproduced by assuming a finite effective height of the mono-layered graphene sheet. Physically, this must be understood by the intrinsically three-dimensional nature of phonon vibrations, which effectively makes the phonon gas in graphene disperse in 3 dimensions. This effective height has been approximated by Nika et al. [30] to $H = 0.65$ nm which gave satisfactory approximation of experimental observations on graphene flake thermal conductivity. This value will be used in the remainder of this section. Secondly, in a sufficiently large sample (width $W > 200$ nm), the distribution of phonon momentum can be considered as continuous, with distribution $\frac{1}{V}g_{1D}(\mathbf{q}) = q/(4\pi H)$. Using this assumption, and with $n = 1/(\exp(\hbar\omega/kT) - 1)$, the thermal conductivity in a propagation direction χ can be computed from the integral

$$\kappa_\chi(T) = \frac{1}{4\pi HkT^2} \sum_{s=1..6} \int_0^{q_{\chi,max}} \hbar\omega_s qv_s^2(q)\tau_{s,tot}(q)\frac{\exp(x)}{(\exp(x)-1)^2}\, dq \quad (6.20)$$

where s goes over all branches in the propagation direction, $x = \hbar\omega/kT$, and $\tau_{s,tot}(q)$ is the total scattering rate for incident phonons on branch s of the propagation direction χ.

To support the following calculations, a full phonon dispersion relation can be computed in the entire graphene Brillouin zone, using a valence force field (VFF) method, as described in Saito et al. [29]. For this purpose, force constants are approximated until the 4th nearest neighbors, with constants found in [29]. The full phonon dispersion relation and representation of the graphene Brillouin zone are represented in Figs. 6.14 and 6.15.

In the following derivation and simulations, we assume that the graphene ribbons are suspended in vacuum, and no heat is exchanged between the ribbon and its environment. In this case, the total phonon scattering rate is a combination of impurities, boundaries, 3 phonon decays, and surface roughness such that, using Mathiessen's rule

$$\tau_{tot}^{-1}(q) = \tau_I^{-1}(q) + \tau_B^{-1}(q) + \tau_U^{-1}(q) + \tau_R^{-1}(q) \qquad (6.21)$$

The semi-classical impurity and boundary scattering rates are given by $\tau_B = v(q)/W$ and $\tau_I = S_0 \Gamma q \omega^2(q)/4v(q)$ from Srivastava [133], with Γ Gruneissen's constant and S_0 the area of a cell of the FBZ. Those rates are fairly easy to include in numerical computations, and we will focus on the 3 phonon and impurity scattering rates, for which the analytic expressions are not given.

In the case of three phonon Umklapp processes, we shall distinguish between two classes of behavior. The first case corresponds to the decay of a phonon into two phonons (emission, type I), while the second case consists in the recombination of two phonons into one phonon (absorption, class II). Umklapp processes are by definition elastic; consequently, decay rules must satisfy conservation of momentum and energy. Incidentally, using $\{\mathbf{b_i}\}_i$ the set of destination nearest neighbor reciprocal cells (Fig. 6.15.c), the decay of a phonon $\mathbf{q}_0$ into $\mathbf{q}_1, \mathbf{q}_2$ happens such that

$$\begin{aligned} \mathbf{q}_0 + \mathbf{b}_i &= \mathbf{q}_1 + \mathbf{q}_2 \\ \omega_0(\mathbf{q}) &= \omega_1(\mathbf{q}) + \omega_2(\mathbf{q}) \end{aligned} \qquad (6.22)$$

Similarly, for type II Umklapp processes, phonon absorption happens in
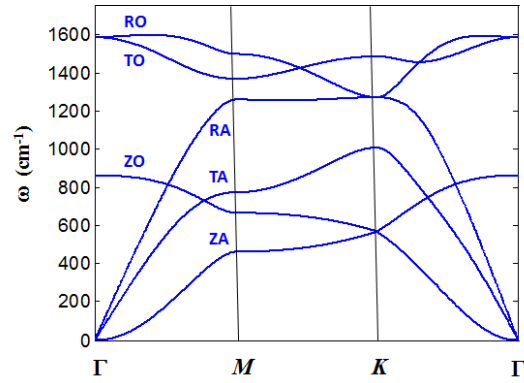
Figure 6.14: Computed dispersion relation in graphene based on a DFT algorithm where force constants are approximated until the 4th nearest neighbors. Six branches are observed: out of plane acoustic, in plane tangential acoustic, in plane radial acoustic, out of plane optical, in plane radial optical and in plane tangential optical.
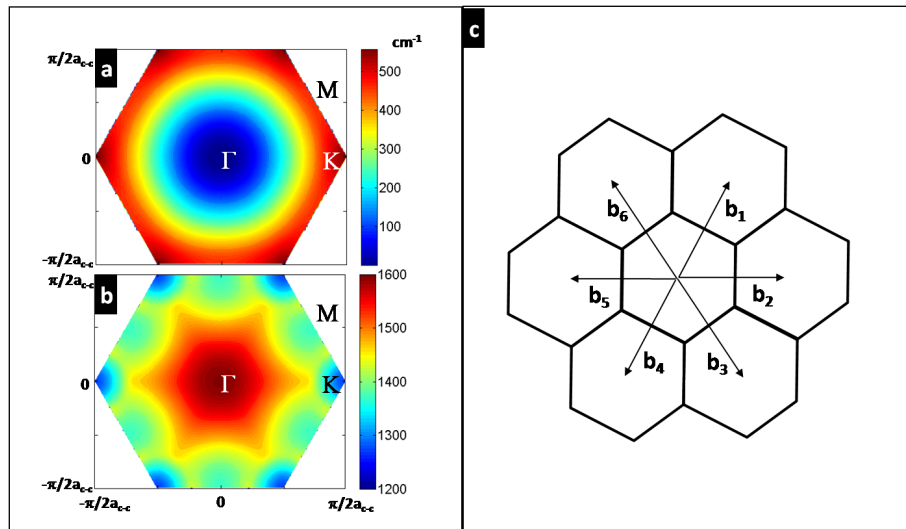
Figure 6.15: Example of a 2D dispersion relation computed in the graphene reciprocal lattice from a valence force field (VFF) method. (a) Out of plan acoustic. (b) In plane bond-bending optical mode. (c) Representation of the graphene reciprocal lattice with the basis of reciprocal vectors to the first nearest reciprocal cells.

accordance with conservation of energy and momentum

$$
\begin{aligned}
\mathbf{q}_0 + \mathbf{q}_1 &= \mathbf{b}_i + \mathbf{q}_2 \\
\omega_0(\mathbf{q}) + \omega_1(\mathbf{q}) &= \omega_2(\mathbf{q})
\end{aligned} \tag{6.23}
$$

Equations 6.22 and 6.23 above, with conditions on allowed combinations

of phonon momentum vectors, define sections of the plane. For instance, given an initial phonon momentum $\mathbf{q_0}$ and branch λ_0, the set of decay momentum $\mathbf{q_1}$ from branch λ_1 that allows decay with a branch λ_2 is a line in the 2D first Brillouin zone. Subsequently, it is possible to approximate the delta function in order to materialize energy conservation in the energy space. Computationally, handling the delta function requires particular care for its diverging nature. For instance, we may use one of the analytical representations of the delta function

$$\delta(\omega) = \lim_{\delta=0} \left(\frac{\delta}{\omega^2 + \delta^2} \right) \tag{6.24}$$

In computations, keeping the value of δ small but non-zero, gives an approximation to the Dirac function which respects the properties of symmetry and unicity of the functions, and avoids introducing biases in computation of integrals. As a result, approximating $\delta(\omega_0(\mathbf{q_0}) - \omega_1(\mathbf{q_1}) - \omega_2(\mathbf{q_2}))$ can be achieved by the following steps:

- For a given $\mathbf{q_0}$ in the propagation direction, compute all allowed combinations of $\mathbf{q_1}$, $\mathbf{q_2}$ satisfying momentum conservation.

- For the allowed transitions, compute the phonon frequencies from the dispersion relation tables pre-computed using the VFF method.

- Using the phonon frequencies above, compute delta functions for allowed transitions between bands

Figure 6.16 shows an example of how allowed 3-phonon transitions between branches can be represented using an approximation to the Dirac delta function. Given the law of energy and momentum conservation, the 3-phonon scattering rate is computed from the following sumation [30] for absorption

$$\tau_U^{-1}(\lambda_0, \mathbf{q_0}) = \frac{\hbar \Gamma^2}{3\pi \rho v_{\lambda_0}^2(\mathbf{q_0})} \sum_{\lambda_{1,2}} \int \int \omega_0(\mathbf{q_0}) \omega_1(\mathbf{q_1}) \omega_2(\mathbf{q_2})$$
$$\times \left[N(\omega_1) - N(\omega_2) \right] \delta(\omega_0(\mathbf{q_0}) + \omega_1(\mathbf{q_1}) - \omega_2(\mathbf{q_2})) d\mathbf{q_1} \tag{6.25}$$

and similarly for emission:

$$\tau_U^{-1}(\lambda_0, \mathbf{q_0}) = \frac{\hbar \Gamma^2}{3\pi \rho v_{\lambda_0}^2(\mathbf{q_0})} \sum_{\lambda_{1,2}} \int \int \omega_0(\mathbf{q_0}) \omega_1(\mathbf{q_1}) \omega_2(\mathbf{q_2})$$
$$\times \left[N(\omega_1) + N(\omega_2) + 1 \right] \delta(\omega_0(\mathbf{q_0}) - \omega_1(\mathbf{q_1}) - \omega_2(\mathbf{q_2})) d\mathbf{q_1} \tag{6.26}$$
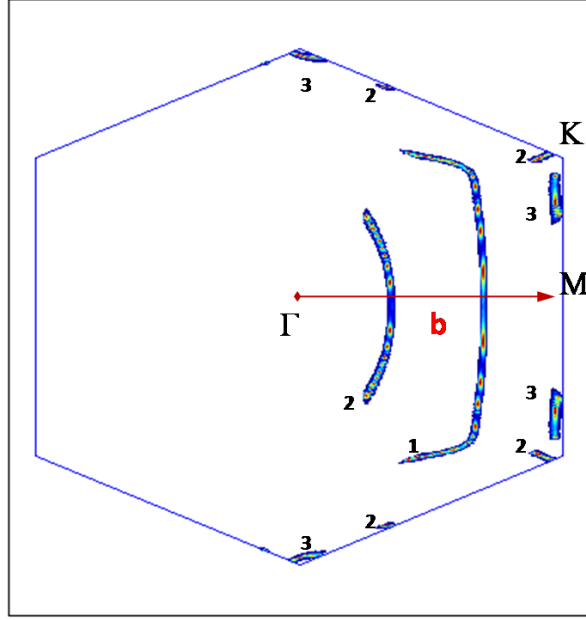
Figure 6.16: Lines of allowed transitions satisfying energy and momentum conservation for phonon recombination. 1. $ZA + ZA -> ZA$ with $\mathbf{q_0} = 0.4 \times (\Gamma M)$. 2. $ZA + ZA -> ZA$ with $\mathbf{q_0} = 0.75 \times (\Gamma M)$. 3. $ZA + TA -> TA$ with $\mathbf{q_0} = 0.4 \times (\Gamma M)$.

Hence, it is possible to compute 3 phonon decay rates in graphene at a given temperature, as represented in Fig. 6.17 for $T = 300$ K. Umklapp scattering rates are computed for initial phonons $\mathbf{q_0}$ propagating in the $\boldsymbol{\Gamma}$-direction. For both acoustic and optical branches, absorption is dominant at low phonon momentum. Incidentally, in ZA, TA, RA, TO, and RO branches, phonon emission only happens when a critical value of $\mathbf{q_0}(\omega)$ is reached, which allows decay of a phonon in two states. Given the results above at room temperature, and neglecting for now other types of decay, we can deduce that heat propagation in graphene mostly occurs on acoustic branches. On these branches, dispersion is limited at low q by absorption, with a decay time in the order of 10 ps. Similarly, at high q, heat conduction in graphene is mostly driven by phonon emission, with a decay time in the order of 0.1 ps. In Fig. 6.17, the dependency of the Umklapp rate in $v_\lambda(\omega)^2$ clearly appears.

Figure 6.18 represents the total scattering rate of type I and type II Umklapp processes combined, at $T = 300$ K and $T = 200$ K. Using initial propagation along the $\boldsymbol{\Gamma}$-$\mathbf{K}$ direction, we notice the effect of temperature on the rates. Importantly, the Umklapp rate for low energy phonons is reduced by a factor ≈ 15 from room temperature to 200 K. On the other hand, optical
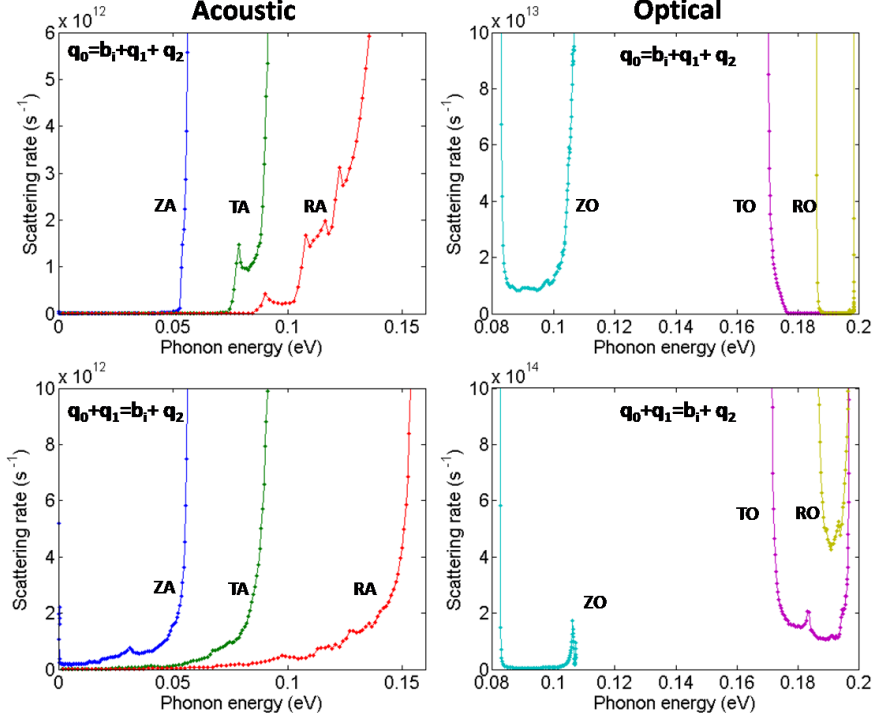
Figure 6.17: Computed Umklpapp scattering rates at $T = 300$ K. Top figures represent phonon emission, bottom figures represent phonon absorption.

Umklapp rates remain unchanged, except for the out of plane optical peak which is decreased by ≈ 50 % at $T = 200$ K. Additionally, temperature strongly predetermines the occupation of the branches of initial propagating phonons $\mathbf{q_0}$. This is not reflected on the rates above, but through the computation of the thermal conductivity integral in equation 6.20.

Computations of elastic 2 phonons decay processes spawn from a similar model. In a model of heat conduction in graphene ribbons, we consider the following phonon scattering mechanisms: 3 phonons Umklapp, point defect, classical boundary, and surface roughness. Point defect scattering is calculated from the following rate [30]:

$$\tau_I^{-1}(q) = \frac{S_0 \Gamma}{4} \frac{q}{v(q)} \omega^2(q) \tag{6.27}$$

where S_0 is the area occupied by one atom of carbon in the lattice, and $\Gamma = 1$ is a screening factor. The value of 1 used by Nika et al. [30] provides the best agreement with experimental results. Similarly to scattering in nanowires, the classical boundary scattering rate in graphene ribbons is derived from
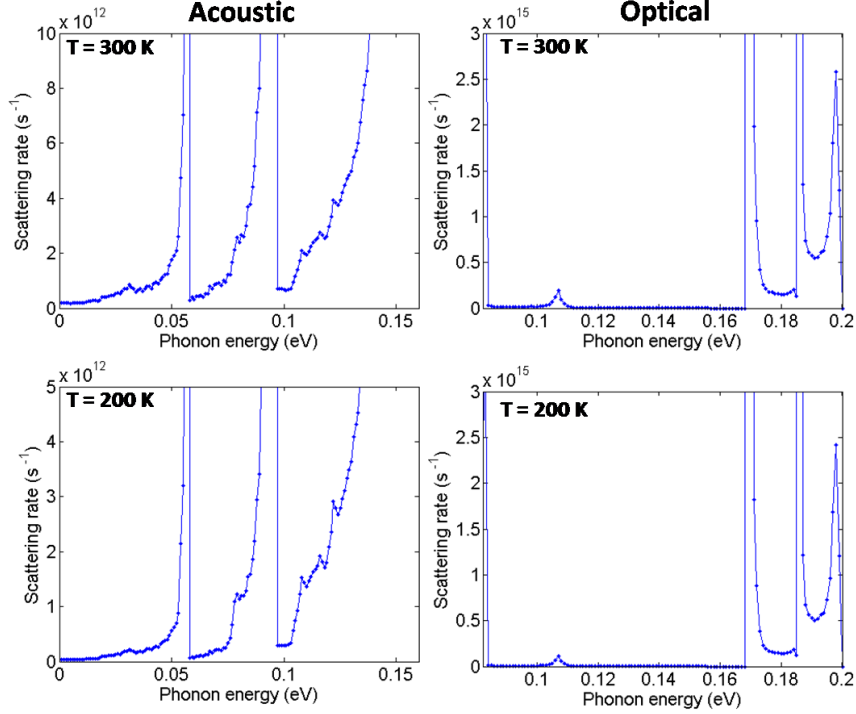
Figure 6.18: Total Umklapp scattering rate for optical and acoustic branches in bulk graphene. Top figures: $T = 300$ K. Bottom figures: $T = 200$ K.

the analytic expression

$$\tau_B^{-1} = \frac{v(q)}{W} \tag{6.28}$$

where W is the width of the ribbon. Figure 6.19 depicts point defect and classical boundary scattering rates in graphene as computed from the rates above. Given equations 6.27 and 6.28, it is remarkable in Fig. 6.19 that the rates diverge or converge at the roots of null velocity of sound. This consequently limits branch occupation at high phonon momentum.

So far, we derived an expression of thermal conductivity in a graphene ribbon. Based on full branch computation, we calculated 3 phonon decay rates, as well as classical boundary and point defect scattering rates. In order to account for edge roughness, we develop below a similar framework to compute a surface roughness phonon transition matrix element for 1D edges. Here also, we adopt the model of a quasi particle phonon that scatters from a series of constrictions along the propagation direction (Fig. 6.20.a).

