Doctoral Thesis

Scanning probe techniques for dopant profile characterization

Author(s):
Stangoni, Maria Virginia

Publication Date:
2005

Permanent Link:
https://doi.org/10.3929/ethz-a-005061495

Rights / License:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Scanning Probe Techniques for Dopant Profile Characterization

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Technical Sciences

presented by
MARIA VIRGINIA STANGONI
Laurea di Dott. in Ingegneria
University of Cagliari
born March 4\textsuperscript{th} 1972
citizen of Italy

accepted on the recommendation of
Prof. Dr. Wolfgang Fichtner, examiner
Dr. Vito Raineri, co-examiner

2005
Abstract

In the present work, the imaging and the quantitative doping profiling capabilities of Scanning Capacitance Microscopy (SCM) and Scanning Spreading Resistance Microscopy (SSRM) for nanometer scale devices are investigated.

Special attention has been paid to the development and the optimization of dedicated processes for the preparation of suitable cross-sectioned samples. In fact, the control of surface characteristics such as roughness, state density, and fixed charges has been demonstrated to be a crucial factor for reproducible and quantitative results.

The exploration of the physical limits of the SCM technique has been assisted by two- and three-dimensional simulations of the measurement process. In particular, device simulation has been applied to quantify the maximum achievable accuracy in the determination of doping profiles and in the delineation of the electrical junction. In this respect, a novel technique is proposed, which represents an efficient alternative to the capacitance spectroscopy for the quantitative localization of the electrical junction in bipolar samples. The capabilities of SCM have been assessed with own measurements and with experimental data from a round robin experiment, which has involved several European laboratories.

SSRM measurements have been carried out on the same samples characterized by SCM and the peculiarities of both techniques are discussed. In particular, this work points out the impact of the different parasitic components that contribute to the formation of the SSRM signal, leading to the observed deviation of the measured characteristics from the curves predicted by the spreading theory. It is also shown that for this reason the quantitative doping profiling by
SSRM requires an accurate case-by-case calibration. The SSRM measurement process has been simulated in two and in three dimensions both to quantify the weight of the parasitic series resistances and to ascertain the intrinsic limits of the technique in the delineation of the electrical junction in bipolar samples. These simulations demonstrate that due to the strong modification of the local carrier density through the injecting tip, SSRM cannot reach the same junction delineation capabilities as for SCM.

Progress has also been made in increasing the accuracy of the conversion of the free carrier SCM and SSRM profiles into the related doping profiles. For this purpose, a numerical solution scheme of the reverse modelling problem based on the use of Artificial Neural Networks has been developed.

The assessment of both techniques demonstrates that SCM and SSRM exhibit a large degree of complementarity. SCM has shown to be superior to SSRM in the case of complex large bipolar structures, whereas SSRM provides the excellent lateral resolution required to image nanometer scale devices.
Riassunto

Questo saggio tratta delle potenzialità di visualizzazione bidimensionale e di profilometria di drogaggio della Scanning Capacitance Microscopy (SCM) e della Scanning Spreading Resistance Microscopy (SSRM) applicate a dispositivi a semiconduttore di scala nanometrica. Poiché per l’ottenimento di risultati riproducibili e quantitativi si sono dimostrati cruciali aspetti quali la rugosità, la densità degli stati di superficie e delle cariche fisse, particolare attenzione è stata dedicata all’ottimizzazione dei processi di preparazione dei campioni di entrambe le tecniche.

L’esplorazione sperimentale dei limiti fisici delle due metodiche è stata assistita da dati ottenuti mediante simulazione dei processi di misurazione in due e tre dimensioni con un avanzato simulatore di dispositivi. Questo ha consentito di determinare la massima risoluzione raggiungibile nella determinazione di profili di drogaggio e nella delineazione di giunzioni elettriche. A questo riguardo è stata sviluppata una nuova metodologia di localizzazione quantitativa della giunzione elettrica in campioni bipolari, che rappresenta una valida ed efficiente alternativa alle tecniche di spettroscopia di capacità.

Le prestazioni in termini di microscopia della SCM sono state accertate mediante il confronto di dati sperimentali propri e di misurazioni acquisite nel quadro di un progetto internazionale che ha coinvolto diversi laboratori europei.

Le peculiarità della SSRM sono state caratterizzate con gli stessi campioni utilizzati per la SCM, in modo da poter effettuare un confronto diretto. Queste misurazioni hanno evidenziato l’importanza delle diverse componenti parassite nella formazione del segnale SSRM, ciò che costituisce la causa principale della notevole discrepanza ri-
levata fra le caratteristiche sperimentali e quelle predette dalla teoria dello spreading di corrente. È stato inoltre dimostrato che la quantificazione dei profili di drogaggio acquisiti mediante SSRM richiede un’accurata fase di calibrazione preliminare da effettuarsi caso per caso.

Il processo di misura della SSRM è stato simulato in due e tre dimensioni. Questo ha permesso di valutare il peso relativo delle resistenze serie parassite, consentendo così di accertare il limite intrinseco della tecnica nella delineazione della giunzione elettrica in campioni bipolari. Le simulazioni dimostrano inoltre che, a causa della forte distorsione introdotta nella densità dei portatori dall’iniezione di corrente attraverso la punta, la SSRM non può raggiungere le stesse capacità di delineazione della giunzione elettrica appurate per la SCM.

Lo sviluppo di uno schema risoluzione numerica del problema della modellizzazione inversa basato sull’utilizzo di reti neurali artificiali ha consentito di migliorare la precisione nella conversione del profilo dei portatori liberi nel corrispettivo profilo di drogaggio. Questa nuova tecnica può essere usata sia per la SCM che per la SSRM.

Lo studio delle prestazioni in termini di microscopia della SCM e della SSRM ha confermato che entrambe le tecniche presentano caratteristiche complementari. La SCM si è dimostrata superiore alla SSRM nel caso di strutture bipolari estese e complesse, mentre la SSRM essendo capace di una risoluzione laterale eccellente, si è rivelata una tecnica più adatta alla visualizzazione di dispositivi su scala nanometrica.
# Contents

Abstract i  
Riassunto iii

1 Introduction 1
   1.1 Traditional 1D characterization techniques . . . . . . . 3
   1.2 AFM-Atomic Force Microscopy . . . . . . . . . . . . . . . 5
   1.3 SCM and SSRM simulation . . . . . . . . . . . . . . . . . 6
   1.4 Outline of this work . . . . . . . . . . . . . . . . . . . . . 7

2 Experimental details 11
   2.1 Scanning Capacitance Microscopy . . . . . . . . . . . . . 11
      2.1.1 1D MOS capacitor . . . . . . . . . . . . . . . . . 11
      2.1.2 SCM operating principle . . . . . . . . . . . . . . . 14
   2.2 Scanning Spreading Resistance Microscopy . . . . . . . 18
      2.2.1 SSRM theory of operation . . . . . . . . . . . . . . 20
      2.2.2 Probe contact and resolution of SSRM . . . . . . . 22
      2.2.3 Electro-mechanical model . . . . . . . . . . . . . . . 24
   2.3 Probes for SCM and SSRM . . . . . . . . . . . . . . . . . 26
   2.4 Features of scanning microscopes . . . . . . . . . . . . . . 28
      2.4.1 Concentration resolution . . . . . . . . . . . . . . . 28
      2.4.2 Spatial resolution . . . . . . . . . . . . . . . . . . 28
      2.4.3 Spatial accuracy (depth resolution) . . . . . . . . . 30
      2.4.4 Gradient resolution (sensitivity) . . . . . . . . . . . 31
   2.5 Sample preparation for silicon cross-sections . . . . . . . 32
   2.6 Measurements on beveled samples . . . . . . . . . . . . . 36
CONTENTS

4.5.4 2D simulation of SSRM on pn junctions . . . . . . 127
4.5.5 SSRM 3D simulations . . . . . . . . . . . . . . . . 137
4.5.6 3D SSRM simulation of a doping step . . . . . . 138
4.5.7 3D SSRM simulation of p-staircase . . . . . . 139
4.5.8 3D simulation of SSRM on a pn junction . . . 141
4.6 SSRM imaging of a bipolar transistor . . . . . . 143
4.7 SSRM imaging of DMOS and MOS transistors . . . 145
4.8 Summary . . . . . . . . . . . . . . . . . . . . . . . . . 148

5 Conclusions
5.1 Dopant profiling application of SCM and SSRM . . . 150
5.2 Electrical junction delineation . . . . . . . . . . . . . 151
5.3 Main differences between SCM and SSRM . . . . . . 152
5.3.1 Probes . . . . . . . . . . . . . . . . . . . . . . . . . . . 153
5.4 SCM and SSRM simulation . . . . . . . . . . . . . . 153
5.5 Microscopy application of SCM and SSRM . . . . . 155
5.6 Quantification of SCM and SSRM . . . . . . . . . . . 156
5.7 Outlook . . . . . . . . . . . . . . . . . . . . . . . . . . 156

Bibliography 159
Acknowledgements 169
Curriculum Vitae 171
Chapter 1

Introduction

The scaling of electronic components requires proper tools for the characterization and measurement of very small semiconductor devices. The International Technology Roadmap of Semiconductors (ITRS) [1] projects a fifteen-year horizon of the developments for future technology generations [2]. The metrology roadmap summarizes the requirements for physical analysis and characterization. Well-matched metrology and process tools reduce the ramping times for pilot lines and factories. An appropriate combination of process and appropriate metrology tools is necessary to maximize productivity while maintaining an acceptable cost of ownership. Better characterization processes can reduce the cost of manufacturing and time-to-market for new products.

Physical characterization requirements include two/three dimensional (2D and 3D) defect and dopant profiling, imaging for very small feature sizes and fast imaging methods for measurements of features across wafers. Microscopy is used in most of the core technology processes where two-dimensional distributions, i.e., images of the shape and structure of integrated circuit, reveal important information. Most of the available techniques are of easy application to microscopy, while is still a challenging issue their use for dopant profiling and junction delineation.

The electron beam techniques operate in vacuum. They include: Scanning Electron Microscopy (SEM), EBIC (Electron Beam Induced
Introduction

Current), and Transmission Electron Microscopy (TEM). Especially for TEM, a thorough preparation of the sample is required. Application of SEM for doping profiling of silicon is feasible [3], although the poor contrast and several of the drawbacks (e.g. contrast reversal, no discrimination between p and n type, need of calibration, critical surface condition) make the use of this method problematic [4]. Appreciable measurements have recently been made for semiconductor material like silicon carbide [5]. The resolution possible with this method is not very high as indicated in Table 1.1. Application of TEM to silicon dopant profiling [6], results in a very high resolution (few nm) with poor contrast for doping concentrations below $10^{18}$ cm$^{-3}$. Problems include specimen charging and excess carrier generation by illumination, existence of an amorphous layer at the sample surface reducing the contrast, and the contribution to the diffraction caused by mechanical stress and crystal defects in the sample.

Simplicity and high performance of Scanning Probe Microscopes make them excellent metrology tools. In recent years, SPM metrology capabilities have improved significantly in resolution and probe fabrication, data processing, and detection sensitivity. Several two-dimensional dopant profiling and microscopy techniques have been developed on the basis of scanning probe microscopy-based methods:

- Kelvin Force Probe Microscopy (KFPM) measures the contact potential difference, which is the difference in work function between tip and sample. Application in high vacuum improves the lateral resolution to atomic scale [7] [8]. The technique is not useful as dopant profiler.

- Dopant selective etching followed by Atomic Force Microscopy (AFM) topographic analysis is in excellent agreement with traditional techniques, but needs an optimized etching process and calibration [9] [10] [11].

- Scanning Tunneling Microscopy (STM) with spectroscopy (STS) measures the tunneling current between tip and sample as a function of the applied voltage, revealing the profile of a pn junction with a spatial resolution on the order of 10 nm [12].
1.1 Traditional 1D characterization techniques

Tunneling techniques always require a current to flow, and no oxide is supposed to grow on the sample.

- Scanning Capacitance Microscopy (SCM) [13] [14] [15] [16] [17] [18], and Scanning Spreading Resistance Microscopy (SSRM) [19] [20] [21] [22] are the most promising techniques and are treated in detail in this work.

The main advantages of scanning probe microscopy are the high spatial resolution of cross sectioned samples and the use in a non-vacuum environment. Table 1.1 summarizes the major characteristics of the most common used techniques.

The increasing importance of doping profiling for process control finds another field of application in the calibration of process modeling. Since the electrical properties of semiconductors are associated with the doping distribution in the lattice, the calibration of advanced device modeling requires experimental profiles and electrical characteristics [23]. Hence, spatial resolution becomes an important feature of profiling tools, and it has been continuously improved in SPM techniques, with a minimum practical limit set by the atomic spacing.

<table>
<thead>
<tr>
<th>Method</th>
<th>Spatial res.</th>
<th>Dyn. range [cm$^{-3}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEM</td>
<td>5-10 nm</td>
<td>$&gt;10^{16}$</td>
</tr>
<tr>
<td>Electron holography</td>
<td>Debye length-limited</td>
<td>$&gt;10^{18}$</td>
</tr>
<tr>
<td>SCM</td>
<td>Debye length-limited</td>
<td>$10^{15}$-$10^{20}$</td>
</tr>
<tr>
<td>SSRM</td>
<td>about 10× tip radius</td>
<td>$10^{15}$-$10^{20}$</td>
</tr>
</tbody>
</table>

*Table 1.1: Comparison of two-dimensional profiling methods.*

1.1 Traditional 1D characterization techniques

In the past, traditional characterization techniques have been used for dopant profiling. Secondary Ion Mass Spectrometry (SIMS) and Spreading Resistance Profiling (SRP) only provide 1D profiling of the dopants or the charge carriers in semiconductors and are not adequate
for two-dimensional purposes. SIMS is a vacuum-based technique, which relies on the removal of material from a solid specimen by ion sputtering and on the analysis of the secondary ionized species [24]. Only ionized atoms can be analyzed by passing through an energy filter and a mass spectrometer. SIMS has good detection sensitivity for many elements, though not as high as electrical or optical methods. Among the beam techniques, it has the highest sensitivity, detecting dopant densities as low as $10^{15}$ cm$^{-3}$. It allows simultaneous detection of different elements at a depth resolution of 1 to 5 nm, and it can give lateral surface characterization on a scale of several microns. It is a destructive method since the very act of removing material by sputtering leaves a crater in the sample. The secondary ion signal is converted to impurity density through standards of known dopant profile. SIMS determines the chemical impurity density, not the electrically active impurity density. For example, implanted, non-annealed samples give SIMS profiles very close to the theoretical profile. Electrical measurements give different results, with the ions not yet electrically activated. SIMS and electrical measurements agree quite well for activated samples.

Comparison of SIMS dopant distributions with profiles measured with Spreading Resistance sometimes show a discrepancy in the lowly doped portion of the profile, giving deeper junction position. This discrepancy is not fully understood, being likely caused by cascade mixing and knock on of dopant atoms by the sputtering beam, contributing to slightly deeper junctions, or by the limited dynamic range of the SIMS instrument. When sputtering from a high-doped region near the surface to a low-doped region deeper in the sample, the crater walls contain the entire doping information. Any stray signal from the crater walls deposited in the crater floor is added to the crater floor signal giving the appearance of a higher dopant density and hence a deeper profile.

Spreading resistance profiling has been in use since the 1960s, when it was introduced for the first time by Mazur and Dickey [25]. It is mainly used to generate resistivity and carrier density depth profiles. The technique, originally used for junctions tens of microns deep, is today capable of profiling very shallow junctions.

The SRP probe consists of two aligned probes that are stepped along the semiconductor beveled surface, the forces applied are on the
order of 50000 μN and the probe penetration is around 10 nm [26]. The resistance between the probes is measured at each location. The probe size and bevel angle set the spatial resolution of this technique on the order of several tens of nanometers.

Spreading resistance measures a carrier distribution along a beveled surface [see chapter 2]. The redistribution of carriers distorts their profile, therefore they account for SRP-determined junction depths that are usually shallower than those measured by SIMS. Spreading resistance profiling is a comparative technique. Calibration curves are generated for a particular set of probes using samples of known resistivity. Comparison of the spreading resistance data with the calibration profiles is necessary and sufficient for uniformly doped samples. For samples containing \( pn \) or high-low junctions, additional corrections are necessary [24].

1.2 AFM-Atomic Force Microscopy

The Atomic Force Microscope (AFM) belongs to the family of scanning probe microscopes. It makes use of a very sharp probe to detect superficial properties of the materials on nanometer and sub-nanometer scale. The AFM probe is located at the free end of a cantilever that is attached to a piezo scanner able to move in vertical direction, and in a restricted (maximum size 60×60 μm\(^2\)) horizontal plane with a sub-nanometer precision.

In particular, the AFM is commonly applied for topography characterization in three operating modes: non-contact, tapping, and contact AFM. Changes in topography are detected in several manners. The most common way is the use of a laser beam pointing at the reflecting cantilever rear, and being bounced off to a split photo detector. While scanning the tip on the surface any change in the topography is accommodated by the probe via an appropriate feedback system. All the cantilever deflections cause the laser beam to hit the photo detector at different spots, which corresponds to different topography coordinates, and which are mapped and stored by a dedicated software. The lateral resolution of an AFM topography profiler is on the order of nanometers. Figure 1.1 shows the main parts of the apparatus. When operating in contact mode, the AFM can be coupled to a
particular sensor, and, with the appropriate cantilever, it can be used also to characterize electrically the area of interest.

Scanning Capacitance Microscopy (SCM) and Scanning Spreading Resistance Microscopy (SSRM) are two techniques based on AFM that measure variation of capacitance and resistivity, respectively. They were developed as 1D and 2D dopant profiling tools for semiconductors. Early work was made by Williams on SCM proving the detection of dopant densities with measurements of local capacitance by a scanning tip [27], Shafai for a Scanning Resistance system able to detect resistance variation [28], and Vandervorst with a first illustration of the capabilities of the formerly-called nano-Spreading Resistance Profiler [29]. The present work aims to define the state of the art of SCM and SSRM, with a quantitative point of view whenever possible.

1.3 SCM and SSRM simulation

Numerical simulation was employed to model the measurement system in order to explain and clarify the microscope output data. The software simulation package used is the ISE [30] TCAD software, which includes a large range of tools. Among them, boundary and grid generation (MESH and DEVISE) and device simulation (DESSIS) have
been used for our purposes. The necessity to run parametrical sets of simulation required the use of a tool like GENESISE, which is a graphical user interface that integrates into one environment all the simulation programs in order to design, parameterize, set up, and execute full clusters of files.

Two-dimensional simulation of SCM and SSRM is, so far, the easiest way to model the systems. For SCM, a physics-dependent grid is the heart of the simulation model, and the use of a boundary conforming mesh is mandatory. The three-dimensional approach is problematic mostly because of the large number of mesh nodes that produce very long CPU simulation times. The scanning tip modeled as a contact and the spatially parameterized doping profile can be used to emulate the scanning movement of the probe. For large doping gradients and large depleted $pn$ junctions the necessity of a wide set of simulation makes the time problem a crucial point.

For quantification of SSRM a three-dimensional simulation model must be also set up. The absence of depletion regions in this case allows to set a smaller refinement box or shell reducing the overall number of nodes. Problems arise with the definition of the tip contact, ohmic or Schottky boundary conditions might affect in a different way the conduction mechanisms and produce not the same results. Attention must be paid to this problem. Full three-dimensional simulation with boundary conforming offset is feasible with DEVISE, a boundary generation tool, and NOFFSET3D, which generates in 3D a region or boundary conforming mesh grid.

1.4 Outline of this work

In spite of all the really significant latest improvements of SCM and SSRM, reliability and quantification of measurements are not automatically obtained.

Many aspects of the measurement must be controlled: accurate sample preparation, tip shape and wear out, quality of the contacts, and calibration. The probe/sample interaction is the key point of these techniques, but also the primary source of error for all probing microscopes. All features are discussed in this work with a quantitative intent, hence, a detailed modeling of the techniques is carried
out, leading to the interpretation of the measured data. Particular emphasis is given to calibration.

The Atomic Force Microscope used for the measurements is a DI 3100 [31], equipped only in a second time with a QUADREX extension and new sensors for SCM and SSRM.

This work is divided in three main chapters and conclusions.

Chapter 2 contains a description of the two techniques under study. Scanning Capacitance and Scanning Spreading Resistance are explained from a theoretical point of view: the fundamental principles, the electrical equipment for detection and data interpretation, and the electrical model representing the system. A paragraph is devoted to the resolution of SCM and SSRM, its definition and its quantification. Special attention is paid to the description of the manual sample preparation for cross-sections, followed by a short description of the beveling method and its artifacts. The samples used for the SCM and SSRM measurements are described.

Chapter 3 contains a detailed analysis of Scanning Capacitance Microscopy, measurements on test samples, simulation of the technique and application to the evaluation and understanding of the measured data. A full quantitative analysis of electrical junction delineation is described, resulting into a quantitative application of this method. The problem of contrast reversal is presented and solved basing on physical considerations.