From equation 2.13, the roughness spectral density of a rough 1D edge is
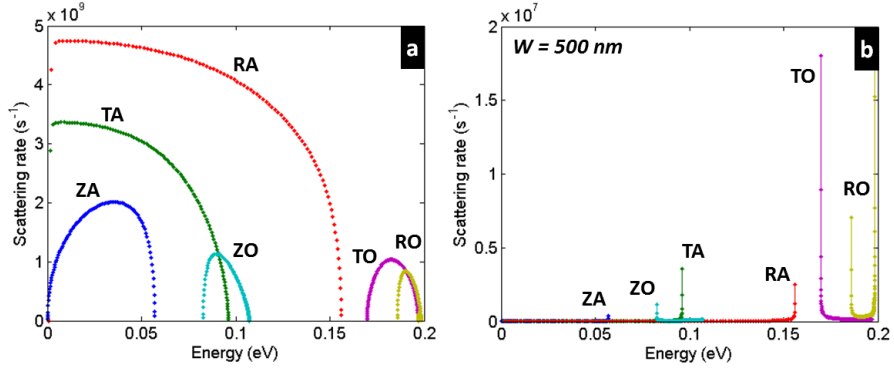
Figure 6.19: (a) Phonon - point defect scattering rate for all incident branches in the **Γ-K** direction. The screening factor is $\Gamma = 1$. (b) Classical boundary scattering in graphene ribbons, for ribbon width $W = 500$ nm.
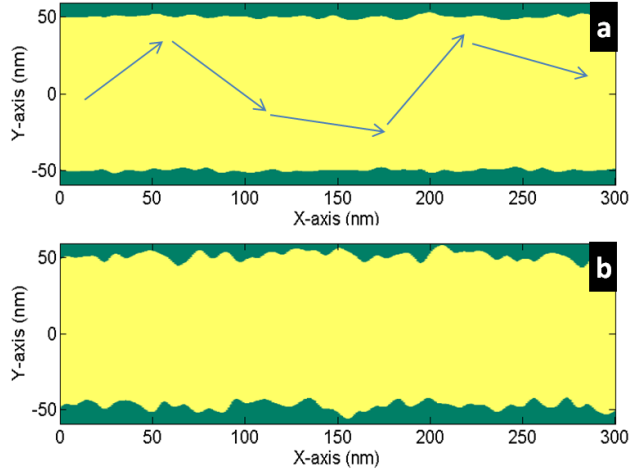
Figure 6.20: (a) Model of semi-classical phonon scattering in a rough graphene ribbon, $\Delta = 1$ nm. (b) Model of a rough graphene ribbon, $\Delta = 3$ nm. $L = 6$ nm and $W = 100$ nm throughout.

(notice the dimension of m³)

$$\Delta(q) = \sqrt{\pi}\Delta^2 L e^{q^2 L^2/4} \tag{6.29}$$

where Δ is the roughness RMS and L the auto-covariance length. Using similar definition of the phonon-surface roughness matrix element, we derive the following expression:

$$|\langle \mathbf{k}|H'|\mathbf{k}\rangle|^2 = \frac{4\gamma^2}{3A}\omega^2(\langle n'\rangle + 1)\Delta^2(\mathbf{k} - \mathbf{k}') \tag{6.30}$$

where A is the area of the ribbon. The probability of transitions from **k** to

96

$\mathbf{k}'$ due to surface roughness is directly proportional to the matrix element derived above. As a consequence of the ω^2 factor, low frequency phonons experience little thermal resistance from surface asperities. In addition, the dependence on the autocorrelation length L, which represents the average width of roughness peaks, shows that smoothly varying surfaces favor scattering events of the specular type.

Similarly to semiconductor nanowires, we use a Gilat-Raubenheimer (GR) [111, 112] scheme to compute the surface roughness scattering rate integrals in graphene (see Fig. 6.21)

$$\tau_{i,j}^{-1}(E) = \frac{2\pi}{\hbar N_i(E)} \int_{E'=E_i} \frac{|\langle \mathbf{k}|H'|\mathbf{k}'\rangle|^2}{\nabla_{\mathbf{k}'} E'(\mathbf{k}')} \, dS \qquad (6.31)$$

where $\tau_{i,j}^{-1}$ is the transition rate from phonon branch i to j due to surface roughness and $E'(\mathbf{k})$ is the phonon dispersion relation on the destination branch j. In order to account for the frequency dependence of phonon-surface roughness interference process, full dispersion relations are used, which are obtained from Dresselhauss' valence force field model. The GR scheme approximates surface integrals on the first Brillouin zone (FBZ) discretized over 200×200 cells, which achieves sufficient accuracy in the 20 K -350 K temperature range [113]. While the Gaussian approximation may be put at fault in the strong roughness limit where $\Delta/L > 1$, the effect of the L term tends to average out the contribution of L. We noticed only little deviation of the predicted thermal conductivity in the strong roughness limit, and consistently used a value of $L = 22$ Å which provides an experimental fit with Balandin's observations [30].

Subsequently, κ is computed for GNRs of various width and roughness RMS Δ. We assume heat propagation in a graphene sheet along the $\mathbf{\Gamma} - \mathbf{K}$ direction, and a temperature range of 20 K - 350 K (Figs. 6.22 and 6.23). Although no reliable data on the roughness of small wires is available yet, our study from chapter 2 shows feasibility of rough edges with a model roughness RMS ranging from 3 Å to 5 nm in the case of extremely rough edges. Experimentally, such rough edges have not been manufactured yet with sufficient reproducibility, though, similarly to electroless etching of nanowires, one could imagine a technique of edge etching in severely randomizing experimental conditions (extremely low/high pressure, etc.). It is noticeable that the effect of reduced dimensions is stronger in rough ribbons in comparison to the ideally smooth case. In 6.22 and 6.23, GNR thermal conductivity is predicted to decrease in rough ribbons (high RMS Δ). The $(W/\Delta^2)^2$ behavior is apparent in the variations of κ with Δ, showing that phonon–surface
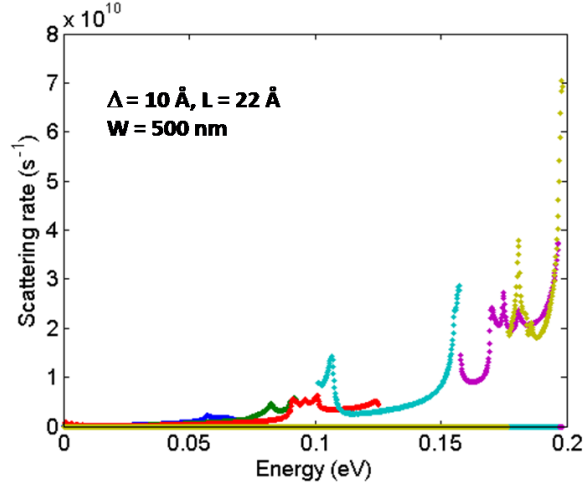
Figure 6.21: Graphene phonon - surface roughness scattering rate computed for each incident branch from a Gilat Raubenheimer scheme. Ribbon width $W = 500$ nm, roughness rms $\Delta = 1$ nm, $L = 2.2$ nm.
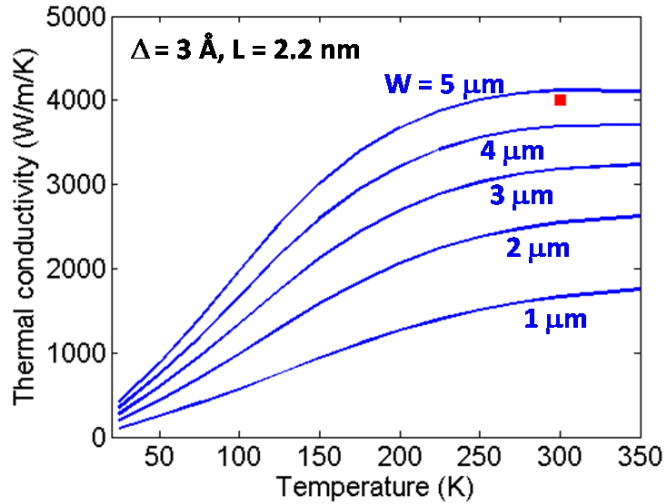
Figure 6.22: Graphene thermal conductivity computed from Holland's formalism in ribbons of varying width. Phonon-surface roughness is included with RMS $\Delta = 3$ Å, and $L = 22$ Å. Square is included from experimental observation of Nika et al. [30]

roughness scattering is the dominant resistive process for $\Delta > 2$ nm. We noticed that for greater width W, κ converges to its bulk values, with a peak in thermal conductivity around 90 K.

Low GNR thermal conductivity results from the combination of all concurrent interference processes mentioned above: isotope, impurity, Umklapp, classical surface, and phonon-surface roughness. While impurities tend to also reduce mobility, isotope scattering limits thermal conductivity and
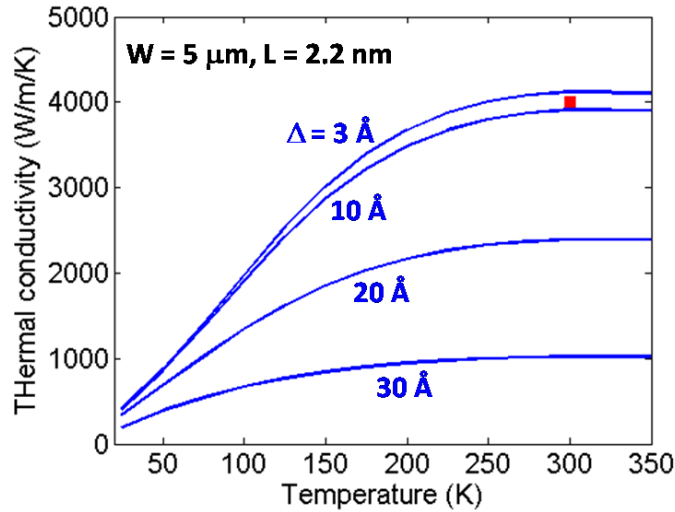
Figure 6.23: Graphene thermal conductivity computed from Holland's formalism in ribbons of varying roughness RMS. Phonon-surface roughness is included with ribbon width $W = 5$ μm, and $L = 22$ Å. Square is included from experimental observation of Nika et al. [30]

keeps electrical conductivity unchanged. In particular, the effect of artificially high Δ decreases thermal conductivity below 1000 W/m/K in suspended GNRs of width 5μ m, and $\Delta = 3$ nm. Importantly, in the case of smooth GNRs, agreement is found with the experimental data point of Nika et al. [30]. The ability to reduce thermal conductivity in graphene while keeping the excellent electronic conduction of the material shows again promise for thermoelectric applications. Nevertheless, in order to obtain accurate expectations for the figure of merit, the effect of surface roughness needs to be included into estimates of the power factor σS^2. The authors are currently unaware of such models and therefore recommend further investigation in this direction.

6.6 Perspectives on Roughness Limited Thermal Conductivity

In summary, we first modeled thermal properties of artificially rough Si, Ge and GaAs nanowires based on a perturbative treatment of the interaction between lattice vibrations and surface asperities. This approach, based on a full 3D phonon dispersion in each material, accurately accounts for the frequency dependence of phonon scattering processes resulting from surface roughness, isotope, boundary, and anharmonic decays. Our model predicts

$\kappa \propto (D/\Delta)^2$ for thin nanowires, yielding remarkably low thermal conductivity below 1 W/m/K in rough etched Ge and GaAs nanowires of diameter $D < 115$ nm. This model was subsequently extended to new mono-layered material. In this case, physical laws have to be reworked to essentially account for the 2D nature of these materials. In this respect, graphene presents considerable theoretical advantages to work in details of the computation of all heat resistive processes in semiconductor devices. Thus, using a computed full phonon dispersion relation, it is possible to derive the branch to branch decay processes for 3 phonons Umklapp scattering, and classical boundary scattering. In view of this, we extend the results on nanowires to graphene nano-ribbons. In this case, low thermal conductivity below 1000 W/m/K is also reported for ribbon width $W < 5$ μm and $\Delta = 3$ nm. Additionally, decreasing width and increasing roughness are expected to further decrease thermal conductivity. As a result, for both semiconductor nanowires and nano-ribbons, the ability to reduce thermal conductivity while keeping good electronic conduction would reveal those materials as potential candidates for efficient site-specific, on-demand, integrated thermoelectric energy conversion solutions.

In particular, for the same diameter and roughness rms and at room temperature, GaAs and Ge nanowires are expected to have approximately 5 times and 10 times lower thermal conductivity respectively than Si nanowires. Similar scattering mechanisms nevertheless limit electron mobility, making it necessary to model charge transport in order to find optimal thermoelectric figures of merit ZT. In this scope, it is noticeable that isotopes have no influence on electronic conduction, and surface roughness may only have limited impact at low D where electrons tend to be pushed away from the interface. Hence, those two factors are expected to be significant design parameters in the engineering of efficient semiconductor nanowire thermoelectric devices.

In addition, we demonstrated in this chapter a remarkably strong effect of surface roughness on the thermal conductivity of thin semiconductor nanowires and nano-ribbons. The approach presented here may generally be extended to semiconductor nanowires and nano-ribbons of various materials, sizes, and direction of thermal propagation using a full-phonon dispersion relation with the inclusion a phonon – surface roughness scattering matrix element. Beyond the class of unidirectional semiconductor based thermoelectric energy converters, previous experimental devices [99, 100, 97, 98, 101, 102] based on classical high efficiency thermoelectric materials such as Bi_2Te_3 and Sb_2Te_3 faced potential challenges in their inte-

gration onto conventional microscale semiconductor platforms. Though not competing with the high thermoelectric performance $ZT > 2$ of these devices, showing that nano-engineered semiconductor nanowires achieve thermoelectric figures of merit close to $ZT = 1$ is nevertheless of considerable importance, given their natural and convenient compatibility with highly available electronic system on a chip technologies. In that respect, in semiconductor nanowires, charge carriers are naturally pushed away from the surface. As a result, surface roughness is expected to minimize thermal conductivity in greater proportion than electrical conductivity and Seebeck coefficient. With that perspective, the present work theoretically addressed the trade-offs between different scattering mechanisms in unidirectional semiconductor devices necessary to achieve low thermal conductivities through surface nanostructuring. Nevertheless, a refined understanding of the impact of surface roughness on the power factor is required in order to further explore possible designs of efficient embedded on-site thermoelectric nanowire coolers or generators. It is generally assumed that semiconductors with small bandgap and high effective mass present higher thermopower. As a result, applying a similar modeling framework to InSb could reveal interesting nanowire properties for thermoelectric applications.

Finally, we chose in this section to use an integral solution to heat conduction in semiconductors. This semi-classical assumption is in general a sufficiently accurate solution to the Boltzman transport equation when the dimensions of the device are several orders greater than the quasi particle mean free path. For greater accuracy, inclusion of the derived scattering rates could be included in a phonon Monte Carlo simulation. Proceeding with the latter model presents the clear advantage of allowing coupled electronic and phonon transport simulation, through a common set of Monte Carlo simulation parameters [134]. Nevertheless, assumptions beyond the Boltzmann transport equation are limiting in the scope of capturing the quantum nature of charge and heat carrier. The wave nature of energy carriers is expected to be of tremendous importance in interference effects happening at device scales below the carrier mean free path. Consequently, in the following chapter, we propose an alternative approach to modeling surface limited transport in semiconductors which revolves around a Green function treatment of the energy carriers. Specifically, we will investigate the case of electronic conduction in 3D rough semiconductor nanowires and 2D patterned sheets.

CHAPTER 7

QUANTUM ENHANCED SURFACE SCATTERING IN CONDUCTING CHANNELS BELOW 10 NM

7.1 On the Approach to Electronic Conduction in Channels Below 10 nm

We discussed in previous chapters the effect of surface roughness on the heat and electrical conduction in semiconducting channels. In both chapters 5 and 6, we considered a semi-classical model in which quasi-particles moved according to classical physics, in free flights interrupted by scattering events. Those scattering events were derived from perturbation theory, as a perturbation to the Hamiltonian of the system. In the case of electronic transport, Schrödinger's equation was solved self-consistently to take into account charge relocation due to quantum confinement. However, the effects of quantum confinement on the electronic band structure and the phonon dispersion relation have not been considered in the previous sections. The latter is a good approximation as long as the devices considered were intrinsically three-dimensional, and the contribution of sub-bands to the conduction processes remained negligible [135].

In this chapter, our objective is to get a finer understanding of the effect of surface roughness on the process of conduction in channels where the transport properties mostly result from the quantum nature of carriers. Using statistical series of real 2D and 3D geometries of electronic conduction channels, we derive a detailed description of the resistance incurred by wave-like charge carriers from the constrictions along the propagation direction in channels below 10 nm. In particular, we introduce a 3D non-parabolic recursive Green function (RGF) approach to study conduction in ultra-thin conduction channels, quantum modulated transistors, and patterned semiconductor nano-ribbons. The latter approach is consistent with previous results obtained from a similar approach [136], and shows, in the absence of other scattering mechanisms, that surface roughness contributes to an increased resistivity for high energy carriers in Si NW of diameter $D < 10$ nm, while having little effect on the lower energy levels. Based on the same

approach, we exhibit increased electronic resistivity and negative resistance in asymmetrically patterned semiconductors ribbons.

The implications of the latter study based on an RGF approach are two-fold. First, as we are researching an optimal design for rough nanowires and nano-ribbons used as thermoelectric devices, it is possible that best performance occurs from quantum processes. Indeed, while considering the wave-like nature of electronic and heat carriers and their inherently different wavelengths, it is conceivable that the size of geometrical fluctuations of the wire will strongly interfere with only heat waves, and little with electronic waves. With this approach, one could design an electronic crystal - phonon glass, similarly to our design goals in chapter 6 using semi-classical concepts. It must be kept in mind that even at scales around and below 10 nm, it is still appealing to design interfering geometrical patterns through a randomizing etching process or chemical growth. In this case, roughness results from random conditions in the fabrication environment [7] and, as such, our statistical approach to surface roughness keeps its validity. Second, electronic devices based on the quantum nature of carriers have already been proposed [137], and in order to study the feasibility of such devices, it is necessary to understand what the impact of real geometries will be on their performance. Here, we find particular interest in the quantum modulated transistor [136]. In this configuration, full use is made of the wave-like nature of electrons, as a gate voltage is applied to control the interference of carriers with a geometrical pattern along the transport direction. Thus, we would like to understand the implication of the real shape of these geometrical patterns on the operation of the quantum modulated transistor. Finally, we would like to bring to the reader's attention that, in a recent work from H. Hahm [135], it was concluded from coupled semi-classical and full-quantum approach that the critical dimension of nano-channels at which strong quantum confinement effects become relevant is on the order of 10 nm.