Chapter 4 discusses the SSRM application to 1D doping profiling and to the junction delineation. Calibration of a standard sample and a comparison with data measured in independent laboratories are presented. The quantification of the resulting data is investigated, with a discussion on the parameters that should be accounted for in order to explain the experimental results. This includes series resistances, which change the linear theoretical rate of the calibration curve. A wide set of simulations is presented: 2D calculations reveal the strong geometry-dependent character of SSRM. Three-dimensional simulations are made to compare the theoretical resistance results with the measurements. Several problems in setting up the simulations are also discussed.
1.4 Outline of this work

Finally, the results are summarized and compared in Chapter 5, which contains a survey of the results with an outlook to possible future activities.
Chapter 2

Experimental details

This chapter contains a description of the operating principles of Scanning Capacitance Microscopy (SCM) and Scanning Spreading Resistance Microscopy (SSRM). Both the techniques require probes with particular shapes and materials, together with a careful preparation of the sample surface. The sample cross-sectioning procedure is described in detail, followed by the illustration of the advantages of beveled samples. Finally, the problem of the quantification of the measured profiles is presented. Using an analytical solution of the reverse modeling problem and a novel approach based on Artificial Neural Networks (ANN).

2.1 Scanning Capacitance Microscopy

The SCM relies on the detection of the local free carrier concentration by measuring the local depletion capacitance that is formed below a conducting scanning probe. The system probe, thin oxide, and silicon can be interpreted as a microscopic MOS capacitor.

2.1.1 1D MOS capacitor

A one-dimensional MOS capacitor consists of a layer of oxide separating one metallic electrode called gate and one semiconductor electrode
Experimental details

(Figure 2.1). The equivalent circuit is simply composed of two capacitances: a capacitance in accumulation defined by the oxide thickness, $C_{ox}$, and a doping dependent capacitance in depletion, $C_d$. The capacitance-voltage curves, obtained by ramping the gate voltage and measuring the capacitance $C = \Delta Q / \Delta V$, depend on the sweep rate, and thus on the frequency of the applied voltage and the minority carrier lifetime $\tau$ in the semiconductor. For a $p$-doped semiconductor, the low-frequency and high-frequency CV curves are plotted in Figure 2.2. The low-frequency response is characterized by a voltage ramping time larger than the minority carriers lifetime $\tau$. On application of a strongly negative gate voltage, the holes accumulate below the oxide exceeding the concentration in the bulk. The MOS capacitance is $C_{ox}$. As the voltage is ramped toward less negative values, the majority carriers accumulating below the oxide are reduced and the total capacitance decreases till the voltage reaches the flat band value. Increasing the voltage the semiconductor depletes under the gate. The total MOS capacitance is now the series capacitance of $C_{ox}$ and $C_d$, till the voltage becomes large enough for an inversion layer to form under the oxide. When the concentration of minority carriers is comparable to the concentration in the bulk the differential capacitance is equal again to $C_{ox}$.

At high frequency, when the sweep rate is comparable to $\tau$, the minority carriers cannot respond fast enough. Once the MOS is in-

![Figure 2.1: 1D MOS capacitor and the simplified equivalent circuit.](image)
2.1 Scanning Capacitance Microscopy

![Figure 2.2: Capacitance vs. voltage curves for low frequency voltage ramp (dotted), very high frequency voltage ramp (dashed), AC small signal detection (solid).](image)

verted below the oxide, the minority carriers cannot follow voltage changes and they do not contribute to the total capacitance. The total capacitance saturates at $C_d$.

At very high frequency, when the voltage ramp rate is much faster than $\tau$, the minority carriers have no time to reach the oxide surface, thus they do not contribute at all to the total capacitance. The capacitance $C_d$, decreases towards deep depletion [32].

In SCM the capacitance detection system uses three different voltage signals:

- a $V_{DC}$ bias
- a $V_{AC}$ bias in the kHz range
- a $V_{AC}$ bias in the UHF range (about 1 GHz).

The DC bias is applied to the tip/sample in order to set the system working point, and the AC voltage executes the fast sweep in the kHz range. The frequency of the AC signal is high enough to obtain a
Experimental details

high-frequency CV curve as described in Figure 2.2. The amplitude of the UHF voltage is small such that it does not interfere with the system output. It is only used as sensing small signal by the capacitive sensor. Some previous studies have highlighted that indeed the UHF might affect the output signal [33] [34].

In reality, the SCM measurement system can be represented by a 3D MOS capacitor. Therefore it is affected by additional parasitic effects like the stray capacitance, due to the presence of the cantilever above the sample (that can be compensated by the microscope controller), and the slight change in the depleted capacitance $C_d$ caused by the fringing fields between the probe and the sample [34], as will be explained further in chapter 3. Moreover, other artifacts arise from the fact that the MOS capacitor might present non-idealities, like oxide trapped charges and interface states at the oxide/semiconductor surface, causing the shift of the CV curves along the x-axis and their stretching out, respectively. However, these artifacts are not taken into account in this work since they are not relevant issues when considering the operating principle as described in the following sections.

2.1.2 SCM operating principle

The possibility to measure local capacitances with a nanometer resolution is given by a very sharp conductive probe with a tip radius on the order of tens of nanometers. This probe, attached at the free end of a cantilever, is scanned on the oxidized surface of the semiconductor as depicted in Figure 2.3. The tip-sample system is biased at a given $V_{DC}$ voltage, theoretically at the flat band condition. The charges are repelled from and attracted towards the surface by means of the kilohertz $V_{AC}$ bias. A larger depletion region is formed in samples with a lower carrier density, and the dynamic capacitance is measured by the detector.

Capacitance detection

The sensor is a high frequency resonator sensitive to very small capacitance changes. It bases on the principle of the RCA capacitive pick up [35]. This is a GHz resonant circuit where the tip/sample system is part of the resonator. A simpler sketch of the circuitry is
2.1 Scanning Capacitance Microscopy

**Figure 2.3:** Scanning Capacitance Microscopy system.

**Figure 2.4:** UHF resonator performing capacitive sensor.
in Figure 2.4. If the UHF signal is small it has a minimal influence on the output signal \[34\] \[36\], but, at the same time it allows the resonator to accomplish the measurements. The oscillator is coupled with a resonator circuit composed of coupling inductances, the HF tuning varactor capacitance \(C_v\), the MOS (tip/sample) capacitance \(C_s\), and the stray capacitance (cantilever/sample) \(C_p\). All three capacitances are biased at \(V_{DC}\) bias and modulated by \(V_{AC}\). In the latest generation of microscopes, the earlier varactor is substituted with a VCO (voltage-controlled-oscillator) to allow a more flexible and stable tuning of the resonator. In such a system, any slight change in the MOS capacitance \(C_s\) changes the resonant frequency of the circuit, with a consequent large drop in the output signal that makes such a kind of detector sensitive to capacitance variations on the order of \(10^{-19}\) F. The output signal is demodulated and represents the \(C_s\) as a function of the time. This circuit is followed by a real quadrature lock-in amplifier (precise and flexible with gains and filters), able to single out mono-frequency signals. It takes as inputs the reference signal \(V_{AC}\) and \(C_s(t)\) and it outputs the instantaneous peak value of \(C_s(t)\). A successive signal processing yields \(\Delta C/\Delta V\) that is the actual output of the microscope (Figure 2.5). Tuning on the linear part of the characteristic by \(V_{DC}\), and the small amplitude of \(V_{AC}\), assures a quasi-linear output and therefore the smallest harmonic distortion [37].

![Figure 2.5: Simplified block diagram of the scanning capacitance microscope.](image)

Lock-in amplifiers are used to detect and measure very small AC signals. They require a reference AC signal, whose frequency is equal
2.1 Scanning Capacitance Microscopy

to the frequency of the signal to be measured. The reference signal and the generic input signal are

\[ V_{\text{ref}}(t) = V_{\text{ref}} \sin(\omega_r t + \theta_r) \quad (2.1) \]

\[ V_{\text{sig}}(t) = V_{\text{sig}} \sin(\omega_s t + \theta_s) \quad (2.2) \]

The lock-in amplifier multiplies the reference by the input signal using a phase sensitive detector (PSD), generating the output

\[ V_{\text{psd}}(t) = V_{\text{ref}} V_{\text{sig}} \sin(\omega_r t + \theta_r) \sin(\omega_s t + \theta_s) \]

\[ = \frac{1}{2} V_{\text{ref}} V_{\text{sig}} \cos[(\omega_s - \omega_r)t + \theta_s - \theta_r] - \]

\[ \frac{1}{2} V_{\text{ref}} V_{\text{sig}} \cos[(\omega_s + \omega_r)t + \theta_s + \theta_r] \quad (2.3) \]

with two AC signals at frequencies \( \omega_s - \omega_r \) and \( \omega_s + \omega_r \). At this point, the \( V_{\text{psd}} \) passes through a low pass filter. The width of the low pass filter establishes the bandwidth of the lock in. If \( \omega_s \neq \omega_r \) no output is measured, whereas if \( \omega_s = \omega_r \) the output of the filter will be

\[ V_{\text{psd}} = \frac{1}{2} V_{\text{ref}} V_{\text{sig}} \cos(\theta_s - \theta_r) \quad (2.4) \]

that is a DC signal proportional to \( V_{\text{sig}} \). In order to measure \( V_{\text{sig}} \), \( \theta_r \) is adjusted to set \( \theta = \theta_s - \theta_r = 0 \) and \( V_{\text{psd}} \) becomes:

\[ V_{\text{psd}} = \frac{1}{2} V_{\text{ref}} V_{\text{sig}} \cos(\theta) = \frac{1}{2} V_{\text{ref}} V_{\text{sig}} \quad (2.5) \]

\[ V_{\text{psd}} \sim V_{\text{sig}} \quad (2.6) \]

In a quadrature lock-in amplifier, there are two PSDs, one delivers an output proportional to \( \cos(\theta) \) and the other an output proportional to \( \sin(\theta) \). The first is usually called \( X \) and defined as the “in phase” component and the second \( Y \) defines the “quadrature” component of the signal

\[ X = V_{\text{sig}} \cos(\theta) \quad (2.7) \]

\[ Y = V_{\text{sig}} \sin(\theta) \quad (2.8) \]
The phase dependence is removed once the magnitude of the signal is computed

\[ R = \sqrt{X^2 + Y^2} = \sqrt{V_{\text{sig}}} \]  \hspace{1cm} (2.9)

If the input signal is not a pure sine wave it will be composed of many sine waves with different amplitude, frequency and phase. The PSD, multiplying the input to a pure sine wave will extract the Fourier component of the input signal at the reference frequency. This means that in SCM the operation in the non-linear part of the CV curve will produce a loss of information, although the most important feature of the lock-in is the response to noise. In general, the input signal is not noise-free. If the amplitude of the signal is much smaller than the noise it might be impossible using normal filtering to separate the signal from the noise. The advantage of a lock-in amplifier is that any noise at other frequencies but the reference is removed by the low pass filter.

### 2.2 Scanning Spreading Resistance Microscopy

Scanning Spreading Resistance Microscopy (SSRM), like Scanning Capacitance Microscopy, has been developed as a microscopy tool and 2D carrier profiler for semiconductors. It is implemented on an AFM system and uses a dedicated SSRM sensor that performs the measurements of the local resistance under the tip in contact with the sample surface (Figure 2.6). In order to realize a good electrical contact it becomes necessary to push the tip with forces that are higher than those usually applied in contact AFM, such that the thin native oxide is penetrated.

A constant voltage is applied to such a system and the injected current is measured. A good electrical backside contact is required in order to avoid any series resistances that might influence the measurements. The contact/semiconductor electro-mechanical model has been studied \[38\] \[39\] \[40\] intensively in the past. It proves the existence of an elasto-plastic deformation of the semiconductor crystal underneath the contact due to the high pressures applied. According to this model the electrical contact is created by means of the induced
plastic phase, which is a region of high conductivity characterized by a metallic-like behavior.

**Resistance sensor**

The resistance sensor is a logarithmic amplifier, which provides a logarithmic voltage output for a linear input current. This circuit is very accurate over seven decades of operation. Temperature compensation and linearity optimization is included in the circuitry. The relation between the input current and output voltage is expressed by

\[
V_{out} = +\log\left(\frac{I \times 10^6}{V_{DC}}\right) \text{ if } V_{DC} > 0 \tag{2.10}
\]

\[
V_{out} = -\log\left(\frac{I \times 10^6}{V_{DC}}\right) \text{ if } V_{DC} < 0 \tag{2.11}
\]

The output of the system is given in Volt, and the transfer curve relating the voltage scale to the resistance scale is represented in Figure 2.7. The detector is calibrated such that a change of 1 V corresponds to a change in resistance of one order of magnitude, and an output

*Figure 2.6: Scanning spreading resistance operating principle.*
of 0 V corresponds to 1 MΩ resistance. The dynamic range covers doping concentrations ranging from $10^{14}$ up to $10^{20} \text{ cm}^{-3}$.

### 2.2.1 SSRM theory of operation

The theoretical approach to the spreading resistance can be described by a simple model consisting of a flat circular ohmic contact and a hemispherical ohmic contact. The resistance between these two contacts can be calculated analytically [41] [42].

Under given conditions this model describes in first approximation the tip/sample system. The radius of the circular contact is $a$ and the distance between the center of it and the semi-spherical contact is $r$, as shown in Figure 2.8.

The resistance is given by

$$R_r = \frac{\rho}{2\pi} \int_0^r \frac{dr}{(a^2 + r^2)} = \frac{\rho}{2\pi a} \tan \frac{r}{a}$$  \hspace{1cm} (2.12)

where $\rho$ is the resistivity of the medium between the contacts. The case of a semi sphere radius of $r=\infty$, is a good approximation of the

---

**Figure 2.7:** Transfer characteristic of the logarithmic amplifier.
2.2 Scanning Spreading Resistance Microscopy

Figure 2.8: Geometry of a flat circular tip and a semi sphere.

spreading problem from a microscopic tip contact to a semi-infinite volume. Equation 2.12 reduces to

\[ R = \frac{\rho}{4a} \]  \hspace{1cm} (2.13)

The circular contact can be assumed as the tip contact surface and the infinite semi-volume as a uniform sample. In reality, the size of the sample is not infinite but its typical size is much larger than the tip radius. It has been calculated \[24\] that if the backside contact is grounded, about 80\% of the total tip voltage drops within a distance that is on the order of five tip radii.

Therefore, it can be assumed that in a SSRM system the resistance is described by Equation 2.13. This expression indicates that, in the case of perfectly ohmic contacts, the spreading resistance depends upon the radius of the tip and the resistivity of the sample, only. The knowledge of the resistance, hence of the resistivity, is related to the carrier concentrations by the following formula

\[ \rho = \frac{1}{q(\mu_n n + \mu_p p)} \]  \hspace{1cm} (2.14)

This indicates that the resistivity depends on the product of the free carrier concentration to their mobility. With the use of the Irvin’s curves in Figure 2.9, the resistivity is immediately converted to the doping concentration under the assumption of full activation of the doping impurities. Calibration becomes easier than in SCM, although additional artifacts arise as will be explained in chapter 4.


2.2.2 Probe contact and resolution of SSRM

The previous theoretical model assumes a flat probe contact at the surface. In reality the tip is not flat, and the high forces applied to the probe lead to the indentation of the probe tip (made of hard material) into the sample. Therefore, assuming a probe modeled as a semi-spherical tip with radius of curvature $r_c$, the indented contact results in a radius $a$ smaller than $r_c$ \[28\]. Moreover, the pressure applied to the probe, after causing penetration, induces a phase change in the underlying semiconductor depending on the magnitude of the force, the semiconductor type and the doping (Figure 2.10). The transition to the high-conducting phase appears where the maximum shear stress occurs, i.e. at about $0.48a$ from the tip apex \[40\], and extends in three directions as the pressure increases. According to this, once it reaches the probe, it produces an ohmic contact under the tip \[43\]. The hypothesis that it affects also the resolution (related to the “effective” tip contact size) is still under investigation. Early studies have highlighted a phase transition in semiconductors under pressure, related to a change in volume of the crystal and a consequent
drop in the resistivity, which brings the semiconductor into a metallic phase. Minomura [39] proposed two different possible transitions: solid-to-liquid phase or a most probable solid-to-denser-solid phase. For all semiconductors under investigation, (except GaAs) a denser solid phase was hypothetically assumed and a resistance of $10^{-4} \, \Omega \, \text{cm}$ was measured above the pressure threshold characteristic of each material. Later studies of Jamieson using X-ray diffraction [44] [45]

![Figure 2.10: Contact radius due to tip indentation.](image)

on silicon, germanium and compound materials revealed that their crystal structure changes at high pressure to a metallic *white-tin* structure. A further investigation over a larger pressure range (0-50 GPa) was done by Olijnyk [46]. It showed that, for silicon, the increase of the pressure transforms the initial Si(I) to Si(II) or $\beta$-tin phase into more complex phases. In contrast, for germanium the behavior is much simpler: $\beta$-tin remains up to 51 GPa. Hu [47] and Zhao [48] described the hysteresis led by the transitions on application and release of pressure in silicon. Several metastable phases result from decompression, and in general case an elasto-plastic deformation takes place in the semiconductor. In more recent studies, the phase transition was classified differently for compound semiconductors of II-VI, III-V, and IV group, for which the existence of the $\beta$-tin is proven not to exist, but other crystal changes are induced [49].
**Silicon**

Silicon at normal temperature and pressure is in the diamond semiconductor phase. A compression under hydrostatic pressure induces several metallic phases above 11.5 GPa. On application of non-hydrostatic (shear) stresses this value is reduced to about 8 GPa [38] [50] [51]. The transformation from the cubic crystal structure to $\beta$-tin produces a volume reduction of 22%, and upon release of the load an intermediate phase between Si(I) and Si(II) with a volume reduction of 8% is reported. Unlike hard materials and low-compressibility materials, silicon exhibits a small shear modulus and bulk modulus. It is then relatively easy to change its shape and volume through an external stress. Undoped silicon and $n^+$ doped silicon exhibit larger hardness than $p^+$ doped silicon [52].

### 2.2.3 Electro-mechanical model

A simple circuit model describing the operation of SSRM is shown in Figure 2.11. The total resistance is composed of a series connection of

- the resistance of the diamond probe $R_{probe}$,
- the contact resistance $R_c = \rho_c / \pi a^2$ between probe and semiconductor,
- the mechanical contact resistance between probe and semiconductor $R_m = \rho_m / \pi a^2$,
- the spreading resistance (i.e. for a non-indented tip) $R_{sp} = \rho / 4a$, and
- the back contact resistance $R_{bc}$.

If no load is applied, due to the high doped diamond coating of the probe and the large back contact, the resistances $R_{probe}$ and $R_{bc}$ have a minor role and they can be neglected. Therefore the total contributions are

$$R_{total} = R_c + R_m + R_{sp} \quad (2.15)$$

The resistance $R_c = \rho_c (V, \varphi_B) / \pi a^2$ is a diode-like non-linear function of the electrostatic barrier $\varphi_B$ between probe and semiconductor,
Figure 2.11: Schematic of a simple SSRM circuit model.

and of the applied voltage $V$, which depends on the type of conduction mechanisms (thermionic emission, tunneling). The resistance of the mechanical contact $R_m$ can be defined as the non-voltage-dependent contribution of the tip/sample Schottky structure. A contact resistance is in general defined as

$$R_m = \frac{\rho_m(P)}{A}$$  \hspace{1cm} (2.16)

where the contact resistivity $\rho_m$ is a function of the pressure applied, and $A$ is the macroscopic area in contact. In case of SSRM, the probe and sample are electrically connected by loading the tip against the surface of interest, with a macroscopic contact area $A=\pi a^2$. Asperities are present in the contacted surfaces reducing the effective contact area, which results into an increase of $R_m$. Therefore, for negligible or low pressures applied, the following occurs

$$R_c + R_m > R_{sp}$$  \hspace{1cm} (2.17)

or

$$\frac{\rho_c + \rho_m}{\pi a^2} > \frac{\rho}{4a}.$$

This means that the overall contact resistivity is too high to allow a SSRM measurement to be accomplished [53].

The presence of a load applied to the tip induces the aforementioned phase change, yielding a high-conductive region that consequently transforms the tip-semiconductor interface from a Schottky to a pseudo-ohmic contact. This transformation entails two local changes
in the semiconductor: the shrink of the band gap and a reduction of interface states. Either of the two effects causes a lowering of the barrier at the interface, making $\rho_c$ also load-dependent. Hence, $\rho_c(V, \phi_B)$ has a much lower value at high pressure and $R_c \leq R_{sp}$. Furthermore, the increase of the tip load results into a larger contact area making the total effective contact surface larger. At a certain pressure the effective contact area becomes high enough, yielding $R_m \leq R_{sp}$ and enabling the system to measure the spreading resistance $R_{sp}$.

Thus, many effects occur once the tip is loaded and biased above the threshold. The first positive consequence is the possibility to neglect the parasitic series resistances over the spreading resistance. These mixed contributions, however, produce a non-linear resistance-resistivity response, which requires an accurate calibration of the measurements for the quantification of SSRM data [54].