The lateral size of the conduction channel at which the wave-like nature of carriers must be taken into account is critical in understanding when the following approach must be used. In this respect, a simple back-of-the-envelope calculation instinctively confirms this assumption and introduces the problematic behind the model which we will introduce in this chapter. For now, let us consider a 2D model in which the confinement width varies along the conduction direction (Fig. 7.1). For a nanowire width D and an average deviation Δ of the confinement width, the first transverse eigen-energies of the Schrödinger equation vary along the conduction direction
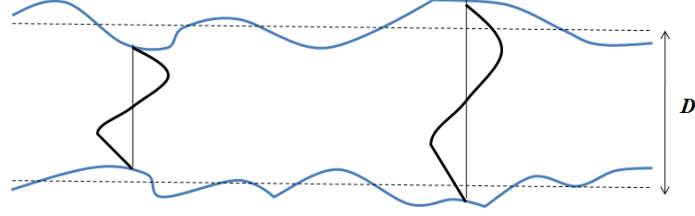
Figure 7.1: Model of the variation of the quantum confinement width along the propagation direction on a nano-channel.

between, on average:

$$E_1^{min} = \frac{\hbar^2}{2m^*(D+\Delta)^2} \tag{7.1}$$

$$E_1^{max} = \frac{\hbar^2}{2m^*(D-\Delta)^2} \tag{7.2}$$

For any material with $\Delta = 0.5$ nm, the variation in the first energy level $(E_1^{max} - E_1^{min})/E_1^{av}$ is of 20% at $D = 10$ nm, 10% at 20 nm, and 4% at 50 nm.

With this in mind, modeling transport in structures with confinement lengths on the order of 10 nm and below requires additional attention to the quantum interferences which appear from the variation of the width of the waveguide. In particular, it can be expected that the roughness scattering rates will be strongly modified by the varying confinement width and the appearance of allowed discrete eigen-modes which can be transmitted in the perturbed waveguide. Besides, in the absence of impurities, transport in such quasi-1D channels as nanowires is limited by phonon scattering at high temperatures whereas, at low temperatures, heat and electronic conduction are mostly limited by scattering of quasi-particles with the device boundaries. Similarly to the scattering rates derived in chapters 5 and 6, we can assume that the full-quantum surface roughness scattering rate is, with a good approximation, temperature invariant. Therefore, we propose in this chapter to derive this new scattering rate in the low temperature regime, with the idea to extend this rate to the high temperature regime with only minimal change.

The low temperature regime presents the advantage of behaving closely to the predictions resulting from Landauer's formalism [138, 139]. In the case of a device performing 1D transport between two reservoirs of Fermi energies $E_{F,1}$ and $E_{F,2}$ and an energy dependent transmission probability $T(E)$ accounting for various scattering mechanisms, Landauer's prediction
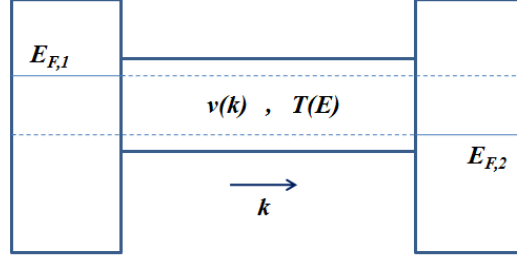
Figure 7.2: Model of 1D conduction between two reservoirs with different Fermi levels. The transmission probability $T(E)$ reflects the effect of scattering in the channel.

gives, for instance, the electrical current from reservoir 1 to reservoir 2 7.2:

$$I = \frac{e}{\pi} \left[\int_0^\infty v(k) f_1(k) T(E) \, dk - \int_0^\infty v(k) f_2(k) T(E) \, dk \right] \qquad (7.3)$$

where f_1 and f_2 are the reservoir distribution functions characterized by their respective Fermi energies. In 1D and at low temperatures, the energy dependent velocity and the density of states cancel to a simple constant, yielding the simple expression

$$I = \frac{2e}{h} \int_{E_{F,1}}^{E_{F,2}} T(E) \, dE \qquad (7.4)$$

For an energy invariant transmission probability T, the channel presents a conductance which is directly proportional to the fundamental conductance $\mathcal{G} = 2e^2/h$. It is also noticeable that the Fermi energies depend directly on the voltage applied on the reservoirs. Nevertheless, in the case of real contacts, a non-negligible part of the resistance measured effectively occurs at the contacts. Thus, in order to properly derive the transmission coefficient of the channel—which will truthfully, and only, account for our boundary scattering model—one must keep in consideration a potential drop at the contacts. In practice, this potential drop can be explained by a charge reordering at the interfaces between the channel and the reservoirs. Such effects have already found an analytical derivation in the literature [12]. At this point, we will simply keep in mind that, from the knowledge of the current flowing in the channel and the effective potential drop at its left and right boundaries, it is possible to estimate the conductance of the channel, which, in turn, can be directly translated into a scattering rate according to Drude's theory.

In the following sections, we focus on the problem of the scattering of an

electron quantum wave-packet with the boundaries of a structure similar to the one shown in Fig. 7.1. In this respect, many option are available to deal with the transmission of different energy modes on a quantum structure. The mode matching formalism [12] is a viable solution; however, it seems less adapted to a system with rough boundaries and has been often considered unstable [136]. A very popular alternative is the non-equilibrium Green function [140, 141] (NEGF), in which the retarded green function is derived from inverting the Hamiltonian of the system. This approach still remains computationally expensive. Finally, the recursive Green function approach [136, 137, 142] offers a valuable alternative since it reduces the computational cost by exploiting the structure of the Hamiltonian, and offers the possibility to be generalized to arbitrary geometries. It is important to note that such a method has been applied indifferently to both electronic transport [137] and heat transport [143].

7.2 The recursive Green function Approach

While the full-quantum behavior of a given conduction channel would be given by the solution of the Schrödinger equation, the numerical approximation to such solutions still constitutes a computational burden for general geometries and 3D systems. However, as discussed in chapter 5, obtaining solutions to the Schrödinger equation on 1D or 2D cross sections is achievable within a reasonable time-frame [90, 85, 92]. It is also remarkable that, in the case of rectangular cross sections or finite 1D cross sections with hard wall boundary conditions, an analytic expression exists for the eigen-functions and eigen-vectors of the Schrödinger equation. Besides, as we are mainly interested in silicon nanowires with the prospect to extend our results to general nanostructures, including graphene nano-ribbons, it is necessary that the following computational scheme be seamlessly applicable to either 1D or 2D cross-sections.

The Green function is commonly used in physics to find the solution to general partial differential equations. In the case of the Schrödinger equation

$$(H(\mathbf{r}) - E)\,\psi(\mathbf{r}) = 0 \tag{7.5}$$

where H is the Hamiltonian of the system and E its energy, the Green function $G(\mathbf{r}, \mathbf{r}')$ is the solution to the equation

$$(H(\mathbf{r}) - E)\,G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \tag{7.6}$$

where $\delta(\mathbf{r} - \mathbf{r}')$ is the Dirac function. Without entering in a detailed discussion on Green functions, let us merely mention the following result, which will be central for the remaining of the discussion. In the event of an applied perturbation potential $V(\mathbf{r})$, the Green function becomes the solution to the perturbed problem

$$\left(H(\mathbf{r}) + V(\mathbf{r}) - E \right) G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \tag{7.7}$$

If the system is discretized over N points in space $\{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N\}$, solving Schrödinger equation will yield the N first eigen-energies E_n. Additionally, in a vector representation $\mathbf{V} = \{V(\mathbf{r}_1), V(\mathbf{r}_2), ..., V(\mathbf{r}_N)\}$, where $\mathbf{G}$ becomes a $N \times N$ matrix, the Green function of the perturbed system may be derived from the unperturbed Green function $\mathbf{G}_0$ according to Dyson's equation

$$\mathbf{G} = \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V} \mathbf{G} \tag{7.8}$$

Since we are primarily concerned with the transport properties of our conduction channels, it is also remarkable that the transmission and reflection coefficients for the nano-channel may be directly obtained from the transverse energy eigen-values and the Green function of the system. Thus, we shall inspect in the next paragraphs the RGF approach to computing the transmission probability $T(E)$ and introduce its novel application to the derivation of quantum enhanced surface roughness scattering rates in sub - 10 nm nano-channels.

We present here the 2D version of our approach which is the first to have shown physically consistent results and sets a rather understandable ground for a more general discussion. The essence of the problem is represented in Fig. 7.3. In order to calculate the transmission and reflection coefficients for an electron moving along a conduction channel and impinging on geometrical variations, we exploit the fact that the plane wave motion can be described by the tight-binding (TB) Hamiltonian on a lattice with periodicity δx, where δx is much lesser than the electron wavelength λ ($\delta x \ll \lambda$).

In order to compute the scattering amplitude, we choose a scheme similar to the one introduced by Sols et al. [137], in which the 2D or 3D transmission channel is modeled by a 1D tight binding chain, where each site represents a cross section of the whole tight binding channel. For a cross-section containing N lattice sites, the TB Hamiltonian yields N eigen-modes. We consider that our portion of rough nano-channel is in contact with two semi-infinite and perfectly smooth nano-channels of the same diameter. To understand how a mode m entering the channel at the left is transmitted into a mode n
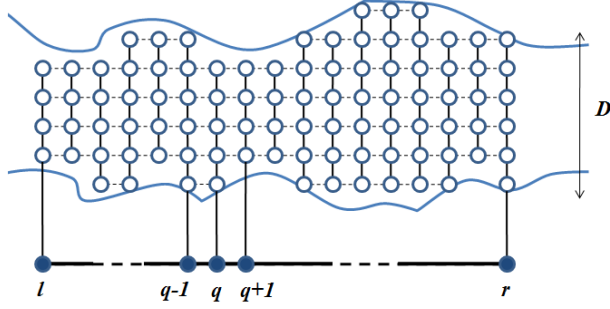
Figure 7.3: 2D Model of the recursive Green function approach in a rough nanowire. Each transverse chain is represented by a single site in the 1D system, and the propagators G are computed to connect the energy modes between neighbor sites along the chain.

on the right, the Green function matrix $\mathbf{G}_{lr}$ needs to be computed. If one quasi-particle enters the channel with a momentum k, its total energy E, in a cross-sectional mode m of energy E_m, is given by

$$E = \frac{\hbar^2 k^2}{2m^*} + E_m \tag{7.9}$$

where m^* is the electron effective mass in the semiconducting material considered. In the absence of applied potential, the TB scheme requires that a hopping potential V should be added in order to retrieve the continuum case when δx is decreased to 0. To enforce this condition, the hopping potential should be

$$V = \frac{\hbar^2}{2m^*(\delta x)^2} \tag{7.10}$$

With these considerations, the transmission probability t_{mn} from an eigenstate $|m>$ at the leftmost site l to an eigenstate $|n>$ at the rightmost site r is

$$t_{mn} = -2iV(\sin\theta_m \sin\theta_n)^{1/2} e^{i(\theta_m l - \theta_n r)} < n|\mathbf{G}_{rl}|m > \tag{7.11}$$

with $\theta_n \equiv k_n \delta x$ and k_n is the wave vector available for longitudinal motion, as indicated in (7.9).

For 1D cross-sections with hard walls, a closed analytic form of the eigenstates can be derived. For a rough 2D cross section, it is possible to compute the first eigen-states within a reasonable time frame. Once the first relevant cross-sectional eigen-states are computed for each site of the 1D longitudinal TB chain, the whole problem sums up to the computation of the Green function matrix $\mathbf{G}_{rl}$. To serve this purpose, we make use of Dyson's equation in a fashion similar to that of Sols et al. [137], although with minor

modification to account for the fact that each site is treated independently and no ensemble of neighbor sites exhibit in general a closed form for their propagator $\mathbf{G}$.

Since the propagator $\mathbf{G}$ is known for a perfectly smooth semi-infinite lead, the algorithm functions in a recursive way. Once the right lead is attached, at the step q of the recursion, the propagator $\mathbf{G}^R_{r,q+1}$ between site r and $q+1$ on the right portion of the channel is known. The site q is then attached in order to determine the new propagators $\mathbf{G}^{L+R}_{r,q}$ and $\mathbf{G}^{L+R}_{q,q}$ such that

$$
\begin{aligned}
\mathbf{G}^{L+R}_{r,q} &= \mathbf{G}^R_{r,q+1}\mathbf{V}_{q+1,q}\left(\mathbf{I}-\mathbf{G}^L_{q,q}\mathbf{V}_{q,q+1}\mathbf{G}^R_{q+1,q+1}\mathbf{V}_{q+1,q}\right)^{-1}\mathbf{G}^L_{q,q} \\
\mathbf{G}^{L+R}_{q,q} &= \left(\mathbf{I}-\mathbf{G}^L_{q,q}\mathbf{V}_{q,q+1}\mathbf{G}^R_{q+1,q+1}\mathbf{V}_{q+1,q}\right)^{-1}\mathbf{G}^L_{q,q}
\end{aligned}
\tag{7.12}
$$

where the subscript L describes the propagator of a single site, which has a closed form as specified in [137]. The matrices $\mathbf{V}_{q,q+1}$ effectively make the connection between the neighbor sites by providing the correlation between transverse modes in neighbor TB cross-sections such that

$$
(\mathbf{V}_{q,q+1})_{nm} = V\sum_j \psi_n^*(j)\psi_m(j)
\tag{7.13}
$$

where $\psi_n(j)$ and $\psi_m(j)$ are the transverse wave functions computed on the jth cross-sectional mesh point at the sites q and $q+1$ respectively. The index j spans the ensemble of possible cross sectional mesh points which connect the sites q and $q + 1$. At the end of the recursive algorithm, the leftmost semi-infinite lead is finally attached.

7.3   Non-Parabolic Bands

In the model presented above, equation 7.9 implicitly assumed parabolic bands. While this assumption is accurate a low energies, it is known that drastic deviations from parabolic bands appear at higher energies. The critical energy at which transition away from parabolic bands occurs may in some cases be relatively low, in particular in silicon $\mathbf{L}$-valleys, where deviation is observed at only 0.1 eV above the bottom of the conduction band [86]. However, as we develop a frequency dependent approach to the study of interferences between electron waves and geometrical features, we would like to capture the higher order energy variation resulting from band non-parabolicity. In addition, in narrow channels, in the case of high carrier densities, or under high fields, the carrier energy will often be well above

the bottom of the conduction band. In this case, the validity of equation 7.9 may be questionable, and non-parabolicity must be considered.

In order to account for non-parabolicity in the $E - \mathbf{k}$ relationship, several options naturally appear. On the one hand, one could expand the carrier energy as a Newton series in the wave-vectors

$$E(\mathbf{k}) = \sum_{n=1}^{N} a_{2n}\mathbf{k}^{2n} + o(\mathbf{k}^{2N+2}) \tag{7.14}$$

One could subsequently use this expansion to the order N directly in the Schrödinger equation to include non-parabolicity [144, 145]. It is also possible to use this relationship to estimate an energy dependent effective mass to include in Schrödinger's equation; however, this approach has led to erroneous results [145].

Another method consists in expanding the wave-vector as a Newton series in the carrier energy

$$\frac{\hbar^2\mathbf{k}^2}{2m^*} = E(1 + \alpha E + \beta E^2 + ...) \tag{7.15}$$

This is the most common approach, and has been widely applied to a variety of materials and device structures [146, 147, 126]. The second order expansion is often referred to as the Kane approximation.

Based on Kane's truncation, we can rewrite Schrödinger's equation using the total energy

$$E = \frac{1}{2\alpha}\left[\sqrt{1 + 4\alpha\frac{\hbar^2\mathbf{k}^2}{2m^*}} - 1\right] \tag{7.16}$$

In the case of quasi-1D transport along the x-axis, induced by a $2D$ confinement along the transverse direction by a potential $V(y, z)$, and in the absence of confinement in the longitudinal direction, it is possible to introduce the ansatz for the electron wave function

$$\psi(x, y, z) = e^{ik_x x}\zeta(y)\xi(z) \tag{7.17}$$

This form allows one to distinguish between transverse and longitudinal energy levels such that $E = E_x + E_{//}$. In particular, using the effective mass tensor $m^* = \{m_x^*, m_y^*, m_z^*\}$, we retrieve [18, 148]

$$E_x(1 + \alpha E_x) = \frac{\hbar^2 k_x^2}{2m_x^*} \tag{7.18}$$

It is interesting to note the approach of Godoy et al. [18] who retrieve,

through a series expansion over the energy dependence of equation 7.16, the expression of the coupling between transverse and longitudinal energy levels such that, if the potential inside the boundaries of the wire is uniformly zero

$$E_{//}(1 + 2\alpha E_x + \alpha E) = \frac{\hbar^2 k_y^2}{2m_y^*} + \frac{\hbar^2 k_z^2}{2m_z^*} \tag{7.19}$$

Here, the energy level $E_{n0} = (\hbar^2 k_y^2/2m_y^* + \hbar^2 k_z^2/2m_z^*)$ is the energy level which would be retrieved by solving the tight binding Schrödinger equation in the transverse cross section under the parabolic expression. As a result, it is possible to include non-parabolicity in the recursive Green function scheme as follows. The particle enters a section of nanowire with a known total energy E. In each cross-section, the value of E_{n0} is computed through a TB Schrödinger solver. Subsequently, it results from equation 7.19 that the energy available for propagation at energy level n is shifted down to the value

$$E_x^n = \frac{1}{2\alpha}[\sqrt{1 + 4\alpha(E - E_{n0} + \alpha E^2)} - 1] \tag{7.20}$$

This value of energy is finally used in the expression of θ_n in computations of $\mathbf{G}_{q,q}$ and transmission coefficients in equation 7.11, such that $E_x^n = 2V(\cos\theta_n - 1)$.

From the equations above, it appears that the principal effect of band non-parabolicity is a shift down of the energy levels in the transport direction. Evidence of this phenomenon is revealed from the computations carried out in the next sections (Fig. 7.13 on page 124)

Finally, we would like to emphasize the fact that, despite our choice of 3D geometry, the same technique is similarly applicable to the case of 2D transport with no modification, provided the use of the E_x and $E_{//}$ notations. Our approach will subsequently be used to simulate transport of charge carriers longitudinally confined in 2D ribbons and 3D wires.