### 2.3 Probes for SCM and SSRM

While probes and cantilevers used for SCM and SSRM require respectively particular characteristics, some requirements are common for the two applications since they are both high resolution scanning probe techniques in contact mode [55]:

- The tip should be as sharp as possible, ideally with atomic size. In practice, with the most up-to-date technology process, it is possible to fabricate probes with a radius of curvature of about ten nanometers. The importance of the size of the apex of the probe is crucial since it is one of the components that define the resolution of the techniques.

- A second important requirement is the hardness. For contact techniques, the rupture during scanning has to be avoided since it may reduce the sensitivity of the tip and the control of the measurements. If the tip is coated, a hard coating layer prevents the wear out of the conductive outer layer.

- The electrical conductivity of the probe has to be high, preferably similar to the conductivity of a metal.
2.3 Probes for SCM and SSRM

- A good reflectivity of the top surface of the cantilever is also necessary for the photo-detection laser system.

![Diagram](image)

**Figure 2.12:** (left) SSRM and (right) SCM probe shape requirements.

All these characteristics are essential for SCM and SSRM, which separately require some other different features: scanning capacitance involves a soft contact and forces applied to the cantilever on the order of nN. The cantilever stiffness is typically in the range 0.1-5 N/m. Spreading resistance needs a hard contact, thus, in this case, the cantilever must be stiffer (20-100 N/m) because forces up to mN are applied, depending on the hardness of the material under measurement. The nature of SSRM does not allow to perform good topography images, for which a soft contact is indispensable. The shape of the probe is also different for SCM and SSRM as shown in Figure 2.12.

In the picture on the left the typical shape of a SSRM tip is sketched. The wide aperture angle allows to apply larger forces and to make the tip more resistant to the friction and the indentation through the native oxide. For SCM it is better to use tips with higher aspect ratio in order to minimize the electrical tip/sample interaction (fringing fields). The probes commercially available for SCM are silicon metal-coated and silicon-doped-diamond-coated. The first are sharper (radius 10-30 nm) but they wear out faster, whereas the second are thicker (60-100 nm) and much more resistant to the mechanical damage. A comparison under operation of the two probes reveals that the performance in scanning of the diamond coated tips is higher in terms of reproducibility of the measurement [56]. The high conductivity of such kind of tips is due to the high boron doping of the diamond coating layer. The limitation of the diamond coating
is the reduced sharpness due to the conformal coating of the probe. However the reduction of the coating film makes the wear out faster. For SSRM, as well as for SCM diamond coated probes are available. The actual trend is going towards full diamond probes for both the techniques. The advantages are the higher resolution due to the sharp probe apex, the better conductivity, and the mechanical robustness.

2.4 Features of scanning microscopes

The characteristics of instruments like SCM and SSRM are defined in this paragraph following a previous denomination of De Wolf [40]. Here they are summarized and compared for the two techniques.

2.4.1 Concentration resolution

The concentration resolution indicates the minimum variation of concentration that can be detected by the system (Figure 2.13). The average noise level in the measurements sets the minimum of the concentration resolution.

![Figure 2.13: Definition of the concentration resolution.](image)

2.4.2 Spatial resolution

The spatial resolution is defined as the minimum distance, between two doping steps or spikes, that the system is able to detect, see Figure
2.4 Features of scanning microscopes

Figure 2.14: Definition of the spatial resolution.

2.14. SCM spatial resolution is variable and related to the tip used for the measurements and to the sample under investigation. The presence of low doping concentrations in the specimen yields large depleted regions and to an averaging effect due to this interaction. In case of high doping the depletion has a small impact because its size is comparable to the tip size, which is responsible for the carrier averaging effect. The presence of different concentration ranges within a sample, implies a site-dependent resolution. The minimum value is reached for very high concentrations and corresponds to the size of the probe contact. The maximum can be observed for low doping concentrations and is on the same order of magnitude of the depletion region under the probe. The spatial resolution in SSRM is defined by the tip/sample volume of interaction. It is considered here as the volume where the tip voltage drops by 80-90% of the initial value. Therefore, the spatial resolution in SSRM does not depend on the doping but only on the geometrical characteristics of the probe tip and on the distribution of the spreading current. Another aspect is that this parameter does not vary through the measure but remains fixed with the probe tip dimension, considering its size not affected during the measure. To give an estimation of this extent, assuming the characteristic radius of a diamond coated probe that might be ranging between 40 nm to 60 nm, the interaction volume is evaluated to be on the order of hundreds of nanometers. This value might be
one order of magnitude smaller if the effective contact size is reduced by the asperity of the diamond coating and the indentation of the tip. A prior characterization of the tip shape, dimension, and degree of indentation, would lead to a more precise evaluation of the spatial resolution.

2.4.3 Spatial accuracy (depth resolution)

The spatial accuracy indicates the ability of the system to locate edges and other geometrical features respect to the doping distribution, see Figure 2.15. The detection of features smaller than a certain size is a relevant characteristic for the application of SCM and SSRM as microscopy techniques or, in the case of dopant profiling, for the detection of edges. The acquisition of topography or deflection maps at the same time as capacitance or resistance maps, is a great support to locate particular features and to improve the spatial accuracy. In literature, embedded oxide layers of tens of nm have been detected [57] with the use of SCM and SSRM. The estimation of the oxide thickness is done comparing the size of the measure to the tip dimension. Recently, a study on SOI devices [58] has demonstrated the higher spatial accuracy obtained by the full diamond probes with the possibility to detect embedded layers of oxide as large as 3 nm.

Figure 2.15: Definition of the spatial accuracy.
2.4.4 Gradient resolution (sensitivity)

The definition of gradient resolution or sensitivity refers to the ratio between the carriers gradient and the corresponding output of the system. Sensitivity to carrier distribution changes is very high for SCM and SSRM detectors.

The impact on the SCM and SSRM output of two different gradients $\delta N/\delta x_1$ and $\delta N/\delta x_2$, is

$$\frac{\delta N}{\delta x_1} > \frac{\delta N}{\delta x_2} \text{ if } \delta x_1 < \delta x_2$$  \hspace{1cm} (2.19)

Thus, the steeper the slope the higher the ratio between carriers gradient and the output of the system. This leads to a decreased sensitivity, see Figure 2.16. Moreover, the sensitivity is affected by the interaction volume of the tip probe. In fact, the averaging effect is much stronger on steep gradients, i.e. comparable to the depleted region under the tip (low doping) or the tip size (high doping), than on profiles distributed on larger extents.
2.5 Sample preparation for silicon cross-sections

The preparation of cross-sectioned silicon sample for SCM is a critical step. The main purpose of this operation is to prepare a specimen as ideal as possible. Hence, a sample with a flat and smooth surface, a homogeneous oxide thickness, and a reduced density of interface states. These characteristics are necessary to obtain reproducible results for every sample. SSRM does not require to grow an oxide. Furthermore, the surface condition is not very critical, since the quality of the signal can be easily controlled by the pressure applied to the probe. The method for the preparation of SCM and SSRM samples has become a standard. The final result depends strongly on the environmental cleanliness and on human skills whether executed manually or using the polishing machine. The following procedure is used for manual polishing.

Equipment and consumables

The basic equipment necessary for cross-sectioned sample preparation consists of

- Tungsten tip for cleaving or fine diamond saw,
- Epoxy resin (bi-component) and conducting epoxy glue,
- Rotating plate with variable speed, water tap and sink,
- Diamond polishing disk films,
- Polishing cloth for Colloidal Silica,
- Micro organic soap,
- Hydrofluoric acid and hydrogen peroxide,
- Hot plate and U.V. system,
- Silicone rubber.
Mounting of the sample

A small piece of the silicon sample is cut out the wafer with a diamond saw or by cleavage. The organic residues and particles are removed by immersion for few minutes in acetone (in a glass container) with simultaneous use of ultrasounds. More effective cleaning procedures (e.g. RCA cleaning) are not necessary, since additional contaminations are introduced in the next steps of the preparation. At the end of the cleaning the sample is washed under running deionized water. The sample is dried with compressed air (or preferably with nitrogen) and glued, using an epoxy resin, to a dummy silicon or a glass that was previously cleaned in the same way (Figure 2.17).

![Dummy-sample sandwich](image)

**Figure 2.17:** Dummy-sample sandwich.

In order to have the thinnest and even layer of glue that guarantees the best adhesion, a pressure should be applied to the sandwich till the resin is cured at 150°C for 30 minutes. It is important to identify which part of the sandwich is the dummy and which part is the sample because grinding and polishing require a specific direction depending on the respective position of the sample and of the dummy silicon. The backside of the sample is contacted electrically to the sample holder. This requires the sample to be sputtered with gold to improve the electrical contact. Before sputtering, the native oxide on the top of the contacting surface is removed by grinding. The sandwich is now ready to be glued to the aluminum sample holder. For this purpose a conductive epoxy resin is used. A small amount of it is applied to the sample holder, where the specimen is attached, by applying a slight pressure. The epoxy resin needs to be cured for 50 minutes at 150°C.
Grinding and Polishing

The procedure includes the grinding of the top surface using a rotating plate and diamond films in the following order of grain size: 30-15-9-6-3-1-0.5-0.1 μm. The films consist of precision graded diamond particles bonded onto a polyester film. In order to avoid to round off the sample edges, which heavily affects the measurements, the direction of rotation of the plate is from the dummy to the sample (Figure 2.18). During grinding and polishing the surface has to be continuously monitored with an optical microscope in order to check possible non-uniformity in grinding and/or the presence of undesirable particles, which have to be removed under running deionized water using for example a cotton swab.

At the end of the polishing phase the surface must be as scratchless, clean and flat as possible. Final polishing is made with Colloidal Silica that is a colloidal suspension of a mixture of abrasive particles dispersed throughout a chemically active liquid carrier. This combination provides a simultaneous chemical and mechanical polishing action. The direction of rotation for the final polishing with Colloidal Silica is opposite to the direction of the grinding. It is not required any pressure on the sample. The time necessary for polishing is about one minute, depending on the depth of the chemical etch allowed. The best SCM results have been obtained from samples with a very smooth surface with a RMS roughness better than 1 nm, which have been prepared using Colloidal Silica for a longer time.

*Figure 2.18: Sample grinding with diamond lapping films.*
2.5 Sample preparation for silicon cross-sections

Colloidal Silica performs a doping selective etch, which reduces the surface flatness. Nevertheless, it is not usually so relevant for the measurements. Very small particles of Colloidal Silica might be embedded in the sample. In this case it is necessary to clean the sample thoroughly under running deionized water with a cotton swab, while applying a certain pressure. A Micro Organic soap is also used to remove the particles of dust and water residuals and it might be necessary to use an ultrasound bath. Monitoring of the cleaning process should be done at each step, either with the optical microscope or by AFM. At this point the specimen is ready to be measured with SSRM.

**Oxide growth (only for SCM measurements)**

Once the sample is perfectly clean the next step consists of removing the native oxide by a 1:10 aqueous solution of hydrofluoric acid (HF). HF also contributes to passivate the dangling bonds, which may introduce unwanted interface states. Since metal and glass are etched by HF, it is necessary to use Teflon tools (container and tweezers). Few drops (2-3) of diluted HF on the surface of the sample are enough to remove the native oxide (around ten seconds). Immediately after, the sample is rinsed under running deionized water for 20-30 seconds. The sample surface should be checked again to verify that is clean. The oxide growth is performed in two ways:

- In air at 300°C under ultraviolet for 40 minutes.

- At 70°C into hydrogen peroxide for 10 minutes. When using the method of wet growth it is necessary to use silicone rubber to cover the conductive epoxy glue to avoid etching and redeposition of the silver. The sample is immersed in hydrogen peroxide (70°C) for ten minutes. It is important to put the sample far under the liquid surface, otherwise small floating particles might adhere on it. The sample has to be rinsed after the treatment. After checking again the state of cleanliness of the specimen, the oxide is annealed on hot plate at 120°C in air for 20 minutes. Finally, the sample is ready to be measured with SCM. Both the oxidation techniques result in an oxide thickness of 3-4 nm. It
has been shown [59] that an oxide thickness of 3 nm is the minimum required to avoid leakage through the oxide and to obtain a high capacitance signal. Clearly, the preparation involves the use of chemicals, which can react with metals or other materials that might be present on the sample if it is for example a real devices in a wafer. This makes the preparation more important, and special attention has to be paid at each use of chemicals. Sometimes, when metal redeposition may occur, it is preferable not to use HF and perform the measure on the native oxide.

The effect on the measures of these two methods has been investigated by Yabuhara [56]. He showed that the dry oxidation in combination with the use of diamond probes provides the best reproducibility of the results. On the contrary, the wet oxidation delivers more homogeneous oxide layers, and for this reason it is more suitable for samples with a large surface.

2.6 Measurements on beveled samples

Beveled samples were used in the past for Spreading Resistance (SPR) measurements. At present their use is diffused also for SCM and SSRM. The main advantage to perform probe measurements on a bevel rather than on cross-sections is the geometric magnification factor along the main bevel axis as shown in Figure 2.19. The sketch

![Figure 2.19: Cross sectioned (left) and beveled (right) p++n junction.](image)
on the left side of Figure 2.19 shows an orthogonal cross-section. The junction is located at a distance $d$ from the edge. The same structure beveled at an angle $\alpha$ is represented on the right side of Figure 2.19. In this case the depth of the junction from the edge is $D > d$ and the enlargement factor $M$ is defined as

$$M = \frac{D}{d} = \frac{1}{\sin \alpha}. \quad (2.20)$$

The magnification of the distances along the main section axis increases the geometrical resolution, basing on the fact that, if no smaller tip sizes are available the sample needs to be enlarged. The magnification factor $M$ depends on the bevel angle, for a $M = 10$ the bevel angle is $\alpha = 5^\circ 44'$. The values obtained from beveled structures have to be demagnified in order to get back the original dimensions. The application of this method presents a major artifact, known as carrier spilling. This is the redistribution of the carriers into the semiconductor due to the change in geometry. It causes a distortion of the original carrier distribution and leads to some measurement problems [60]. One of the drawbacks affects the electrical junction delineation, especially for very shallow structures. The bevel introduces a systematic shift in the location of the electrical junction that cannot be avoided.

**Beveled samples preparation**

Sample preparation for bevels has been described in detail in [61]. The procedure consists in rigorous steps intended to produce a scratch free and uniform surface.

**2.7 Sample description**

In this work the samples measured with Scanning Capacitance and Scanning Spreading Resistance are silicon uni- and bipolar structures. Some samples are produced for calibration and their doping concentration is accurately known. The unipolar staircases, the implanted junctions, and the bipolar transistor are produced at IMEC (Leuven), and have been used for a round robin across several independent laboratories. The main purpose of these samples is to define the calibration
\textbf{Figure 2.20:} Resistivity (solid) and impurity density (dashed) profiles of the p-type staircase.

\textbf{Figure 2.21:} Impurity density (SIMS) vs. resistivity profile (SRP) of the p-type staircase.
curves, and to establish the sensitivity and repeatability of the results. The low-doped junction has been manufactured at Infineon (Villach).

### 2.7.1 Unipolar samples

The unipolar samples are two ($p$- and $n$-type) silicon epitaxial staircases. This kind of structure combines the following requirements: a wide dynamic range and a homogeneous concentration for a sufficient width in the layers. Thus, they are adapted for calibration of SCM and SSRM.

![Figure 2.22: Resistivity (solid) and impurity density (dashed) profiles of the n-type staircase.](image)

**P-type silicon**

The $p$-type staircase (T8) [62] is composed of seven different epitaxial layers with a doping concentration ranging from $7.5 \times 10^{19}$ down to $3.8 \times 10^{15}$ cm$^{-3}$. The substrate concentration is $2.7 \times 10^{15}$ cm$^{-3}$. Each layer is separated from the neighboring by higher doped buffers. The
reason is due to the fact that the sample was designed for the calibration of SRP and the buffers provide an optimization of the resistance correction factor. The width of each layer is about 5 μm. The impurity and resistivity profiles measured by conventional SRP are shown in Figure 2.20. The doping profile was calculated from the measured resistance profile (not available) and compared with the doping profile acquired with SIMS. This was consistent with the SRP measurements. The doping density respect to the resistivity is plotted Figure 2.21. The presence of high-doped buffers makes the resistivity data set non monotonic. As shown in Figure 2.21 the discrepancy between SIMS and SRP measurements is lower than 10%.

![Figure 2.23: Impurity density (SIMS) vs. resistivity profile (SRP) of the n-type staircase.](image)

**N-type silicon**

A similar staircase was manufactured for n-type doped silicon. The n-sample (T9) is composed by four layers of different doping starting from $1.4 \times 10^{19}$ down to $5 \times 10^{15}$ cm$^{-3}$ on a low doped substrate ($4.2 \times 10^{14}$...
cm\(^{-3}\)). Every layer is separated by highly doped buffers. The structure, whose doping distribution and resistivity is represented in Figure 2.22 is about 30 µm large. As shown in Figure 2.23 discrepancies in the 10% range are observed in close vicinity of the buffers.

### 2.7.2 pn junctions

The bipolar samples are three \(pn\) junctions, whose doping profiles are plotted in Figures 2.24, 2.25, 2.26. The first two junctions, named I10 and I12, are highly doped implants (\(\sim 10^{20}\)). They are representative for the source-drain implants in transistors [63].

![Doping profile of the I10 p++n junction as measured with SIMS.](image)

**Figure 2.24:** Doping profile of the I10 \(p^{++}n\) junction as measured with SIMS.

The third junction is a low doped junction, with an epitaxial \(n\)-layer doped \(5 \times 10^{15}\) cm\(^{-3}\) (thickness 1.4 µm) on a \(p\) substrate doped 1 \(10^{15}\) cm\(^{-3}\). The advantage to have an epi-structure is that the location of the metallurgical junction is exactly known, without the need to perform any additional measurements. This sample was manufactured to have a wide depletion region for SCM application and for SSRM studies of low-doped bipolar samples.
Figure 2.25: Doping profile of the I12 n++p junction as measured with SIMS.

Figure 2.26: Doping profile of the epitaxial low-doped junction.
2.7.3 Bipolar transistor

This transistor exhibits microscopic features intended to verify the resolution of the microscope. The npn bipolar transistor is drawn in Figure 2.27. The critical dimension is represented by the width of the p-base (0.3 μm).

2.8 Inverse problem

An important aspect of the quantitative use of SCM and SSRM is their use as dopant profiling techniques. The inverse problem consists of the computation of the active dopant density profile from the free carrier profile delivered by SCM and SSRM. Since the measured resistance and capacitance values are expressed as voltages, they first need to be converted into a carrier profile by means of calibration curves. The calibration curves can be obtained either from dedicated calibration samples, or, as in the case of SCM, from dedicated computational tools like SCaMsim [64] and FASTC2D [65]. These software tools provide the 3D capacitance value of a probe with a given shape positioned on the top of an oxidized semiconductor sample with uniform doping.

The experimental calibration procedure is performed preparing a two-specimen sample. The first specimen of unknown concentration
is glued to a second sample of known concentration (e.g. staircase samples) in order to measure them at the same time.

Both the techniques SCM and SSRM lead to a local alteration of the carrier concentration due to the interaction with the probe. This causes an averaging of the measured signal over a distance defined by the size of the interaction volume. This unwanted effect can be usually purged by special deconvolution algorithms [34].

De Wolf [40] has proposed an iterative procedure for SSRM. It consists in solving the equation

\[ R = CF(a, \rho) \left( \frac{\rho}{4a} + R_{cTotal}(\rho) \right) \]  

(2.21)

in respect to \( \rho \), where \( CF \) is the correction factor. A first guess of the solution is provided by the experimental calibration curves, i.e. by

\[ R = \frac{\rho}{4a} + R_{cTotal} \]  

(2.22)

\( \rho \) is calculated at each step of the iteration by taking into account the vicinity of edges, oxides, or other doped layers. This allows to extract a suitable correction factor \( CF \). The calculated doping profile is entered in the loop to provide a new estimate. The stopping criterion is given by the difference between two successive profiles \( \rho_i \) and \( \rho_{i+1} \), which has to be smaller than a given \( \varepsilon \). Once the best resistivity profile has been assessed, the selected carrier and doping profiles are computed by using at the same time the mobility equation

\[ \rho = \frac{1}{n_e q \mu_e + n_h q \mu_h} \]  

(2.23)

and the Poisson equation

\[ \nabla^2 \Psi = -\frac{q}{\varepsilon_s} (N_D - N_A + p - n). \]  

(2.24)

Unfortunately, the use of this technique is restricted to the 1D case because of the limitations due to the computation time.

### 2.8.1 The direct method

The direct conversion method assumes that the local doping density is equal to the local carrier concentration. This approach is used
by SCaMsim and FASTC2D. The assumption is correct if the doping profile does not change too much over distances that can be compared to the local Debye length. This is not the case of abrupt doping profiles, like in ultra shallow junctions, where the direct method is not accurate enough and the conversion requires more sophisticated algorithms.