7.4   Preliminary Results on 2D Transport

Based on the scheme described above, we wish to compute the transmission coefficient of conducting channels in the scope of the novel derivation of surface roughness scattering rates derived from quantum mechanical considerations. In order to get a better understanding of the problem and to calibrate our version of the RGF algorithm, several cases will be considered, at first, in a 2D topology. The extension to 3D geometries solely involves solving the Schrödinger equation on 2D cross-sections instead of

1D cross-sections, and is the object of next section. The 2D model presents two natural advantages. First, we will be able to explore key properties of transport along a rough channel without going through the additional computational burden of dealing with real geometries. Second, the 2D picture of rough conducting ribbons recently attracted a renewed interest in the scope of the extraordinary heat and electrical conduction properties of graphene ribbons. In this context, understanding what processes lead to scattering of electron and heat waves with surface asperities is key to the design of novel graphene based devices.

Here, the quantities of interest are the transmission probabilities $T_m(E)$ of the input modes m and the total transmission probability $T(E)$. The latter quantity is directly related to the number of quanta of conductance which will be excited on the channel at low temperature in the Landauer formalism of (7.4). Indeed, each mode m constitutes a conduction channel which, at best, carries in parallel with the other channels an electrical current with the elementary conductance $2e^2/h$. Thus, the total transmission probability for an input mode m is

$$T_m(E) = \sum_j |t_{mj}|^2 \tag{7.21}$$

where j goes over all allowed energy modes in the output lead. Incidentally, the total transmission probability for all allowed transmission lines in the conduction channel is

$$T(E) = \sum_m T_m(E) \tag{7.22}$$

The first test case consists in a perfectly smooth nano-channel. Here, the channel is connected to two semi-infinite leads of the same diameter $D = 10$ nm. The electron effective mass is $m^* = 0.05m_0$ in order to keep consistency with the results from Sols et al. [137]. This test case presents little interest other than guaranteeing that our RGF accurately predicts a constant transmission probability of 1 for each allowed energy level in the channel given an input energy E (Fig. 7.4). In order to understand the following discussion, it is important to emphasize that an electronic wave entering the nanochannel has either:

- an energy E superior to a given energy level n, and inferior to level $n+1$. Consequently, the n first channels may transmit the wave, which is reflected by a transmission probability of 1 for T_1, T_2, ..., T_n. The total transmission $T(E) = n$.
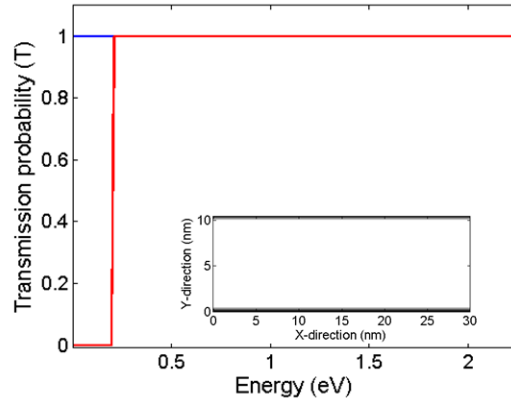
Or

Figure 7.4: Transmission probability for the first two modes in a perfectly smooth wire. Solid line: T_1, dashed line: T_2, inset: geometry.

- an energy inferior to the first energy level. The transmission probability is zero on every channel, and the wave cannot be transmitted.

Here, the total transmission coefficient takes discrete integer values since the geometry is constant along the direction of propagation and all cross-sections are identical. Geometrical fluctuations will introduce mixing between different modes, leading to non integer values of the transmission coefficient.

The trivial case of the smooth ribbon, however, does not provide information about mode leaking. In order to test this situation, we will compare our results to the special case of the quantum modulated transistor (QMT) which was proposed by Sols et al. [136] and subsequently studied using a RGF scheme. Here, the geometry consists in a T-stub structure, with input and output diameters of 10 nm, and a central extension of 10 nm. In essence, the QMT employs a gate voltage to modulate the effective length of the extension and allow resonance of different modes in the channel. Similarly to the results of Sols et al. [136], our model predicts several zeros in the transmission of the first two modes, as well as regions of unity amplitude (Fig. 7.5). It is remarkable that the transmission of the first mode is unequivocally degraded after the appearance of the second eigen-mode, due to mixing between different modes. While the agreement is acceptable, the discrepancy found with the data of [136] is explained by the rather coarse grid which was employed. As a result, the effective extension width was slightly lower than 10 nm, accounting for different resonance peaks. As a concluding remark on the QMT, let us put the emphasis, again, on the fact that the low temperature conductance is directly related to $T(E)$. Hence, the current in the QMT should correspondingly go to zero for certain Fermi
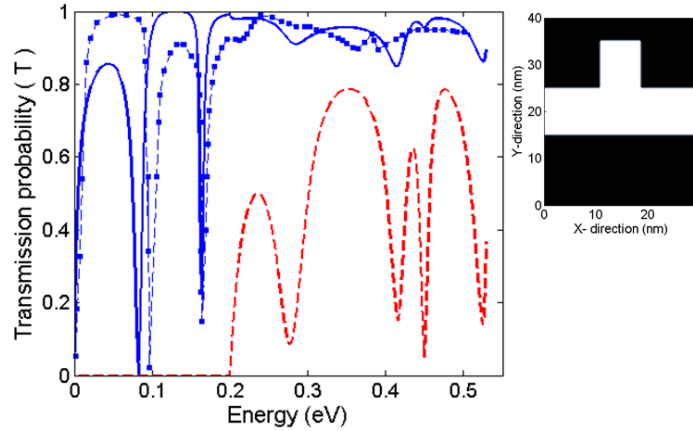
113

Figure 7.5: Transmission probability for a quantum modulated transistor as a function of the input energy E. Solid line: T_1, dashed line: T_2, squares: Sols et al. [136].

energies. Such behavior was experimentally demonstrated [149], although the stark deviations predicted by the theory were not observed.
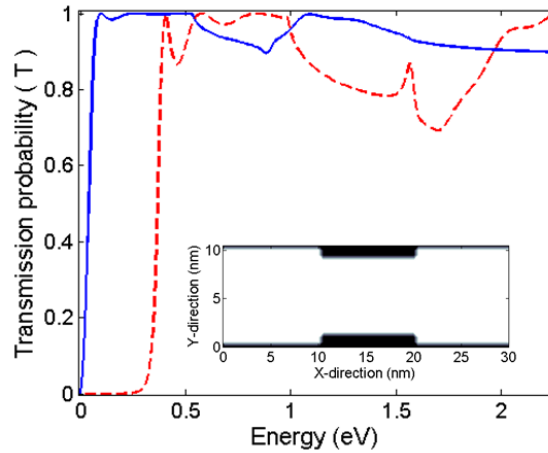
Figure 7.6: Transmission through a constriction as a function of input energy E. Solid: T_1, dashed: T_2.

Next, in the scope of understanding the mechanisms of transport in a rough channel, let us first inspect the case of a single constriction. In this geometry, the width of the 10 nm channel is reduced to 8 nm at the center of the channel, reproducing in this way the average effect of one single roughness constriction of rms $\Delta = 0.5$ nm. Figure 7.6 shows the transmission coefficient in this case. Here, the second mode appears exponentially, at an energy much higher than the 0.2 eV of the perfectly smooth case. The reason for this effect lies in the increase of the second energy level in the
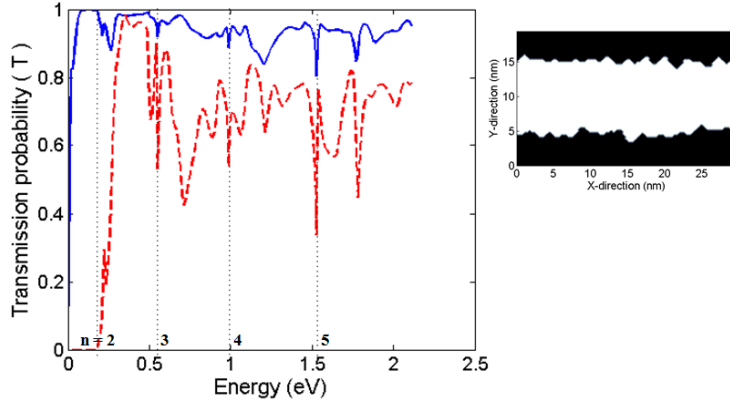
114

Figure 7.7: Transmission through a rough semi-conducting nano-channel as a function of input energy E. Solid: T_1, dashed: T_2.

constriction. Hence, the traveling wave effectively impinges on a potential barrier, and only the modes carrying an energy higher than the second constriction energy may travel through the constriction. In addition, since the constriction length is finite, it is possible to observe tunneling through the barrier. This phenomenon is correctly depicted by the exponential build-up, and later oscillations, of the transmission coefficients of the first two harmonics. Thus, the leaking of the first mode into subsequent modes appears at higher energies than the cross-sectional eigen-energies of the input lead.

Incidentally, we inspect the case of a rough nano-ribbon, which can be thought of as a series of constrictions of limited length (Fig. 7.7). We used for this geometry a channel of average diameter $D = 10$ nm, roughness rms $\Delta = 0.5$ nm, and correlation length $L = 2$ nm. It is noticeable that the effect of tunneling into the first mode is almost washed out in comparison to the single constriction. In general, the transmission of the first mode remains good even at high energies, while the transmission of the second mode is drastically degraded. It is also remarkable that the transmission of each mode is sharply damped at energies corresponding to the eigen-modes of the smooth channel of average diameter D. As of now, no explanation has been proposed which accounts for this phenomenon.

Finally, the total transmission $T(E)$ is computed according to equation 7.22 (Fig. 7.8). The attentive reader should know at this point that the total conductance of the channel, including all transmission lines from different input modes, is directly related to $T(E)$ in Landauer formalism. In essence, $T(E)$ is a measure of the number of quanta of elementary conductance which are involved in the conduction process along the scatterer. As discussed earlier, the QMT shows maxima of conductance close to the ideal case,
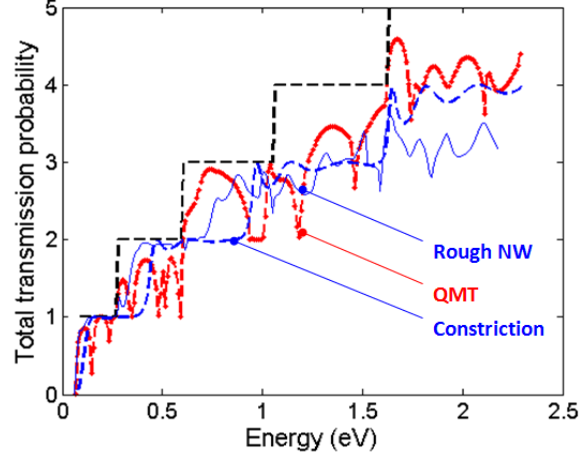
Figure 7.8: Total transmission coefficient $T(E)$ for the perfectly smooth nano-channel (dashed plateaus), QMT, constriction, and rough nano-channel with similar input and output lead diameter $D = 10$ nm.

with the ability to shut off the conduction on certain modes depending on the length of the extension. The nano-channel with only one constriction shows a behavior close to the perfect transmission channel with a smaller diameter equivalent to the cross-section width at the constriction. Spurious oscillations are introduced, in this case, by tunneling of waves through the constriction barrier. The rough nanowire shows a conduction close to ideal at low energies whereas its behavior is starkly altered at higher energies. Since the higher energy harmonics are barely transmitted, the conductance seems to exhibit a convergence toward a finite average value. This is in accordance with our discussion on phonon transmission in rough nanowires, where we have considered that quasi-particles of wavelength λ that is big in comparison to the roughness critical dimensions will see little contribution from the perturbation at the edges.

7.5   3D Non-Parabolic Simulation of Transport in Si NW

In the previous section, considering the case of rough 2D ribbons gave valuable insight into the processes of scattering of electronic waves from geometrical fluctuations. In order to go through a series of constrictions along the propagation direction, the wave needs sufficient energy. In addition, the protuberances lead to the appearance of local QMT configurations which shut off transmission at given frequencies, depending on the height of these asperities. Here, we want to apply our knowledge to the case of real nanowires,

116

including band non-parabolicity and 3D geometries. Hence, we explore the effect of geometrical fluctuations on the electronic transport in rough Si nanowire (NW) thermoelectric devices of diameter $D < 10$ nm. At this scale, the quantum nature of transport is accounted for in the computation of energy dependent transmission coefficients through the recursive Green function algorithm. In the following section, the rough 3D NW geometry is used as a direct input to simulations through the roughness height Δ and autocovariance length L. Using this approach, we will demonstrate that the channel conductance above 0.1 eV is drastically reduced in rough Si NW with high D/Δ ratio. In addition, we will see that the roughness induced resistivity is only increased by 6% on the first energy level of 10 nm Si channels with $\Delta = 7.7$ Å, showing possible application for high thermoelectric figures of merit ZT.

Indeed, we are consistently working under the natural assumption that scattering of charge carriers with the surface limits transport in nanowires, owing to their high surface-to-volume aspect ratio. Recent experimental and theoretical work on such NWs [7, 19] showed that, through careful nano-engineering of intentionally rough edges, it is possible to decouple thermal and electrical transport, making such devices extremely appealing for novel applications in thermoelectric energy conversion. In particular, in thin NW of diameter $D <15$ nm, where electrons are pushed away from the surface, phonons may be scattered more efficiently than electrons by asperities at the NW surface [120], leading to high thermoelectric figures of merit ZT. As discussed in chapters 5 and 6, conventional formalism which accounts for electron – surface roughness scattering is based on the derivation of a perturbed Hamiltonian of the system due to geometrical fluctuations [15, 16] and is adapted to a 2D electron gas (2DEG) conducting near an interface. However, at low NW scales, this approach does not account for the quantum nature of electron transport, where charge carriers theoretically impinge on a series of potential barriers resulting from geometrical constrictions in the transport direction [150]. Thus, in this section we compute an energy dependent electrical quantum conductance in thin Si NW from a recursive Green function (RGF) approach [137], including evanescent modes, non-parabolic bands, a 3D description of NW, and a model for arbitrary rough surfaces through the root mean square (RMS) roughness height Δ and autocorrelation length L. Rough interfaces are directly taken into account through the definition of randomly variable NW geometries, using the method described in section 7.4. Based on this approach, and in the absence of other scattering mechanisms, we will show that surface roughness contributes to an

increased resistivity for high energy carriers in Si NW of diameter $D < 10$ nm, while having little effect on the lower energy levels, thus building good hopes for application of rough thin NW in thermoelectrics.

As a reminder, the quality of thermoelectric devices is measured by the dimensionless figure of merit $ZT = (S^2\sigma/\kappa)T$, where S is the Seebeck coefficient, T the temperature, σ the electrical conductivity and κ the thermal conductivity. In this respect, the standard approach which we have pursued in chapter 6 in order to achieve performance is to design a "phonon glass - electron crystal" with high electron mobility and poor thermal conductivity. Recently, thermal conductivity in Si NW with etched rough edges was experimentally reduced by a factor of about 100 in comparison to bulk crystalline Si [7, 19], to nearly the value of amorphous Si. In order to provide supplementary information for the design of efficient thermoelectric devices, the present model relates the transmission of charge carriers in a rough Si NW conduction channel to the NW diameter, the experimental roughness parameters Δ and L, and the electron energy E. While considering geometrical fluctuations as the sole source of resistivity in NW, the transmission probability $T(E)$ of carriers with energy E is directly linked to the electrical current I flowing between two Büttiker probes held at Fermi energies $E_{F,1}$, $E_{F,2}$ through Landauer's formula $I = 2e/h \int_{E_{F,1}}^{E_{F,2}} T(E)\,dE$. Similarly to what was discussed in the case of ribbons in the section above, the channel conductance, in the ideally smooth case, will vary in steps of integer multiples of the fundamental quantum conductance e^2/h.

The 3D conduction channel is modeled by a 1D tight binding (TB) chain where each site q represents a cross section of the whole 3D TB domain (Fig. 7.3). The NW is assumed in contact with two semi-infinite leads of perfectly smooth NW of the same diameter. The quantum nature of NW transport is accounted by solving the tight binding Schrödinger equation (TBSE) in a series of 2D cross sections along the transport direction, and recursively computing the transmission between adjacent sections through the overlap of the wave functions [137] $(\mathbf{V}_{q,q+1})_{nm} = V \sum_j \psi_n^*(j)\psi_m(j)$, where $\psi_n(j)$ and $\psi_m(j)$ are the transverse wave functions at the jth cross-sectional mesh point at sites q and $q+1$ respectively, and $V = \hbar^2/(2m_x^* dx^2)$ is the hopping potential between adjacent sites. The domain is discretized on a 3D mesh of spacing $dx = dy = dz$ where x is the transport direction. The NW boundaries are defined by a potential step equal to the Si work function $\Phi_{Si} = 4.8$ eV, while the interior of the wire is kept at 0 eV, hence modeling a Si NW suspended in vacuum. In addition, the simulation domain is larger than the actual NW, thus allowing eigenfunctions to diffuse out of the NW
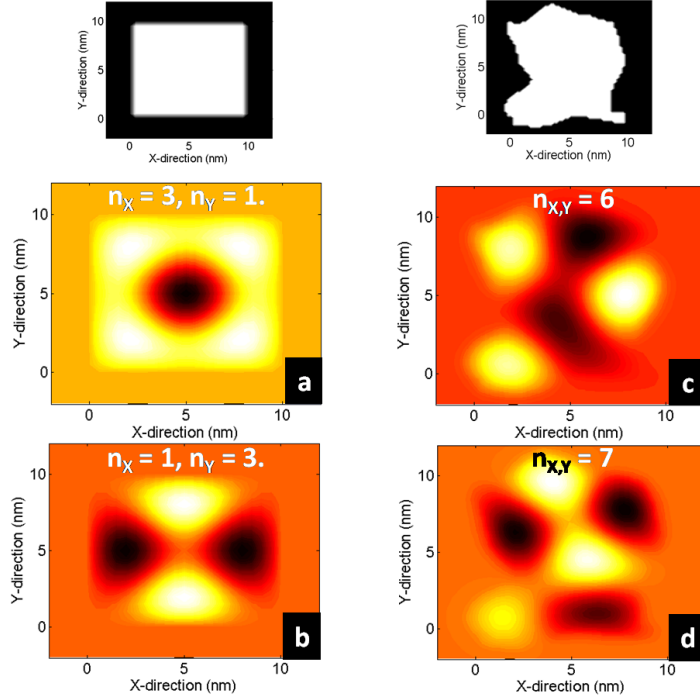
Figure 7.9: (a), (b) Examples of cross-section wave functions in a perfectly smooth NW. (c), (d) Examples of cross-section wave functions of equivalent order in a rough NW with $\Delta = 0.5$ nm.

boundaries (Fig. 7.9).

It is remarkable that this approach allows one to directly size the effect of the nanowire surface roughness on the cross-sectional eigen-functions. From Fig. 7.9, it is seen how surface roughness alleviates degeneracy in the cross section, and introduces a large number of wave-functions in a given range of energies. Indeed, classical quantum theory gives an analytic expression for the eigen-energies in the case of the perfectly smooth square cross-section nanowire

$$E_{n_Y, n_Z} = \frac{\pi^2 h^2 n_Y^2}{2m_y^* D} + \frac{\pi^2 h^2 n_Z^2}{2m_z^* D} \tag{7.23}$$

each resulting from a particular combination of modes along the y and z directions. As such, the eigen-functions corresponding to the pair $(n_Y, n_Z) = (1, 3)$ and $(3,1)$ have the same transverse energy: the transverse states are degenerate. Nevertheless, in rough wires, it appears that such degeneracy is relieved due to the lack of symmetry. Furthermore, it is remarkable to see, as in the case of Fig 7.9.c and Fig 7.9.d, that roughness triggers the appearance of sub-patterns. In these cases, the constriction fosters wave-function configurations which appear like combinations of cross-sectional states in smaller, smoother cross sections, each delimited by the principal
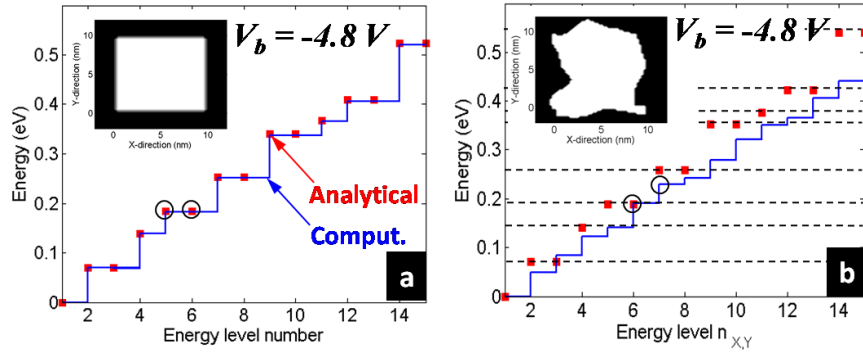
119

Figure 7.10: Computation of the first energy levels in the nanowire crossection corresponding to Fig. 7.9. (a) Smooth square cross section. (b) Rough cross section.

constriction. Hence, in Fig. 7.9, a close analogy is found with combination of sub-states $(1,3) + (1,2)$ and $(3,1) + (1,2)$. Yet, those sub-states are not completely independent from each other, and no simple linear rule may be applied to derive the corresponding states of energy. Figure 7.10 represents the first energy levels for the respective cross-sections mentioned above. In the case of the smooth cross-section, the computed energies match exactly with the analytical, hence validating our Schrödinger solver. Similarly, computations on the rough cross-sections demonstrate how roughness alleviates degeneracy, and shows the increased number of eigen-values in a given energy range.

The number of eigen-values which can be computed in a cross-section depends on the discretization scheme used to model the nanowire. Indeed, if the cross-sections are discretized on $N_y \times N_z$ mesh points, the resulting Hamiltonian will be numerically represented by a $N_y N_z \times N_y N_z$ square matrix. Consequently, a maximum of $N_y N_z$ eigen-functions may be computed. Sizing the computational scheme is then a trade-off between competing effects:

- One needs enough eigen-modes to compute the transmission on the desired energy range, including the effect of evanescence of higher energy modes.

- One needs a grid fine enough to capture the geometrical variations of the nanowire surface due to roughness. Thus, large roughness rms Δ and autocorrelation length L usually allow for coarser grids.

- The required computational time to obtain the eigen-functions and energies on each cross-section scales up with the number of computed
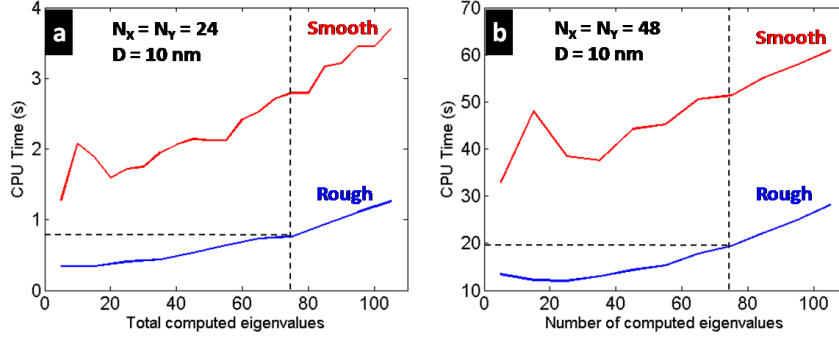
120

Figure 7.11: Computational time needed to solve the eigen-value problem in smooth and tough cross-sections with (a) 576 mesh points and (b) 2304 mesh points. $D = 10$ nm and $\Delta = 0.5$ nm.

eigen-modes and the number of mesh points.

In our simulation, Schrödinger's equation is solved using Arnoldi's iteration and a Choleski factorization on the sparse symmetrical Hamiltonian matrix, as provided by the ARPACK package used in MATLAB [151]. Here, the Hamiltonian remains symmetrical because the nanowire boundaries are only treated as a step in the electrical potential and effective mass. Based on this approach, we provide in Fig. 7.11 an estimation of the computational time required to solve the eigen-value problem in one cross-section. At a given number of approximated eigen-values, the computational effort increases in the case of a finer mesh. It is also interesting to notice that the computational time to resolve the eigen-values of rough cross section is smaller. This is a direct consequence of the disappearance of degenerate states. Finally, we would like to estimate the number of eigen-values that need to be estimated in order to solve the transmission problem in the given energy-range. In this respect, Fig. 7.12 depicts the eigen-values in cross sections of silicon nanowires ($m^* \approx 0.19m_0$) with smooth and rough boundaries over a wide energy range. For greater accuracy in the 0 eV - 1 eV energy range, non-parabolic bands are used. Since surface roughness introduces a higher density of energy modes in a given range, it must be noticed that, in the case of rough wires, the gain in computational effort resulting from the non-degeneracy is balanced by the necessity to approximate a larger number of modes. Hence, in rough silicon wires of 10 nm diameter, it will be necessary to compute about 75 modes in order to capture the effect of higher mode tails on the 0 ev - 1 eV range. In this scope, we estimate that approximating the 75 first modes of Schrödinger's equation in each cross section will require about 20 s of time on a single core 1.66 GHz CPU using
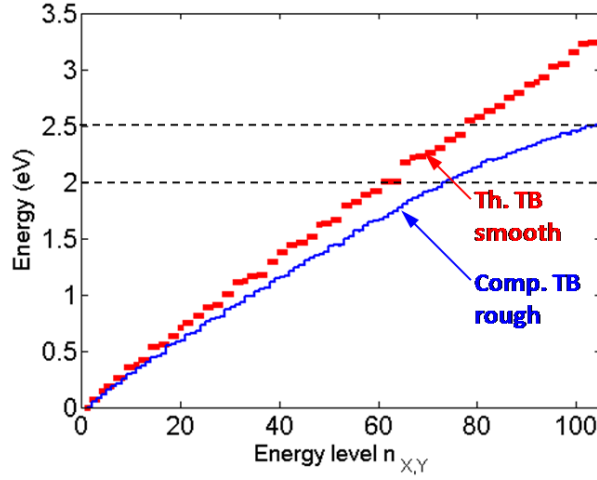
Figure 7.12: 2D cross-section energy levels in rough ($\Delta = 5$ Å) and smooth silicon nanowires, with $D = 10$ nm and $m^* = 0.19m_0$.

the method described above.

Once the eigen-values and eigen-energies are computed, the transmission probability from eigen-state m at the leftmost site l to state n at the rightmost site r is

$$t_{mn} = -2iV(\sin\theta_m\theta_n)^{1/2}e^{i(\theta_m l - \theta_n r)}\langle n|\mathbf{G}_{rl}|m\rangle \tag{7.24}$$

where $\theta_n \equiv k_{x,n}dx$, $k_{x,n}$ is the wave vector available for longitudinal motion, and $\mathbf{G}_{rl}$ is the Green Function (GF) propagator matrix between r and l. The total transmission at energy E is $T(E) = \sum_{m,n} |t_{mn}(E)|^2$. The NW resistivity due to geometrical constrictions is then obtained from $\mathbf{G}_{rl}$. Since the propagator $\mathbf{G}$ is known for a perfectly smooth semi-infinite lead, the algorithm functions in the recursive way described in section 7.2, and pivots on equation 7.12. Non-parabolicity is included in the computation of the longitudinal wave vectors [18, 148]

$$E_{y,z}\left[1 + 2\alpha(E_x + E_{y,z})\right] = E - E_x \tag{7.25}$$

where α is the coefficient of non-parabolicity, $E_{y,z}$ is the TBSE cross section energy, and $E_x = \hbar^2 k_x^2/(2m_x^*)$ is the effective energy available for longitudinal transport. As a result of non-parabolicity, transverse energy levels appear shifted down (Fig. 7.13). Conduction is assumed along the $\mathbf{\Gamma} - \mathbf{X}$ direction with $m_x^* = 0.19m_0$ and $\alpha = 0.5$. Non-parabolic bands allow better accuracy in the computation of roughness limited $T(E)$ at $E > 0.1$ eV.

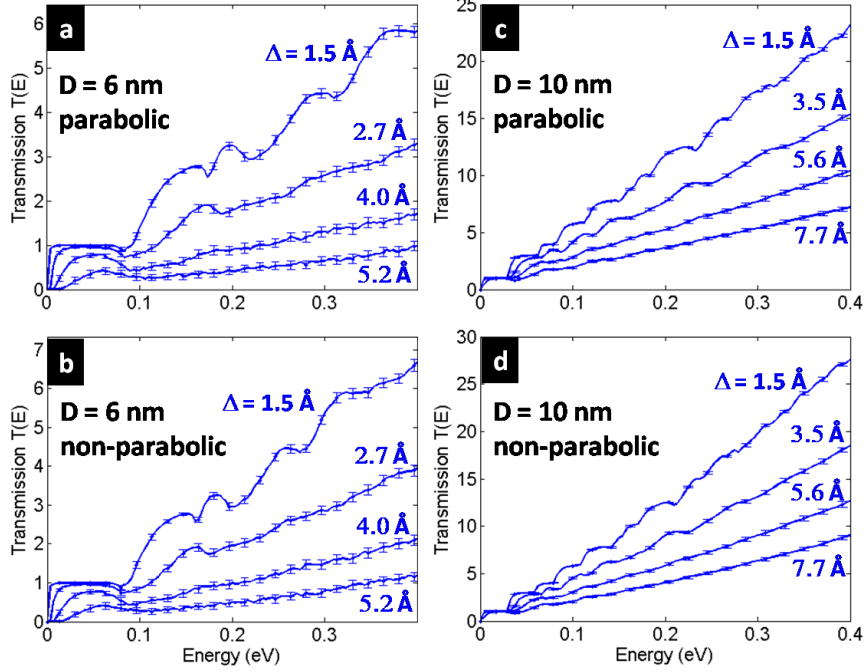We model transport in 100 nm long NW of average width $D < 10$ nm.

Figure 7.13: Effect of Δ on the quantum transmission in rough Si NW of diameter 6 nm (a), (b), and 10 nm (c), (d). $L = 22$ Å throughout.

Surface roughness alleviates the degeneracy of the cross section eigen-modes (Fig. 7.9). Subsequently, the computation of $T(E)$ is averaged over 500 randomly generated NW. The effect of RMS height Δ is compared at different diameters and varying energy (Fig. 7.13). Increasing Δ results in a series of higher potential steps encountered along the x-direction, thus reducing the average transmission on rough NW. The quantum features of conduction disappear at high Δ where the step-like behavior is smoothened as a consequence of carrier back-scattering from surface asperities. In addition, the resistivity incurred at similar RMS height is stronger in NWs of smaller diameter.

Noticeably, surface roughness has limited impact on the transmission of the first energy level, particularly at $D = 10$ nm, in accordance with Wang et al. [150]. Thus, geometrical fluctuations are expected to have only a minimal influence on the low field NW mobility. In this regime of operation, where thermal conductivity varies as $(D/\Delta)^2$ [19], it is expected to achieve high $ZT > 1$. The RMS height Δ measures the average height of roughness peaks or constrictions. This statistical picture is consistent with the case of chemically grown NW, where one may control the amplitude of geometrical variations rather than the actual shape of the edge. In 10 nm NW with $\Delta = 7.7$ Å, we observed constrictions of effective diameter down to 6 nm.

Hence, for high Δ/D ratio, transmission does not occur until sufficiently high energy is reached to tunnel through constriction barriers. Nevertheless, the surface is generally comparable to a random noise around the average diameter. As a result, higher variance of NW behavior is observed when E is near the transition between energy levels. The standard deviation of NW transmission is also increased at high Δ/D ratio, but remains small at the center of energy steps.

As a conclusion, we explored the effect of surface roughness on electronic transport in thin Si NW of diameter $D < 10$ nm based on a recursive Green function algorithm. This approach allows a 3D representation of the NW geometry through the statistical parameters Δ and L, and further includes the effect of non-parabolicity of the Si band structure. As the quantum nature of charge carriers is accounted for, an energy dependent transmission is computed for such devices. With geometrical fluctuations as the sole electronic decay process, conductance at energies $E > 0.1$ eV is drastically decreased in such NW with high Δ/D ratio. Nevertheless, surface asperities have little effect on low energy transport, leading to possibly high thermoelectric figure of merit ZT in this regime of operation. To test such assumption, it is necessary to additionally model thermal transport in channels with dimensions below the average phonon mean free path.

7.6 Effect of Geometrical Fluctuations on a quantum modulated transistor

In section 7.4, we used the quantum modulated transistor (QMT) as an example to validate the RGF approach on results by Sols et al. [136] using a similar scheme. Nevertheless, the device presents interesting characteristics based on quantum modulation of the electronic transmission along its direction of transport. As Sols et al. foresaw valuable application for the device in new generations of transistors operating at very low energies, few factors precluded further study of the latter configuration. Noticeably, the required size of the device, with features below 10 nm, was too costly for fabrication until 3-4 years ago. In addition, the QMT, whose behavior essentially revolves around the waveguide nature of the device, was expected to suffer from the effect of real surfaces, effectively randomizing resonance of the waves in a cavity of varying width. In this section, we use our non-parabolic RGF scheme to extend the results of Sols et al. [136] to the case of GaAs based rough QMTs, taking into account the effect of band non-parabolicity.
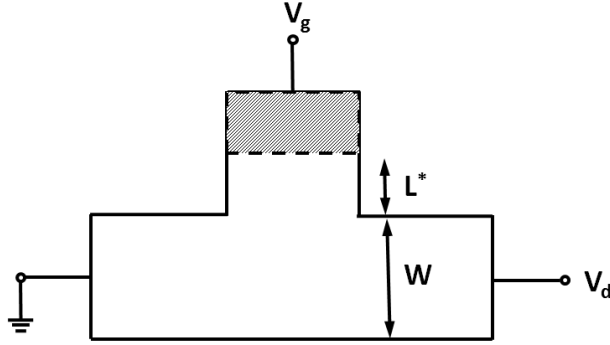
Figure 7.14: Model of a quantum modulated transistor (QMT). Electronic waves are modulated by a stub whose effective length L^* is adjustable by an applied gate voltage, similarly to a T-shaped electromagnetic waveguide.

As a result, we show that, despite a degradation of the damping at cut-off energies, the characteristics of GaAs QMTs remain acceptable for applications in low energy transistors in spite of surface roughness with rms ranging up to 1.5 nm for a 10 nm wide conduction channel.

In essence, the QMT is a T-shaped electronic waveguide (Fig. 7.14). Similarly to an electromagnetic waveguide, the frequencies of resonance in the propagation cavity are controlled by the length of a stub. As a result, depending on the length of the stub, with propagation from left to right, it was shown that transmission is cut off at a series of electronic wavelengths. The main difference in the electronic case consists in the penetration length of electronic waves in the stub, which is effectively adjustable by an applied gate voltage. Incidentally, the energy of transmitted carriers is controlled by an applied voltage V_d. In that respect, the device provides a transistor effect that operates purely on the ground of a modulation of the quantum electronic waves. In addition, a major benefit of the QMT lies in the fact that quantum modulation may in essence be achieved for single carrier transport, bringing effective application of QMTs to low energy transistors.