### 2.8.2 The traditional mathematical approach

A set of five equations links the free carriers distribution to the active doping, the Poisson equation

\[ \nabla^2 \Psi = -\frac{q}{\epsilon_s} (N_D - N_A + p - n) \quad (2.25) \]

the current equations

\[ J_n = q n \mu_n E + q D_n \nabla n, \quad J_p = q n \mu_p E + q D_p \nabla p \quad (2.26) \]

and the continuity equations

\[ \frac{n}{t} = \frac{1}{q} \nabla J_n + G - U_n, \quad \frac{p}{t} = \frac{1}{q} \nabla J_p + G - U_p. \quad (2.27) \]

The computation of the doping profile from a known free carrier distribution (inverse modeling problem) is not obvious, since it requires the inversion of a five non-linear second order differential equation system with boundary conditions. Attempts to solve analytically the inverse modeling problem under somewhat restrictive assumption have been reported in the past [66] [67].

A simpler mathematical formula developed by Kennedy and O’Brien (KOB) [66] expresses the doping as a function of the free carriers concentration in 1D

\[ N_A(x) = p(x) - \frac{e k_B T}{q^2} \frac{d}{dx} \left[ \frac{1}{p(x)} \frac{dp(x)}{dx} \right] \quad (2.28) \]

where \( p(x) \) is the carrier concentration along the \( x \) direction. The substitution of the doping concentration \( N_A \) with the carrier concentration \( p(x) \) in the expression for the Debye length

\[ \lambda_d = \left( \frac{\epsilon_S k_B T}{q^2 N_A} \right)^{\frac{1}{2}} \quad (2.29) \]
yields the KOB formula

\[ N_A(x) = p(x) - \left[ 1 - (\lambda_p)^2 \frac{d^2[\ln p(x)]}{dx^2} \right]. \]  \hspace{1cm} (2.30)

This equation is the analytical solution of the inverse problem. The major inconvenient is the very high sensitivity to the noise, requiring initial smoothing algorithms producing an averaging effect that is not desirable. A possible new different approach is proposed in \cite{68}, suggesting the conversion of carriers to doping profile with the use of neural networks, described in the next paragraph.

### 2.8.3 Solution of the inverse problem with neural networks

An Artificial Neural Network (ANN) is an information processing paradigm inspired by the way the brain process information. It is composed of a large number of highly interconnected processing elements (neurons) working together to solve specific problems. ANNs, like people, learn by examples. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process \cite{69}.

In this case they are applied to the recognition of a smooth profile (carriers) to be transformed into a more abrupt profile (doping). The network is trained first to solve the direct problem, thus, the calculation of the carriers from the doping profile. The direct problem is of easy solution since the doping distribution is univocally related to the carriers. Simulation (DESSIS) of the carriers profile is used to validate the network output. The network is trained in order to obtain a maximum discrepancy of 2.2% respect to the simulation. The sampling procedure is based on the local Debye length. The synthesis of the Artificial Neural Network (ANN) requires the definition of

- the activation function of the units,
- the output function of the units,
- the definition of the optimum number of layers and for each layer
  the number of units,
2.9 Summary

The use of proper probes and a careful surface preparation are requirements that need to be fulfilled in order to obtain quantitative doping profiles from SCM and SSRM experimental data. The sample preparation procedure and the cantilevers to be used are important criteria that define the quality of the results. It has been proven that

- the selection of the most suitable learning method,
- the computation of a complete set of examples to train the ANN.

The structure of the ANN is depicted in Figure 2.28. For the solution of the inverse problem the same network is used, trained with a set of 2500 sequences computed by DESSIS. Accuracy and sensitivity of the ANN are demonstrated with particular profiles, like the SEMATECH calibration standard and doping spikes profiles.

A very good aspect of the network is its immunity to the noise. In fact, even though in absence of noise the result of the network is comparable to the result given by the KOB formula, in case of noisy data the network renders much better results respect to the analytical procedure.

**Figure 2.28:** Structure of the artificial neural network to solve the direct problem.

- the selection of the most suitable learning method,
- the computation of a complete set of examples to train the ANN.
the procedure for quantification of the measured data based on Artificial Neural Networks is more robust and less prone to the noise than the traditional analytical approach.
Chapter 3

Scanning Capacitance Microscopy

This chapter presents different cases, where SCM has been used success-fully both for the determination of 1D doping profiles and for the delineation of the electrical junction. The imaging capabilities of the technique have been tested on real devices, while the quantitative aspects have been assessed by using dedicated test structures. Many aspects of the technique are investigated, in particular the influence of the applied AC and DC bias. This latter affects the signal contrast due to the different working point of the detector. As a consequence, the bias also have an influence on the position of the electrical junction. The quantification of the SCM signal is the weak point of the technique, since complex algorithms have to be used to render the doping profile [34] [68] [70].

3.1 Calibration staircase p- and n-type

The first measurements on a p-staircase (sample already described in chapter 2), show very good results. The monotonic profile reflects the doping concentration variation as shown in Figure 3.1. The sample-tip system is biased with a DC voltage $V_{DC} = -0.2$ V, and an AC voltage $V_{AC} = 0.5$ V. The application of very large AC voltages is not
recommended because of the averaging effect produced. However, $V_{AC}$ has to be large enough to produce an acceptable signal level.

![Figure 3.1: Measured SCM profile of the p-type staircase.](image)

The SCM output is expressed in a voltage level proportional to the instantaneous $dC/dV$ signal. The conversion in capacitance is not straightforward, because it requires calibration data obtained usually by simulation. Nevertheless, since the interpretation of the profile is not ambiguous, in this chapter the SCM output will be expressed with the raw signal units, i.e. in V. The measured SCM profile for the $n$-staircase is plotted in Figure 3.2.

The behavior at a bias of $V_{DC} = -2$ V and $V_{AC} = 0.875$ V is monotonic. The profiles are chosen among others acquired at different biases, basing on the quality of the output contrast. Because of the large range of doping characterizing this structure, large differences in flat band voltages occur between different epitaxial layers and the best solution has to be found.

The calibration curves extracted from these measurements are shown in Figure 3.3 for the $p$-staircase and in Figure 3.4 for the $n$-staircase, respectively. The theoretical calibration curves in Fig-
3.1 Calibration staircase p- and n-type

**Figure 3.2:** Measured SCM profile of the n-type staircase.

**Figure 3.3:** Calibration curve obtained from the measured profile of the p-staircase.
Figure 3.4: Calibration curve obtained from the measured profile of the n-staircase.

Figure 3.5: Experimental and theoretical calibration curves of the p- and n-type staircase samples.
Figure 3.5, calculated with the use of SCaMsim [64], show a qualitative agreement with the measurements. The quantification software SCaMsim predicts the SCM signal of various doping levels basing on a set of parameters that have to be estimated from the real features of the tip and the sample, and from the experimental conditions (tip size and shape, oxide thickness, voltage biases). The model assumed in SCaMsim is ideal, thus it does not account neither for interface states, nor for oxide charges, which are always present in the reality. The effects of non-idealities are described in chapter 2.

### 3.2 Unipolar samples of arbitrary doping profile

The simplest procedure to extract the carrier distribution (and the dopant profile) from SCM measurements is the direct inversion method. It delivers the conversion curve between $dC/dV$ signals and the corresponding dopant level. This is achieved in two different ways. The first procedure bases on the use of experimental calibration samples (e.g. epitaxial samples), which are prepared and measured together with the sample under investigation. This provides an experimental calibration curve extending over five up to six dopant decades. The second procedure bases on dedicated SCM simulators (e.g. SCaMsim) to calculate the $dC/dV$ response of homogeneous doped samples and to compute the conversion curve. Combined approaches are also possible.

Figure 3.6 shows the simulated SCM response of an abrupt dopant transition from $1\times10^{18}$ down to $1\times10^{16}$ cm$^{-3}$ (located at 0 $\mu$m) in conjunction with the distribution of the carriers at a bias $V_{DC}$=0 V. It can be observed that the 10% to 90% width ($W_{10-90}$) of the simulated carrier distribution is in the 30 nm range, while the $W_{10-90}$ of the simulated SCM signal is much wider, i.e. in the 80 nm range.

The response of the SCM is not linear. In fact, the broadening of the SCM curve in the low-doped side (right) is more relevant than on the high-doped side. This effect is due to two concurrent causes, which are correlated. The first cause refers to the tail of the carrier distribution that is larger on the right side of Figure 3.6 because of
the lower doping level. The second cause is related to the interaction region of the SCM probe. The size of the interaction region is given by the width of the depleted region underneath the tip. Of course, the lower the local carriers density, the larger the resulting depletion region. In Figure 3.6 the extension of the interaction zone in the highly and in the lowly doped quasi neutral regions are represented in scale by the dashed area. In the low-doped region the diameter of the interaction zone is about 0.3 μm, such that the probe starts to detect the presence of a gradient in the carrier distribution at a distance of about 0.15 μm.

![Figure 3.6: Simulated carriers distribution (dashed line) and dC/dV response (solid line) in an abrupt n⁺n sample (1 \(10^{18}\) and 1 \(10^{16}\) cm\(^{-3}\), respectively, oxide thickness \(t_{ox}=3\) nm, \(V_{DC}=0\) V, tip radius 20 nm). The dashed regions represent the interaction range of the probe in the quasi-neutral on both sides.](image)

The total SCM signal is the result of an averaging of the free carriers concentration over the whole interaction volume. Due to the size of the interaction zone, the tip diameter (black dot within the dashed surface of Figure 3.6) has just a minor role. The same principle holds for the highly-doped side. In this case, the extension of the interaction zone is much smaller than before (about 0.04 μm). In
this case the interaction zone has the same order of magnitude of the
tip diameter, such that the tip size plays a very relevant role in the
definition of the round off of the transition edge. In summary, three
elements are concurring to define the lateral resolution of SCM in
the case of arbitrary doped unipolar samples: the interaction zone (in
quasi neutral regions), the carrier spilling (in proximity of high dopant
gradients), and the probe diameter (close to the edges of highly doped
structures). All components interact in a complex way, such that the
SCM response is not linear and the resulting broadening effect cannot
be simply described by a transfer function associated to the probe.
Thus, the only solution to estimate the achievable resolution is by
means of simulation.

3.3 Three-dimensional dependence of the
SCM signal

The simulation of the probe-sample system is not straightforward from
different points of view. One among the main issues is related to the
mesh density of the model. This strategy enables to perform accurate
3D simulations of a whole SCM measurement within reasonable com-
putational times. While simulating a SCM measurement the mesh
has to take into account, at the same time, of effects arising on some
nanometer thickness (e.g. inversion layer) and of more macroscopic
regions (e.g. the depletion volume), without leading to the explosion
of the number of nodes.