We model in Fig. 7.15 the total transmission coefficient in a QMT with respect to the carrier energy and the effective length of the stub. Within this scope, we used a stub width and a channel width of 10 nm, and coefficients of non-parabolicity $m^* = 0.05m_0$ and $\alpha = 0.8$ which best reflect the properties of GaAs. Noticeably, the length of the conduction channel is fixed to 30 nm total. While the length of a smooth channel is irrelevant to the problem in absence of other scattering mechanisms (the modal transmission is invariantly equal to 1), the introduction of surface roughness scattering
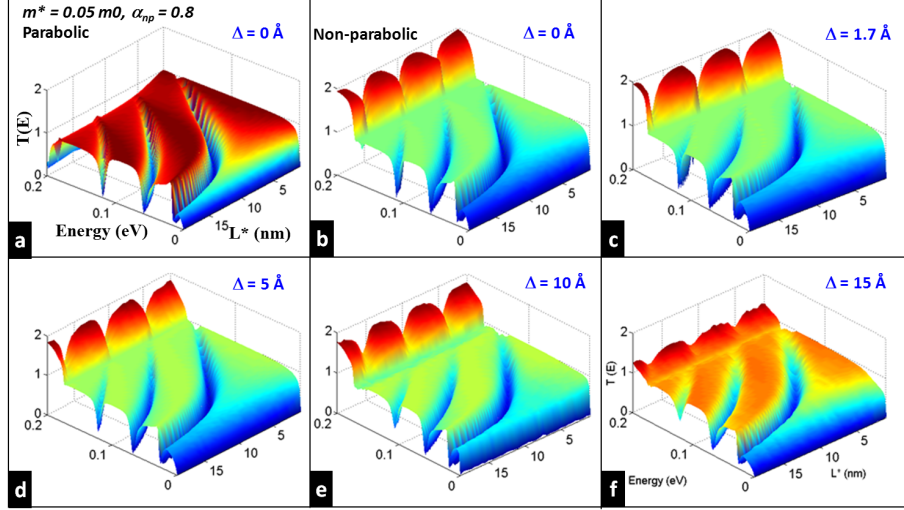
Figure 7.15: Total transmission in a rough quantum modulated transistor as a function of the stub length L^* and carrier energy. (a) Perfectly smooth device with a parabolic band approximation, and $m* = 0.05m_0$. (b)-(f) Non-parabolic band approximation, with $m* = 0.05m_0$ and $\alpha_{np} = 0.8$, corresponding to a GaAs approximation [136]. Roughness varies in the device, with (b) $\Delta = 0$ Å, (c) $\Delta = 1.7$ Å, (d) $\Delta = 5$ Å, (e) $\Delta = 10$ Å, and (f) $\Delta = 15$ Å.

makes transport dependent on the length of the waveguide. Similarly to the case of silicon NWs, the effect of band non-parabolicity results in a decrease of the characteristic energy levels available in the transport direction (Fig. 7.15.a and 7.15.b). Subsequently, the roughness rms of the edges is increased (Fig. 7.15.c to Fig. 7.15.d). In the case of the QMT, the effect of surface roughness smoothens the singularities of the transmission coefficient. In particular, the maximum transmission achievable is attenuated at higher surface roughness. This effect is observed in greater proportions at higher carrier energy levels. In addition, the minimal transmission coefficient at the cut-off energies is increased for higher surface roughness rms. Effectively, high surface roughness reduces the pass to cut-off conduction ratio, thus diminishing the efficiency of the device for transistor applications (Fig 7.16). Nevertheless, the cut-off features remain present at high rms $\Delta > 1$ nm, and the cut-off energies remain unchanged. It is remarkable that computations show a shift in energies for $\Delta = 1$ nm, though we believe that the latter effect is an artifact of the coarseness of the grid on the simulation domain. We conclude by observing that the effect of band non-parabolicity effectively reduces the expected cut-off energy levels, and the surface roughness leaves the device usable for transistor applications on the first energy level. In particular, we have shown that a GaAs based QMT
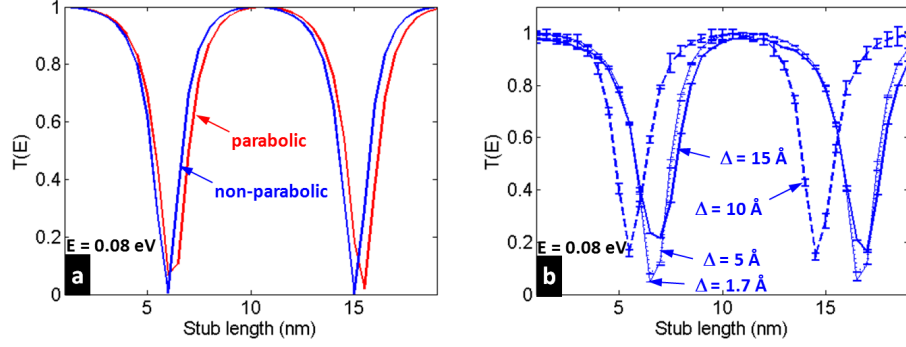
126

Figure 7.16: Total transmission in a rough quantum modulated transistor at energy $E = 0.08$ eV. (a) Effect of non-parabolic bands. (b) Effect of surface roughness rms.

shows practical cut-off features in the 0-0.18 eV energy range, with a max cut-off / pass transmission ratio of 0.16 at a carrier energy of 0.08 eV, for the effective stub length varying from $L^* = 0 - 20$ nm. Finally, we emphasize that the roughness rms effectively makes the overall channel resistivity dependent on the channel length. Overall, the non-parabolic rough RGF scheme appeared as a powerful tool to quickly assess the value of QMT for low energy application given rough geometries. In the last section, we will use the latter approach to examine further prototypical ideas based on benefits from quantum interference of the charge carriers based on channel geometry.

7.7 Perspectives on Mono-Layered Materials

In the previous sections, we derived the impact of geometrical fluctuations on the electronic properties of 2D and 3D conduction channels, taking into account the wave nature of the charge carriers at the nanometer scale. In particular, we introduced a framework based on a Green function formalism which accounts for band non-parabolicity in semiconductor materials. Simulation of transport in 2D sheets is intrinsically advantageous on a theoretical level for its easier inclusion of geometrical effects. More importantly, an absolute confinement in a direction to achieve the state of a 2D electron sheet has strong consequences for the electronic density functions. Recently, experimental advances in nanoscale fabrication processes put graphene at the forefront of promising materials for 2D electronic and heat transport, with electrical mobility and thermal conductivity over an order of magnitude greater than silicon [152, 153]. As such, graphene is a two-dimensional crys-
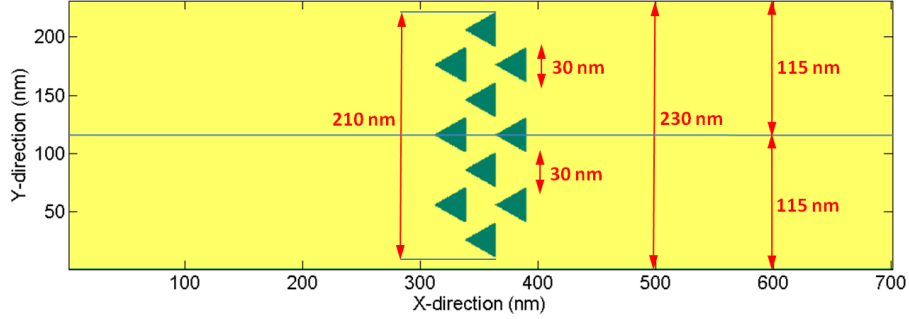
Figure 7.17: Model of transport along a semiconductor ribbon with periodic patterns of asymmetric constrictions along the direction of propagation.

tal of sp^2-bonded carbon atoms, whose electronic properties can be tuned by an applied electric field, thus triggering growing interest for its application in field effect transistors (FET) [154, 155].

The advantage of deriving a general computational scheme lies in the ability to quickly conduct and test thought experiments, and avoid costly prototypical iterations. As we derived in the past sections a general scheme for nano-scale simulation of transport in non-parabolic semiconductors, we will extend in this section the domain of simulations to new types of devices. In the previous chapters, we mainly focused on the statistical fluctuations of surfaces and edges. While it is experimentally impossible to factor out the effect of surface roughness - perfectly smooth surfaces do not exist in reality - one can further imagine a repeatable, periodic device geometry that catalyzes the physical mechanisms in play with surface roughness scattering. In this scope, we propose here an additional study on transport in 2D semiconductor sheets periodically patterned with asymmetric constrictions along the direction of propagation (Fig. 7.17).

The latter geometry raises a priori a series of theoretical questions. First, the periodicity of the patterns could be expected to interfere with the electronic waves along the direction of propagation. Hence, one could imagine that this geometrical configuration be used to amplify the transmitted signal through resonance of particular energy modes. Alternatively, the inverse effect could be observed, in the sense that the repeated patterns could simply shut off transmission at certain wavelengths. Energies at which constructive or destructive quantum interferences are achieved will likely be a function of the periodicity and size of the patterns on the ribbon. Secondly, the intrinsic asymmetry of the patterns along the direction of propagation is of interest. As such, it will be interesting to assess whether electronic transmis-
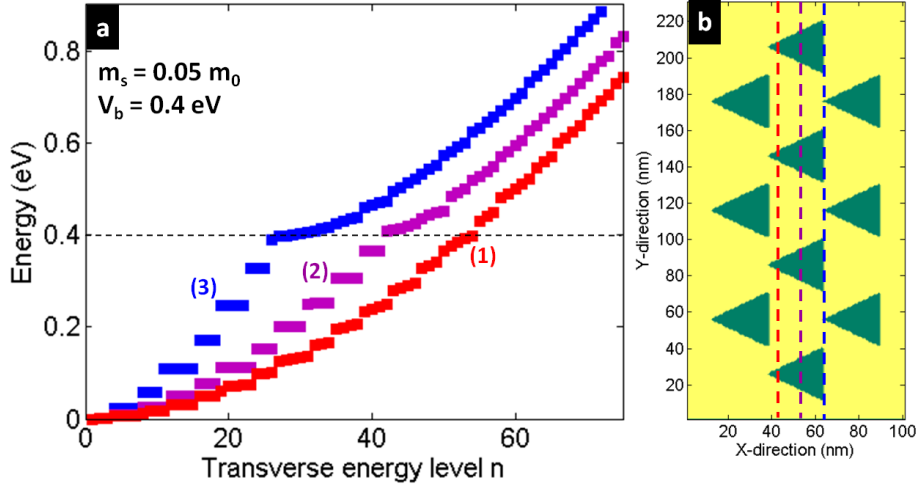
Figure 7.18: (a) Energy levels in three different cross-sections of the patterned structure. (b) Associated geometrical representation of the cross sections for a semiconductor ribbon. Energy levels are derived from solutions to the Schrödinger equation in the cross section adjusted to the energy along the propagation direction. A non parabolic band model is used with $m^* = 0.05m_0^*$, $\alpha = 0$, and and $V_b = 0.4$ V throughout.

sion shows different properties in left-to-right and right-to-left conduction. In this case, asymmetrically patterned nano-ribbons could find interesting applications in polarized devices.

Figure 7.18 depicts the computed energy levels in different cross sections along the propagation direction. Here the patterns are modeled by a potential barrier, with an effective applied bias of $V_b = 0.4$ V. As a result, different energy modes resonate in the periodic potential wells in the cross sections. We use here a square mesh of 1 nm $\times$ 1 nm in order to solve the TBSE in the cross sections. Noticeably, carriers of energy above 0.4 eV in the transverse direction see a quasi-continuum of energy levels above the energy barrier, where most of the quantization occurs from the confinement in the 230 nm ribbon only. When moving from left to right, (moving from cross sections 1 to 3 in Fig. 7.18), carriers see a reduction in the number of allowed energy levels, as the energy steps become wider in narrower potential wells. As a result of the constrictions narrowing along the propagation direction, an increasing number of carriers of energy below 0.4 eV are either back scattered or decay along the propagation direction until an allowed energy level is reached. Similarly, in the reverse transport direction (left to right), carriers first reach narrow constrictions, after which the energy levels adjust to an increasing number of energy levels available in the widening constrictions.

We included the latter computations of cross section wave functions in the
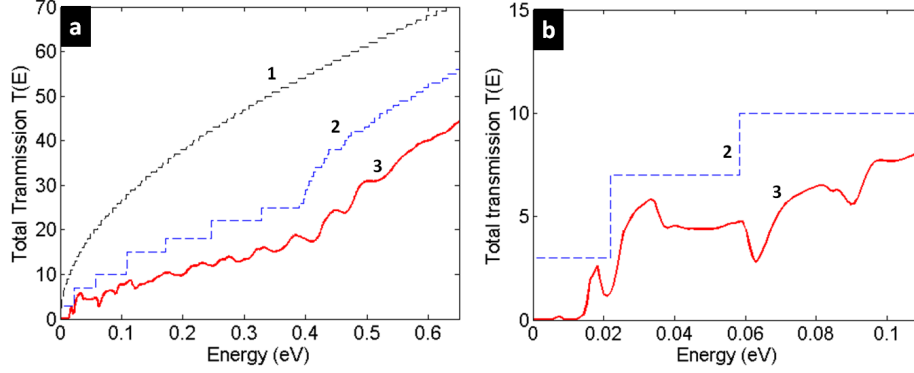
129

Figure 7.19: (a) Total transmission in the asymmetric ribbon depicted in Fig. 7.17. **1**: Total transmission in the channel without patterns. **2**: Coupled parallel stripes of width 30 nm, with identical cross-sections throughout the propagation direction. **3**: Triangle patterns. (b) Close-up view on the 0-0.11 eV energy range. $m^* = 0.05m_0^*$ and $V_b = 0.4$ V throughout.

Green function model in order to simulate transmission across the asymmetric channel (Fig. 7.19). For comparison, curve 1 in Fig. 7.19.a depicts transport in a similar channel without patterns, as a result of the same Green function simulation model, where confinement is only due to the overall ribbon width (230 nm). In addition, transport is simulated in a ribbon with symmetric patterns (curve 2), constituted of 30 nm wide stripes spaced by 30 nm, and an applied bias of 0.4 V. This scenario represents the case where waves are equally confined in potential wells of invariant width along the propagation direction, with the minimal well width encountered in the asymmetric case and the possibility for waves to tunnel through the potential barriers in the cross section. Curve 3 illustrates the total transmission in the asymmetric channel represented in Fig. 7.17. In general, we notice that the total transmission in the asymmetric ribbon is less than in the case of striped ribbon, thus yielding a greater electrical resistivity in the asymmetric ribbon. This should be attributed to the misalignment of patterns along the propagation direction, which prevents clear overlap of the cross section wave functions. Incidentally, it must be noticed that patterned ribbons exhibit transmission curves similar to the non-patterned ribbons above a critical energy level. For the striped ribbons used above, this energy level lies in the 0.4 eV - 0.5 eV energy range. For ribbons with asymmetric patterns, the critical energy lies around 5.3 eV. The latter higher level is due to the non-overlapping wave-functions along the propagation direction, increasing the energy at which an incident wave would experience quantized propagation through successive decays in the direction of transport. It is noticeable

that resistivity is increased in the striped ribbons and asymmetric ribbons respectively. This phenomenon is a result of the smaller number of energy levels available for transmission along the direction of transport. Looking at the asymmetric patterns, an exponential build-up of the transport plateaus reveals tunneling through the potential barriers along the propagation direction, similarly to the rough nano-ribbon case. A series of energy levels allows fewer transitions, accounting for repetitive drops in the total transmission.

Incidentally, we notice that the total transmission curves are identical in the forward and backward directions. We believe this effect is a consequence of the combination of several factors. First, it is expected that most of the quantization results from the size of the smallest constriction. As such, forward and backward traveling waves will experience the same effect from the constrictions of smallest width along the propagation direction. Secondly, we expect the misalignment of the asymmetric patterns to contribute greatly to this effect. Indeed, the misalignment of patterns yields, in two consecutive cross sections of the TB Green function model, misaligning wave functions. Transmission coefficients between the misaligning wave functions of lowest order are expected to be the same in the backward and forward direction, hence greatly contributing to the observation of identical total tranmssion.

In this section, we used the recursive Green function algorithm to study electronic conduction in a regularly patterned semiconductor ribbon. The wave-like nature of the charge carriers was accurately captured and showed the effects of size quantization and interferences, yielding to step-wise conduction and areas of negative resistance in the electronic transport along the ribbon. In addition, the model allows one to drill down and study wavefunctions and eigen-energies in particular cross-sections. As a result, we were able to explain the observed behavior of the configuration of periodic asymmetric patterns. In particular, the latter ribbon configuration presented, a priori, the expectation of different electronic conductivity in the forward and backward direction. We demonstrated the contrary with RGF methodology without having recourse to experimental prototyping of the device. Failure to exhibit asymmetric transport properties was determined to result from the misalignment and size of the patterns. Nevertheless, unidirectional conduction is an interesting property, and we recommend further investigation of different pattern configurations based on a similar approach.

7.8 Conclusion

In this chapter, we introduced a novel approach to estimating the quantum enhanced surface roughness scattering rate for nano-channels of width below 10 nm. As the channel size is decreased to the order of the electron wavelength, it is necessary to take into account the waveguide nature of the problem. The variation of confinement width with the roughness of the channel interfaces will strongly interfere with the energy modes allowed in the cavity. Indeed, in the extreme case, the quantum modulated transistor proves that it is possible to completely shut off a transmission mode with careful engineering of the confinement width. Our model is based on a recursive Green function approach, which, as demonstrated in this chapter, provides a reasonable trade-off between the accuracy with which the problem is treated and its computational cost. Thus, we were able to compute the total transmission probability for a single electron in the low temperature regime for nano-channels with a single constriction, and rough nano-channels made of a series of statistically distributed constrictions. Similarly to our discussion in chapter 6, we exhibited the energy dependence of the roughness scattering term, as high energy particles will experience a stark reduction in their transmission probability when impinging on rough nano-channels. In particular, we explored the effect of surface roughness on electronic transport in thin Si NW of diameter $D < 10$ nm. We introduced in that respect a new RGF method, which allows for a 3D representation of the NW geometry through the statistical parameters Δ and L, and further includes the effect of non-parabolicity of the Si band structure. Based on the latter approach, we accounted for the quantum nature of carriers in ultra-thin conduction channels, as we computed an energy dependent transmission in rough 3D Si NW. With geometrical fluctuations as the only electronic decay process, conductance at energies $E > 0.1$ eV is drastically decreased in NWs with high Δ/D ratio. Nevertheless, surface asperities have little effect on low energy transport, leading to possibly high thermoelectric figure of merit ZT in this regime of operation. In order to test such assumption, it is necessary to transpose the RGF method to thermal transport in thin conduction channels, in order to take into account quantum effects on the heat carriers as a result of geometrical fluctuations.

Subsequently, the same non-parabolic approach was employed to study the effect of real geometries on quantum modulated transistors. The principal effect of band non-parabolicity is a shift down of the energy levels in the transport direction. The QMT function lies in a shut-down of transmission

132

at a series of energies controlled by an applied gate voltage. While band non-parabolicity lowers the values of the series of non-transmitted energy levels, and while surface roughness allows for some conduction along the shut-down modes, their joint effect yet remains minimal, fueling strong hopes for real applications of quantum modulated transistors. As we modeled 2D QMTs for consistency with results of Sols et al. [136], we recommend further investigation of the device with prototyping and 3D RGF simulation to validate the concept on GaAs structures.

Finally, we benefited from the computational efficiency of the RGF scheme in order to assess the behavior of prototypical devices based on regular patterning of semiconductor nano-ribbons. In this case, we have shown limitations in the expected asymmetric conduction due to the intrinsic misalignment of patterns. Mainly, it was possible to determine key effects of the regularity of the patterns on the quantum conductance without going through costly prototyping of the devices in the laboratory. In general the model could be seamlessly applied to transport in graphene, where the major difference lies in the approximation of the particle effective mass. Most importantly, the main difficulty in graphene resides in the Dirac point at the center of the FBZ, where the effective mass approximation does not hold. This difficulty disappears in graphene nano-ribbons, where the boundary conditions enforce band splitting at the Γ point.

We conclude this discussion by stating that such an approach remains equally applicable to phonon transport, as it was recently demonstrated by Yang et al. [143]. In this case, each phonon branch must be considered independently. In research on optimal design, this advance could form, with the model introduced here, a complete framework to assess the viability of sub-10-nm devices based on geometrical patterning of edges as efficient thermoelectric devices.

CHAPTER 8

CONCLUSION AND PERSPECTIVES

In this dissertation, we introduced a series of novel computational schemes that allowed efficient statistical simulations of surface limited electronic, mechanical, and thermal processes in devices at the micro- and nano-scale. As the size of devices is reduced, coupling of different physical processes and the effect of atomistic irregularities at the device surface greatly contribute to the observed performance of micro-contacts, semiconductor wires, and sheets. In addition, such phenomena introduce a high statistical variation in the characteristics of sample devices built on the same design model, and the behavior of integrated systems may result in non-linear phenomena due to the statistical distribution of the performances of similar devices. In this scope, we position this work in a novel multi-scale framework that we introduced in this work. Hence, we base the highest level of the multi-scale hierarchy on statistical Monte Carlo and percolation principles in order to average the time dependent non-linear evolution of integrated systems based on devices with statistically varying characteristics. At the lower scale, the physical characteristics of devices are derived from semi-classical properties of heat and electronic transport in nano-scale devices, the quantum nature of carriers in devices whose dimension go below 10 nm, and the atomistic knowledge of the surface properties.

Throughout this work, we focused our attention on the effect of surface roughness at different scales. In the area of M/NEMS micro-contacts, we modeled the effect of surface roughness on the mechanical, electrical and thermal properties of metallic contacts leading to thermal failure of such devices. Given the scale of the device, statistical inclusion of surface roughness in Kirchhoff's and Euler's classical laws was sufficient to accurately model this class of devices. M/NEMS contacts were subsequently modeled in the non-contacting state, where our percolation framework was used to approximate the time evolution of electro-thermal failure of such devices due to non-linear gas discharge between metallic electrodes. At the nano-scale, we further included the effect of surface roughness on the semi-classical solu-

tion of Boltzmann's transport equation (BTE) in semiconductor channels and mono-layered sheets, where electronic and heat transport is mostly limited by a series of constrictions along the propagation direction. Finally, at scales below 10 nm, surface roughness was directly included in the quantum solution of the system, where the effect of interface randomness on the transmission of quantum wave packets was observed.

In scope of this work, computational efficiency of the simulations is desirable in order to favor cost-efficient exploration of new areas of design without using expensive iterations on experimental prototypes. In order to support this work, we introduced a number of novel statistical schemes. First, we described a statistical model that allows estimation of the roughness rms Δ and autocorrelation length L from the experimental observation of surfaces and edges. In addition, these observable parameters can be used to computationally generate a series of random surfaces. In this scope, we introduced a new model to approximate the contact resistance in MEMS micro-contacts with respect to the applied pressure. The algorithm revolves around a Gaussian approximation of surface roughness, and a model of the intersection between a series of randomly generated rough contacts. This approximation provides a good fit with experimental observations on gold-gold contacts. Based on this model, we report minimal achievable contact resistance of 80 mΩ for applied pressures above 0.3 mN on 1 μm $\times$ 1 μm surfaces of sputtered gold with observed roughness rms $\Delta = 7$ nm and autocorrelation length $L = 60$ nm. In the non-contacting regime, we use a percolation method in order to simulate electro-thermal breakdown in micro-contacts through gas discharge. The model is based on a statistical distribution of conductive defects in the insulating gas due to the presence of sputtered metal created through the repetitive contacting cycles of the device. Through a statistical Monte Carlo method on a series of different random gas configurations, we were able to model the time to breakdown in such devices under electrode separations in the 500 nm - 1μ m range, and applied bias voltage of 1 V - 90 V. The same observable roughness parameters L and Δ were further used in semi-classical models of electronic and heat transport in rough semiconductor nanowires and nano-ribbons. In particular, we studied the effect of surface roughness on electronic and heat resistivity of wires through novel scattering rates based on L and Δ which we derived from Fermi's golden rule and perturbation theory. In the particular case of heat transport, we modeled for the first time the extreme reduction of thermal conductivity in rough silicon nanowires of width below 100 nm, down below 5 W/m/K for $\Delta = 4$ nm and width $W = 20$ nm, close to the limit of amorphous silicon. Using

perturbation theory to average the effect of random surface scattering of phonons over an ensemble of carriers, the model shows excellent agreement with experimental observations. In the case of semiconductor channels of width $W < 20$ nm, we further introduce a 3D non-parabolic recursive Green function scheme to approximate the effect of surface roughness on the propagation of quantum electronic waves along a semiconductor wire of varying rough cross-sections. The model shows physically consistent results and agreement with previously published simulation of quantum devices, and shows for the first time the effect of non-parabolicity and surface roughness on electronic transmission in wires below 10 nm.

Based on the validation of the simulation frameworks presented in this dissertation against experimental results, we further used these models to explore novel design patterns based on careful engineering of surface roughness to utilize coupled properties of electronic and heat transport. In the area of micro-contacts, we report in particular that the minimal achievable contact resistance will be mostly minimized by reduction of L. On the other hand, in order to decrease the necessary applied pressure to reach the minimal resistance, the main parameter to optimize for is Δ. In the area of nanowires, our knowledge of surface limited transport led to the design, in this dissertation, of a an electronic device based on electrostatic confinement of charge carriers in a quasi-1D channel. As a result, such devices show reduced channel resistivity and increased mobility in comparison to standard nanowire field effect transistors. Our further extension of the semi-classical framework to heat conduction led to the consideration of engineered surface roughness as a way to decouple electronic and heat transport in nanowires. Extending the results on silicon channels led to the prediction of extremely low thermal conductivity in germanium and gallium arsenide based nanowires below 1 W/m/K. Such results on electro-thermal decoupling hold promise for the further consideration of Ge and GaAs nano-engineered rough wires as efficient, on demand, site specific, thermoelectric coolers or generators for integrated circuits. In addition, the same model was extended to such novel materials as graphene, where the effect of edge roughness was included in semi-classical models of heat transport in rough suspended graphene nano-ribbons. As a result, we show low thermal conductivity below 1000 W/m/K in such devices, in agreement with recent experimental observation. Finally, extending our recursive Green function model of quantum transport in ultra-thin channels, we explored new types of devices, based on a quantum operation of the transistor effect. In this scope, we showed minimal influence of surface roughness on semiconductor quan-
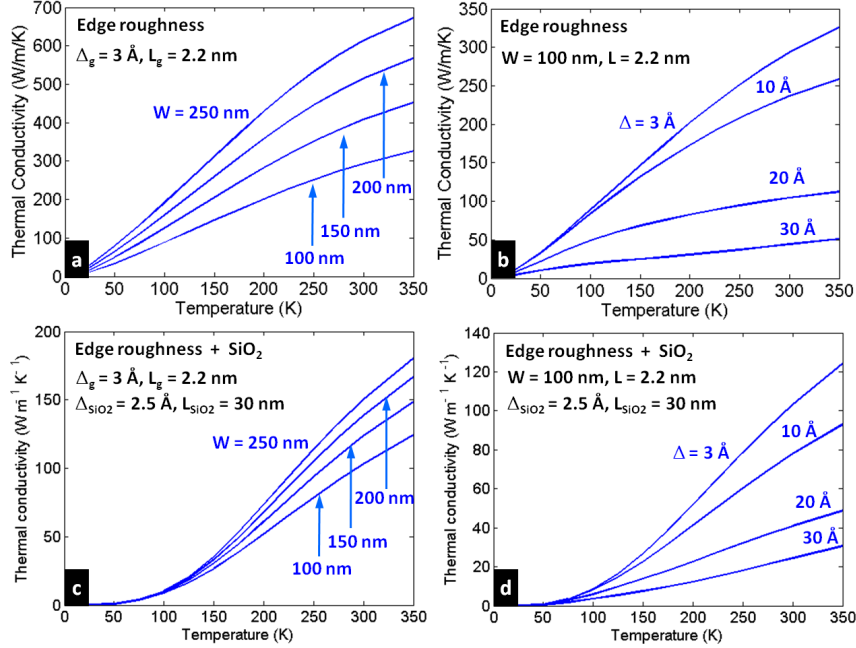
Figure 8.1: (a) and (b): predicted effect of width and edge roughness on the thermal conductivity of suspended graphene nano-ribbons of width $W < 250$ nm. (c) and (d): Additional effect of oxide phonon leaking and oxide roughness in graphene ribbons deposited on SiO_2.

tum modulated transistors (QMT). In addition, we were able to predict the effect of asymmetric engineered patterns embedded in semiconductor sheets.

The existence of the statistical simulation approaches that we introduced in this dissertation and their good performance in accordance with experimental observations leads to a vast field of opportunities for further research. As new materials appear in the laboratory, graphene presents a promising ground for extraordinary electronic and heat transport in 2D mono-atomic layers. In that respect, we explored edge roughness as a design parameter to decouple heat conduction from electronic conduction in suspended graphene for thermoelectric applications. Incidentally, on deposited graphene, the fluctuations of the substrate on which the semiconductor sheet is deposited present a second degree at which surface roughness can be engineered. In particular, sheets will conform to the oxide asperities, and out of plane lattice vibrations may subsequently be shut off with careful engineering of oxide roughness, as we modeled in Figs. 8.1 and 8.2. In view of this, we recommend experimental observations on the effect of edge and surface roughness on deposited rough graphene ribbons of width $W < 200$ nm, for their possible thermoelectric applications. In addition, as we showed the feasibility of real quantum modulated transistors, we further recommend experiments on
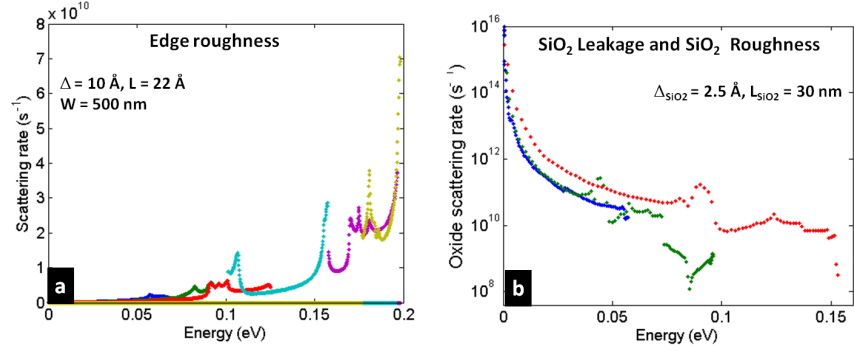
Figure 8.2: (a) Phonon – edge roughness scattering rate in graphene nano-ribbons. (b) Phonon-oxide scattering rate in graphene nano-ribbons due to phonon leaking in the oxide and oxide roughness scattering.

such devices, as they show interesting properties for low energy field effect transistor applications. On theoretical grounds, the multi-physics aspects of thermoelectric phenomena are still full of challenges. In particular, exploration of the effect of device geometry on the coupled thermopower is yet an area of uncertainty. In that respect, we would imagine a further study of these characteristics through inclusion of our semi-classical scattering rates in coupled electron-phonon semiconductor Monte Carlo simulations.

REFERENCES

[1] S. Senturia, N. Aluru, and J. White, "Simulating the behavior of MEMS devices: Computational methods and needs," *IEEE Comp. Sci. Eng.*, vol. 4, no. 1, pp. 30–43, 1997.

[2] K. Entesari, K. Obeidat, A. Brown, and G. Rebeiz, "A 2575-MHz RF MEMS tunable filter," *IEEE Trans. Microw. Theory Tech.*, vol. 55, p. 2399, 2007.

[3] C. Goldsmith, A. Malczewski, Z. Yao, and S. Chen, "RF MEMS variable capacitors for tunable filters," *Int. J. RF Microw.*, vol. 9, pp. 362–374, 1999.

[4] N. Singh et al., "High-performance fully depleted silicon nanowire (diameter below 5 nm) gate-all-around CMOS devices," *IEEE Electron Dev. Lett.*, vol. 27, no. 5, pp. 383–386, 2006.

[5] W. Steinhögl, "Comprehensive study of the resistivity of copper nanowires with lateral dimensions of 100nm and smaller," *Appl. Phys. Lett.*, vol. 97, 2005.

[6] C. Dames and G. Chen, "Theoretical phonon thermal conductivity of Si/Ge superlattice nanowires," *J. Appl. Phys.*, vol. 95, p. 682, 2004.

[7] A. Hochbaum et al., "Enhanced thermoelectric performance of rough silicon nanowires," *Nature*, vol. 451, pp. 163–167, 2008.

[8] A. Boukai et al., "Silicon nanowires as efficient thermoelectric materials," *Nature*, vol. 451, pp. 168–171, 2008.

[9] D. Li et al., "Thermal conductivity of individual silicon nanowires," *Appl. Phys. Lett.*, vol. 83, p. 2934, 2003.

[10] D. Lacroix et al., "Monte Carlo simulation of phonon confinement in silicon nanostructures: Application to the determination of the thermal conductivity of silicon nanowires," *Appl. Phys. Lett.*, vol. 89, 2006.

[11] R. Chen et al., "Thermal conductance of thin silicon nanowires," *Phys. Rev. Lett.*, vol. 101, no. 10, 2008.

[12] S. Goodnick and D. Ferry, *Transport in Nanostructures*, 1st ed. Cambridge University Press, 1997.

[13] S. Yamakawa et al., "Study of interface roughness dependence of electron mobility in Si inversion layers using the Monte Carlo method," *J. Appl. Phys.*, vol. 79, no. 2, 1996.

[14] S. Goodnick, R. Gann, and D. Ferry, "Surface roughness scattering at the Si-SiO2 interface," *J. of Vac. Sci. Tech. B*, vol. 1, pp. 803–808, 1983.

[15] S. Goodnick et al., "Surface roughness scattering at the Si-SiO2 interface," *J. Vacuum Sci. Tech. B*, vol. 1, no. 3, pp. 803–808, 1983.

[16] S. Yamakawa et al., "Electron mobility and Monte Carlo device simulation of MOSFETs," *VLSI Design*, vol. 6, pp. 27–30, 1998.

[17] J. Watling et al., "Si/SiGe heterostructure parameters for device simulations," *Solid State Electronics*, vol. 48, p. 1337, 2004.

[18] A. Godoy, Z. Yang, and U. Ravaioli, "Effects of nonparabolic bands in quantum wires," *J. Appl. Phys.*, vol. 98, p. 013702, 2005.

[19] P. Martin, Z. Aksamija, E. Pop, and U. Ravaioli, "Impact of phonon-surface roughness on thermal conductivity of thin Si nanowires," *Phys. Rev. Lett.*, vol. 112, p. 125503, 2009.

[20] P. Martin, Z. Aksamija, E. Pop, and U. Ravaioli, "Reduced thermal conductivity in nano-engineered rough Ge and GaAs nanowires," *Nano Letters*, vol. 10, p. 1120, 2010.

[21] D. Peroulis, K. Sarabandi, and L. Katehi, "Low contact resistance series MEMS switches," *IEEE Microwave Symp. Dig.*, vol. 1, p. 223, 2002.

[22] S. Majumder et al., "Study of contacts in an electrostatically actuated microswitch," *Sensors and Actuators A*, vol. 93, p. 19, 2001.

[23] H. Kwon et al., "Investigation of the electrical contact behaviors in Au-to-Au thin-film contacts for RF MEMS switches," *J. Micromech. Microeng.*, vol. 18, 2008.

[24] M. Nakamura, "Ballistic thermal conductance of a graphene sheet," *IEEE Trans. Comp., Hybrids, and Manufac. Tech.*, vol. 16, p. 339, 1993.

[25] R. Holm, *Electric Contacts, Theory and Applications*, 1st ed. Berlin:Springer, 1967.

[26] J. Greenwood and J. Williamson, "Contact of nominally flat surfaces," *Proc. Royal Soc. London A*, vol. 295, p. 300, 1966.

[27] B. Streetman and S. Banerjee, *Solid State Electronics*, 5th ed. Englewood Cliffs, NJ: Prentice Hall, 2000.

[28] F. Gamiz et al., "Electron mobility in double gate silicon on insulator transistors: Symmetric-gate versus asymmetric-gate configuration," *J. Appl. Phys.*, vol. 94, p. 5732, 2003.

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

[30] D. Nika, E. Pokatilov, A. Askerov, and A. Balandin, "Phonon thermal conduction in graphene: Role of umklapp and edge roughness scattering," *Phys. Rev. B*, vol. 79, p. 155413, 2009.

[31] B. Jensen et al., "Effect of nanoscale heating on electrical transport in RF MEMS switch contacts," *IEEE J. Microelectromech. Sys.*, vol. 14, no. 5, pp. 935–945, 2005.

[32] K. Ritter and J. Lyding, "The influence of edge structure on the electronic properties of graphene quantum dots and nanoribbons," *Nature Materials*, vol. 8, p. 235, 2009.

[33] F. Elsholz et al., "Roughness evolution in thin-film growth of SiO and NbO," *J. Appl. Phys.*, vol. 98, p. 103516, 2005.

[34] F. Elsholz, E. Scholl, and A. Rosenfeld, "Control of surface roughness in amorphous thin-film growth," *Appl. Phys. Lett.*, vol. 84, p. 4186, 2004.

[35] R. Onions and J. Archard, "The contact of surfaces having a random structure," *J. Phys. D: Appl. Phys.*, vol. 6, p. 289, 1973.

[36] G. Adams and M. Nosonovsky, "Contact modeling–forces," *Tribol. Int.*, vol. 33, p. 431, 2000.

[37] C. Gardiner, *Handbook of Stochastic Methods*, 1st ed. New York, NY: Springer, 1985.

[38] P. Martin, Z. Aksamija, E. Pop, and U. Ravaioli, "Green function treatment of electronic decay in narrow nano-engineered rough semiconductor conduction channels," *J. Phys.: Conference Series*, vol. 193, p. 012009, 2009.

[39] P. Martin and U. Ravaioli, "Prediction of reduced thermal conductivity in nano-engineered rough semiconductor nanowires," *J. Phys.: Conference Series*, vol. 193, p. 012010, 2009.

[40] J. Bryzek, K. Petersen, and W. McCulley, "Micromachines on the march," *IEEE Spectrum*, vol. 31, no. 5, pp. 20–31, 1994.

[41] R. Coutu et al., "Selecting metal alloy electric contact materials for mems switches," *J. Micromech. Microeng.*, vol. 14, no. 8, p. 1157, 2004.

[42] S. Patton and J. Zabinski, "Failure mechanisms of capacitive MEMS RF switch contacts," *Tribol. Lett.*, vol. 19, no. 4, pp. 265–272, 2005.

[43] G. Rebeiz, *RF MEMS Design, Theory and Applications.* Hoboken, NJ: Wiley, 2003.

[44] J. Kruglick and S. Pister, "Lateral MEMS microcontact considerations," *IEEE J. Microelectromech. Sys.*, vol. 8, pp. 264–271, 1999.

[45] D. Hyman and M. Mehregany, "Contact physics of gold microcontacts for MEMS switches," *IEEE Tran. Comp. Pack. Tech.*, vol. 22, no. 3, 1999.

[46] S. Senturia, R. Harris, and B. Johnson, "A computer-aided design system for microelectromechanical systems (MEMCAD)," *J. MEMS*, vol. 1, pp. 3–13, 1992.

[47] M. Griebel, Oeltz, and M. Schweitzer, "An algebraic multigrid method for linear elasticity," *SIAM J. Sci. Comp.*, vol. 25, pp. 385–407, 2003.

[48] Y. Xiao, P. Zhang, and S. Shu, "An algebraic multigrid method with interpolation reproducing rigid body modes for semi-definite problems in two-dimensional linear elasticity journal of computational and applied mathematics," *J. Comp. App. Math.*, vol. 200, pp. 637–652, 2007.

[49] G. Li and N. Aluru, "Hybrid techniques for electrostatic analysis of nanoelectromechanical systems," *J. App. Phys.*, vol. 96, pp. 2221–2231, 2004.

[50] S. Patton and J. Zabinski, "Effects of dielectric charging on fundamental forces and reliability in capacitive microelectromechanical systems radio frequency switch contacts," *J. App. Phys.*, vol. 99, 2006.

[51] B. Jensen and K. Huang, "Low-force contact heating and softening using micromechanical switches in diffusive-ballistic electron-transport transition," *App. Phys. Lett.*, vol. 86, 2001.

[52] O. Bulashenko, O. Kochelap, and V. Kochelap, "Size effect on current fluctuations in thin metal-films - Monte Carlo approach," *Phys. Rev. B*, vol. 99, no. 4, 1992.

[53] J. Simmons and R. Verderber, "Coherent scattering of hot electrons in gold films," *Phys. Rev. Lett.*, vol. 17, no. 1274, pp. 665–667, 1966.

[54] R. Gould and C. Hogharth, "Forming, negative resistance, and dead-time effects in thin films of CaBr2," *Phys. Stat. Sol. (a)*, vol. 23, no. 531, pp. 531–535, 1974.

[55] R. Gould, "Relationship of the current-voltage characteristics to the distribution of filament resistances in electroformed mim structures," *Thin Sol. Films*, vol. 57, pp. 352–357, 1979.

[56] G. Dearnaley, D. Morgan, and A. Stoneham, "A model for filament growth and switching in amorphous oxide films," *J. Non Cryst. Sol.*, vol. 4, 1970.

[57] J. Rossignol, S. Clain, and M. Abbaoui, "The modelling of the cathode sheath of an electrical arc in vacuum," *J. Phys. D*, vol. 36, 2003.

[58] Y. Yurgelenas and H. Wagner, "A computational model of a barrier discharge in air at atmospheric pressure: the role of residual surface charges in microdischarge formation," *J. Phys. D*, vol. 39, 2006.

[59] S. Shimada, "Nano-processing with gas cluster ion beams," *J. Mat. Proc. Tech.*, vol. 144, 2004.

[60] D. Sornette and C. Vanneste, "Dynamics and memory effects in rupture of thermal fuse networks," *Phys. Rev. Lett.*, vol. 68, no. 5, 1992.

[61] M. Soderberg, "Resistive breakdown of inhomogeneous media," *Phys. Rev. B.*, vol. 35, 1987.

[62] C. Penetta and L. Reggiani, "A percolative approach to electromigration in metallic lines," *Comp. Mat. Sci.*, vol. 20, 2001.

[63] A. von Hippel, *Conduction and Breakdown of Gases.* New York, NY : Cambridge University Press, 1958.

[64] A. M. Howatson, *An Introduction to Gas Discharges*, 2nd ed. Pergamon Press, 1976.

[65] J. S. Townsend, *Electricity in Gases.* Clarendon Press, 1915.

[66] G. A. Farral, "Vacuum arcs and switching," *Proceedings of the IEEE*, vol. 61, no. 8, pp. 1113–1136, 1973.

[67] G. Grimmett, *Percolation*, 2nd ed. Springer, 1999.

[68] L. Hu, D. Hecht, and G. Gruner, "Percolation in transparent and conducting carbon nanotube networks," *Nano-Letters*, vol. 4, p. 2513, 2004.

[69] A. Yu et al., "Effect of single-walled carbon nanotube purity on the thermal conductivity of carbon nanotube-based composites," *App. Phys. Lett.*, vol. 89, no. 13, p. 2513, 2006.

[70] G. Pei and J. Kedzierski, "FinFET design considerations based on 3-d simulation and analytical modeling," *IEEE Trans. on Electron Devices*, vol. 49, pp. 1411–1419, 2002.

[71] G. Kathawala, B. Winstead, and U. Ravaioli, "Monte Carlo simulations of double-gate MOSFETs," *IEEE Trans. on Electron Devices*, vol. 50, pp. 2467–2473, 2003.

[72] B. Doyle and S. Datta, "High performance fully-depleted tri-gate CMOS transistors," *IEEE Electron Device Letters*, vol. 24, pp. 263–265, 2003.

[73] A. Sharma and S. Zaidi, "Mobility and transverse electric field effects in channel conduction of wrap-around-gate nanowire MOSFETs," *IEEE Proc.-Circuits Devices Syst.*, vol. 151, pp. 422–430, 2004.

[74] J. Inkson, *Many-body Theory of Solids: An Introduction.* Plenum Press, 1984.

[75] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev. B*, vol. 136, p. 864, 1964.

[76] G. Roach, *Green's Functions.* Cambridge University Press, 1982.

[77] W.-S. Choi and J.-K. Ahn, "A time dependent hydrodynamic device simulator SNU-2D with new discretization scheme and algorithm," *IEEE Trans. on CAD*, vol. 13, p. 898, 1994.

[78] A. Forghieri and R. Guerrieri, "A new discretization strategy of the semiconductor equations comprising momentum and energy balance," *IEEE Trans. on CAD*, vol. 7, p. 231, 1988.

[79] S. M. Sze, *Physics of Semiconductor Devices.* J. Wiley and Sons, 1981.

[80] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductor with application to covalent materials," *Rev. Mod. Phys.*, vol. 55, pp. 645–705, 1983.

[81] A. Duncan, U. Ravaioli, and J. Jakumeit, "Full-band Monte Carlo investigation of hot carrier trends in the scaling of MOSFETs," *IEEE Trans. on Electron Devices*, vol. 45, pp. 90–96, 1996.

[82] K. Hess, *Monte Carlo Device Simulation: Full Band and Beyond.* Kluwer Academic Publishers, 1991.

[83] W. McMahon, A. Haggag, and K. Hess, "Reliability scaling issues for nanoscale devices," *IEEE Trans. on Nanotechnology*, vol. 2, pp. 33–38, 2003.

[84] J. Jakumeit, U. Ravaioli, and K. Hess, "Calculation of hot electron distributions in silicon by means of an evolutionary algorithm," *J. of Applied Physics*, vol. 80, pp. 5061–5066, 1996.

[85] B. Winstead, "Monte Carlo simulation of silicon devices including quantum correction and strain," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2001.

[86] K. Hess, *Advanced theory of semiconductor devices.* New York, NY: IEEE Press, 2000.

[87] A. Pacelli, "Transport phenomena and hot electron effects in scaled MOSFET devices," Ph.D. dissertation, Polytechnic Institute of Milan, 1997.

[88] U. Ravaioli, *MOCA user's guide, MOCA: a 2D full-band Monte Carlo device simulator*, 2002.

[89] P. Martin, "TCAD of confined silicon nanochannels based on a 3d quantum corrected monte carlo approach," M.S. thesis, University of Illinois at Urbana-Champaign, 2007.

[90] A. Trellakis, "Computational approaches to silicon based nanostructures," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2000.

[91] R. Hockney and J. Eastwood, *Computer Simulations Using Particles*. McGraw Hill, 1981.

[92] A. Trellakis and U. Ravaioli, "Lateral scalability limits of silicon conduction channels," *J. Appl. Phys.*, vol. 86, p. 3911, 1999.

[93] A. Sharma, S. Zaidi, S. Lucero, and S. Beck, "Mobility and transverse electric field effects in channel conduction of wrap-around-gate nanowire MOSFETs," *IEEE Proc. Circuits Dev. Syst.*, vol. 151, pp. 422–430, 2004.

[94] S. Yamakawa, H. Ueno, K. Taniguchi, and C. Hamaguchi, "Study of interface roughness dependence of electron mobility in Si inversion layers using the monte carlo method," *J. Appl. Phys.*, vol. 79, pp. 911–916, 1996.

[95] I. Mudawar, "Assessment of high-heat-flux thermal management schemes," *IEEE Trans. on Comp. and Pack. Tech.*, vol. 24, pp. 122–141, 2001.

[96] P. Klemens, "Thermal conductivity and lattice vibrational modes," *Solid State Physics*, vol. 7, pp. 1–98, 1957.

[97] I. Chowdhury et al., "On-chip cooling by superlattice-based thin-film thermoelectrics," *Nature Nanotech.*, vol. 4, p. 235, 2009.

[98] A. Majumdar, "Thermoelectric devices helping chips to keep their cool," *Nature Nanotech.*, vol. 4, pp. 214–215, 2009.

[99] R. Mahajan, C. Chiu, and G. Chrysler, "Cooling a microprocessor chip," *Proc. IEEE*, vol. 94, p. 1476, 2006.

[100] R. Prasher et al., "Nano and micro technology-based next-generation package-level cooling solutions," *Intel Tech. J.*, vol. 94, p. 285, 2005.

[101] K. Hsu et al., "Cubic AgPbSbTe2: Bulk thermoelectric materials with high figure of merit," *Science*, vol. 303, p. 818, 2004.

[102] T. Harman et al., "Quantum dot superlattice thermoelectric materials and devices," *Science*, vol. 297, p. 2229, 2002.

[103] D. Rowe, *Thermoelectrics Handbook: Macro to Nano.* CRC Press, Florida, 2006.

[104] W. Hui and J. Corra, "Seebeck coefficient of thin-film germanium," *J. Appl. Phys.*, vol. 39, p. 3477, 1967.

[105] W. Fulkerson et al., "Thermal conductivity electrical resistivity and Seebeck coefficient of silicon from 100 to 1300 degree K," *Phys. Rev.*, vol. 167, p. 765, 1968.

[106] O. Madelung, U. Rossler, and M. Schultz, *Group IV Elements, IV-IV and II-V Compounds. Part b - Electronic, Transport, Optical and Other Properties.* Springer, 2002.

[107] N. Mingo, "Thermoelectric figure of merit and maximum power factor in III-V semiconductor nanowires," *Appl. Phys. Lett.*, vol. 84, p. 2653, 2004.

[108] D. Broido and T. Reinecke, "Theory of thermoelectric power factor in quantum well and quantum wire superlattices," *Phys. Rev. B*, vol. 64, p. 045324, 2001.

[109] J. Ziman, *Electrons and phonons: the theory of transport phenomena in solids*, 1st ed. Clarendon, 1960.

[110] C. Kittel, *Introduction to solid state physics*, 8th ed. Wiley, 2005.

[111] G. Gilat and L. Raubenheimer, "Accurate numerical method for calculating frequency-distribution functions in solids," *Phys. Rev.*, vol. 144, no. 2, pp. 390–395, 1966.

[112] W. Weber, "Adiabatic bond charge model for the phonons in diamond, Si, Ge, and Sn," *Phys. Rev. B*, vol. 15, no. 10, pp. 4789–4803, 1977.

[113] Z. Aksamija and U. Ravaioli, "Joule heating and phonon transport in nanoscale silicon mosfets," *Electro/Information Tech.*, vol. 1, pp. 70–72, 2007.

[114] M. Lenzi et al., "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. Dev.*, vol. 55, no. 8, pp. 2086–2096, 2008.

[115] C. Glassbrener and A. Slack, "Thermal conductivity of silicon and germanium from 3k to the melting point," *Phys. Rev.*, vol. 134, pp. 1058–1069, 1964.

[116] N. Mingo, "Calculation of Si nanowire thermal conductivity using complete phonon dispersion relations," *Phys. Rev. B*, vol. 68, pp. 168–171, 2003.

[117] M. Holland, "Analysis of lattice thermal conductivity," *Phys. Rev.*, vol. 132, no. 6, pp. 2461–2471, 1963.

[118] E. Pop, R. Dutton, and K. Goodson, "Analytic band monte carlo model for electron transport in si including acoustic and optical phonon dispersion," *J. App. Phys.*, vol. 96, no. 9, pp. 4998–5005, 2004.

[119] J. Lagarias, J. Reeds, M. Wright, and P. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[120] E. Ramayya et al., "Cross-sectional dependence of electron mobility and lattice thermal conductivity in silicon nanowires," *8th IEEE Conference on Nanotechnology*, vol. 1, p. 339, 2008.

[121] M. Asen-Palmer et al., "Thermal conductivity of germanium crystals with different isotopic compositions," *Phys. Rev. B*, vol. 56, p. 9431, 1997.

[122] S. Baraman and G. Srivastava, "Quantitative estimate of phonon scattering rates in different forms of diamond," *Phys. Rev. B*, vol. 73, p. 205308, 2006.

[123] W. Fon, K. Schwab, J. Worlock, and M. Roukes, "Phonon scattering mechanisms in suspended nanostructures from 4 to 40 K," *Phys. Rev. B*, vol. 66, p. 045302, 2002.

[124] S. Lee, D. Cahill, and H. Allen, "Thermal-conductivity of sputtered oxide-films," *Phys. Rev. B*, vol. 52, p. 253, 1995.

[125] D. Strauch and B. Dorner, "Phonon dispersion in GaAs," *J. Phys. Cond. Mat.*, vol. 2, p. 1457, 1990.

[126] G. Nilsson and G. Nelin, "Phonon dispersion relations in Ge at 80 K," *Phys. Rev. B*, vol. 6, p. 3777, 1972.

[127] M. Holland, "Analysis of lattice thermal conductivity," *Phys. Rev*, vol. 134, p. 471, 1964.

[128] K. Novoselov et al., "Two-dimensional gas of massless Dirac fermions in graphene," *Nature Letters*, vol. 438, p. 197, 2005.

[129] E. Hwang, S. Adam, and S. D. Sarma, "Carrier transport in two-dimensional graphene layers," *Phys. Rev. Lett.*, vol. 98, p. 186806, 2007.

[130] J. Chen et al., "Intrinsic and extrinsic performance limits of graphene devices on SiO2," *Nature Nano.*, vol. 3, p. 206, 2008.

[131] T. Yamamoto, T. Watanabe, and T. Mii, "Empirical-potential study of phonon transport in graphitic ribbons," *Phys. Rev. B.*, vol. 70, p. 245402, 2004.

[132] K. Saito, J. Nakamura, and A. Natori, "Ballistic thermal conductance of a graphene sheet," *Phys. Rev. B*, vol. 76, p. 115409, 2007.

[133] G. Srivastava, *The Physics of Phonons.* IOP, Philadelphia, 1990.

[134] Z. Aksamija and U. Ravaioli, "Anharmonic decay of g-process longitudinal optical phonons in silicon," *J. Appl. Phys.*, vol. 105, p. 083722, 2009.

[135] H. Hahm, "Enhancement of Monte Carlo simulations on 3d nanoscale semiconductor devices," Ph.D. dissertation, University of Illinois, Urbana-Champaign, 2008.

[136] F. Sols, M. Maccuci, U. Ravaioli, and K. Hess, "On the possibility of transistor action based on quantum interference phenomena," *App. Phys. Lett.*, vol. 54, pp. 350–353, 1989.

[137] F. Sols, M. Maccuci, U. Ravaioli, and K. Hess, "Theory for a quantum modulated transistor," *J. Appl. Phys.*, vol. 66, pp. 3892–3905, 1989.

[138] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM J. Res. Develop.*, vol. 1, p. 223, 1957.

[139] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM J. Res. Develop.*, vol. 32, p. 306, 1988.

[140] Y. Meir and S. Wingreen, "Landauer formula for the current through an interacting electron region," *Phys. Rev. Lett.*, vol. 72, p. 2512, 1992.

[141] P. Caroli, R. Combescott, and D. Saint-James, "Direct calculation of tunneling current," *J. Phys. C*, vol. 4, p. 916, 1971.

[142] D. Thouless and S. Kirkpatrick, "Conductivity of the disordered linear-chain," *J. Phys. C*, vol. 14, p. 235, 1981.

[143] P. Yang et al., "Thermal transport in a dielectric t-shaped quantum wire," *Phys. Rev. B*, vol. 75, p. 235, 2007.

[144] U. Ekenberg, "Nonparabolicity effects in a quantum well: sublevel shift, parallel mass, and Landau levels," *Phys. Rev. B*, vol. 40, p. 1626, 1989.

[145] A. Persson and R. Cohen, "Reformulated Hamiltonian for nonparabolic bands in semiconductor quantum wells," *Phys. Rev. B*, vol. 38, p. 5568, 1988.

[146] E. Conwell and M. Vassel, "High-field transport in n- type GaAs," *Phys. Rev.*, vol. 166, p. 797, 1968.

[147] G. Bastard, "Theoretical investigations of superlattice band structure in the envelope-function approximation," *Phys. Rev. B*, vol. 25, p. 7584, 1982.

[148] J. Lopez-Villanueva et al., "Modified Schrödinger equation including nonparabolicity for the study of a two-dimensional electron gas," *Phys. Rev. B*, vol. 48, p. 1626, 1993.

[149] K. Aihara, M. Yamamoto, and T. Mizutani, "3-terminal conductance modulation of a quantum interference device using a quantum-wire with a stub structure," *Appl. Phys. Lett.*, vol. 63, p. 3595, 1993.

[150] J. Wang et al., "Theoretical investigation of surface roughness scattering in silicon nanowire transistors," *Appl. Phys. Lett.*, vol. 87, p. 043101, 2005.

[151] R. Lehoucq, D. Sorensen, and C. Yang, *ARPACK User's Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods.* Philadelphia, PA: SIAM, 1998.

[152] F. Ferrari et al., "Raman spectrum of graphene and graphene layers," *Phys. Rev. Lett.*, vol. 97, p. 187401, 2006.

[153] K. Novoselov et al., "Electric field effect in atomically thin carbon films," *Science*, vol. 306, pp. 666–669, 2004.

[154] M. Bae, Z. Ong, D. Estrada, and E. Pop, "Imaging, simulation, and electrostatic control of power dissipation in graphene devices," *Nano Lett.*, 2010.

[155] D. Abdula et al., "Environment-induced effects on the temperature dependence of Raman spectra of single-layer graphene," *J. Phys. Chem. C*, vol. 112, p. 20131, 2008.