This problem can be kept under control by the use of adaptive
mesh schemes defined by concentric shell-shaped boundaries. 3D mod-
els are necessary if quantitative simulations are required. Figure 3.7
shows the comparison between a 1D, a 2D, and a 3D simulation of
a SCM measurement on a sample with a homogeneous dopant con-
centration of $1 \times 10^{16}$ cm$^{-3}$. The first effect that can be observed is
the strong dependency of the amplitude of the dC/dV signal on the
simulation type. The maximum amplitude is reached by the 1D sim-
ulation, while the 3D simulation (with about 60% of the 1D value)
represents the minimum. This behavior is due to the different rele-
vance of the stray fields surrounding the probe. There are no stray
fields in the 1D simulation, while they affect just two out of the four sides of the blade-shaped probe in 2D. In the 3D simulation, the stray fields are generated along the whole perimeter of the probe. Stray fields couple in parallel the main capacitor immediately underneath the probe with smaller capacitors distributed around the probe. When the main capacitor is entering in full depletion, the lateral distributed capacitors (due to the lower electric field) are still completely or partially depleted. Thus the resulting CV characteristic is a mixing of the characteristic of the main capacitor with that of multiple distributed capacitor (with a lower weighting factor).

As shown in Figure 3.8, this effect lowers the slope of the transition from the accumulation to the depletion region. This is also the reason of the apparent shift of the flatband voltage observed in Figure 3.8 when going from 1D to 3D simulations. In real measurements, because of the finite size of the SCM probe, the flatband shift from the ideal value (1D) is due to the long range capacitive coupling that may reach some hundreds of millivolts, especially for low-doped samples. This effect also explains the different $C_{ox}$ and $C_d$ levels observed in the CV characteristics in Figure 3.8. An additional interesting phenomenon

*Figure 3.7: Dependency of the amplitude of the simulated $dC/dV$ on the simulation condition (all flat contacts with the same nominal area, p-type doping $1 \times 10^{16}$ cm$^{-3}$).*
3.3 Three-dimensional dependence of the SCM signal

Figure 3.8: Dependency of $C_{ox}$ and $C_d$ on the different simulation conditions. The 3D curve shows a slight bend up of the curve in the depletion region due to the effect of the stray fields (flat contact with constant area, doping $1 \times 10^{16}$ cm$^{-3}$).

Figure 3.9: Width of the maximum vertical depletion region as a function of the doping concentration ($n_i$ intrinsic carriers density, probe radius 20 nm) from 1D, 2D, and 3D simulations.
related to the effects of the stray fields is the bending up of the CV curve in the depletion region (for high bias voltages) observed in 2D and 3D simulations of low doped samples (Figure 3.8). This behavior is due to the lateral propagation of the depleted region far from the physical boundaries of the probe. At high bias levels the depletion region due to the main capacitor saturates, while the weighting factors associated to the distributed capacitors become more important, leading to an increase of $C_d$.

**Effect of the dimension and lateral resolution**

As seen in Figure 3.6, the broadening effect of a SCM curve is also due to the size of the interaction volume of the probe. As soon as the sample is driven in the full band bending regime, the maximum extension of the depletion width can be estimated according to one-side abrupt junction approximation [71]

$$W = \sqrt{\frac{4 \varepsilon_S kT}{q^2}} \sqrt{\frac{\ln(N/n_i)}{N}} = \alpha \left(\frac{\ln(N/n_i)}{N}\right)^\beta$$

(3.1)

where $N$ is the doping concentration and $n_i$ is the intrinsic carrier concentration. Figure 3.9 shows the width $W$ as extracted from 1D, 2D, and 3D simulations. In this case, $W$ is defined as the distance from the SiO$_2$ interface, for which the free carriers distribution reaches 90% of the equilibrium value. It can be observed that for doping levels exceeding $10^{18}$ cm$^{-3}$, $W$ is the same for all three cases. On the contrary, for lower doping concentrations, the 3D-value is smaller than the 1D-value by a factor ranging from 2 to 5. This behavior is quantitatively represented in Table 3.1, which lists the calculated coefficient $\alpha$ and the exponent $\beta$ in Equation 3.1 for all three cases. The 3D value of $W$ depends on the cubic root of the reduced dopant concentration rather than on the square root, as predicted by the theory.

**3.4 Contrast reversal in SCM**

Contrast reversal in scanning capacitance microscopy arised as one of the first issue to be solved when the SCM use became a popular
3.5 pn junctions

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$ [cm]</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory (1D)</td>
<td>826</td>
<td>0.50</td>
</tr>
<tr>
<td>1D - 90%</td>
<td>1650</td>
<td>0.51</td>
</tr>
<tr>
<td>2D - 90%</td>
<td>252</td>
<td>0.47</td>
</tr>
<tr>
<td>3D - 90%</td>
<td>12</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Table 3.1: Parameters of the power law in equation 3.1 as extracted from Figure 3.9 ($n_i$ and $N$ in cm$^{-3}$).

technique few years ago. A non-monotonic behavior has been observed in staircase samples, especially in low doped layers. Many hypotheses did arise to explain this phenomenon. Malberti [72] interpreted it as related to the use of metal coated tips. The wear out of the metal coating with the consequent exposure of the underlying silicon tip, has been shown to result into contrast reversal phenomena. On the contrary, the change in flat band voltage related to different doping concentrations seemed not to be the main cause of the problem. Even measurements on the same sample at the same tip bias also resulted in an intermittent occurrence of contrast reversal. Stephenson [73] and Duhayon [74] explained this effect by the different impact of oxide charges and of interface states on sample with different doping concentrations. Later investigations pointed out the importance of a well defined procedure for the sample preparation. In fact, samples either prepared under uncontrolled conditions have been observed to produce different unwanted effects, among these the contrast reversal. Since reproducible sample preparation procedures have been developed (especially the low temperature oxidation), contrast reversal is no longer a major issue. By using diamond-coated probes in conjunction with the sample preparation process exposed in chapter 2, contrast reversal has never been observed.

3.5 pn junctions

Scanning capacitance microscopy can also be used to locate the electrical junction in pn-junctions. The accuracy of the delineation depends on the doping concentration. High doped $pn$ junctions provide
the best results. In order to study the procedure that leads to the delineation of the electrical junction, it is necessary to introduce the concept of capacitance-voltage spectroscopy.

The study of the doping concentration by means of the analysis of the capacitance-voltage (CV) curves is called capacitance spectroscopy. This technique is implemented in scanning probe systems with dedicated sensors under the name of Scanning Capacitance Spectroscopy (SCS). SCS has been shown in the past years to provide a proper tool for junction delineation. This technique bases on the fact that the theoretical CV curve is symmetrical when the tip is centered exactly on the electrical junction. In real measurements the CV curve is affected by oxide charges and interface states. In this case, one has to search for the most symmetrical CV curve. The condition of symmetry relies on the definition of electrical junction, which is the place where the electrons and holes are present in the same amount. The response of the system is therefore equal for the two carriers types, yielding an accumulation of electrons for positive tip voltages and an accumulation of holes for negative tip voltages (both produce a maximum capacitance $C_{ox}$). Assuming a theoretical model, for $V_{DC}=0$ V the CV curve has an absolute minimum $C_d$. These assumptions are validated with a simulation of the SCS apparatus on a low-doped junction (already described in chapter 2). This kind of samples represents the worst case to study because of the wide depletion region causing a longer interaction with the tip and a series of CV curves that are very similar to each other.

### 3.5.1 Simulation model and results

The simulation model used for SCM is represented in Figure 3.10. The bias applied to the tip consists of the DC bias and of the AC high-frequency sensing signal. The two-dimensional simulations of the SCM system are performed by DESSIS. The whole set of simulation is executed in the GENESISE environment. The tip is fixed in the center of the model and the scan action is realized by moving the doping distribution along the horizontal direction. The model represents a $10 \times 10 \ \mu$m square silicon domain with a 3 nm ultra thin oxide on the top and a grounded backside contact. The tip is modeled as an ohmic boundary condition shaped as a truncated cone. The effective contact
3.5 pn junctions

**Figure 3.10:** Model showing the dedicated mesh to optimize the accuracy and the convergence behavior (based on Ciampolini [34]).

diameter (on the oxide) is 40 nm.

The tip is immersed in air to take into account the effects of the fringing fields. In order to optimize the numerical accuracy and the convergence of the results, the tip is kept fixed in the center of the structure. DESSIS solves the Poisson’s equation coupled to the drift-diffusion equation over the whole domain. Successively the small signal capacitance of the tip-substrate system is computed.

The simulation produces a series of CV curves at each scanning point across the low-doped junction [75]. According to the shape of the CV curves, five different regions of interest are identified and represented in Figure 3.11. Each region width is listed in Table 3.2 for a flat and a truncated cone tip. The five cases, which are observed, are summarized in Figures 3.12-3.15. Regions A and E are far enough from the metallurgic junction, such that they do not experience any interaction with the depletion region. Therefore, the resulting CV curves in Figure 3.12 are similar to those measured in unipolar samples. In region B (Figure 3.13) the weak overlap between the interaction region of the tip with the depletion region of the pn junction
**Figure 3.11:** Low-doped sample subdivided in five regions according to the SCS simulation results.

<table>
<thead>
<tr>
<th>Region</th>
<th>Data</th>
<th>Interval [nm]</th>
<th>Width [nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Flat</td>
<td>Trunc.</td>
</tr>
<tr>
<td>A</td>
<td>sim</td>
<td>$[-\infty, -920]$</td>
<td>$[-\infty, 1210]$</td>
</tr>
<tr>
<td></td>
<td>exp</td>
<td>[not relevant]</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>sim</td>
<td>$[-920, 220]$</td>
<td>$[-1210, 410]$</td>
</tr>
<tr>
<td></td>
<td>exp</td>
<td>$[-800\pm100, 200\pm100]$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>exp</td>
<td>$[600\pm200, 0\pm200]$</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>sim</td>
<td>$[-260, -140]$</td>
<td>$[-280, -190]$</td>
</tr>
<tr>
<td></td>
<td>exp</td>
<td>cannot be resolved</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.2:** Classification of the different regions defined in Figure 3.11 and their width.
causes the formation of an inversion layer that is not completed due to the lack of carriers. Thus, the CV curves in the inversion region start bending upwards, though not reaching the same capacitance value as in accumulation ($C_{ox}$). This results into an asymmetric V-shaped curve. In region C (Figure 3.14), the distance of the probe from the metallurgic junction is much shorter than the interaction zone, such that it completely enters the depletion region of the $pn$ junction, and a complete inversion layer can be formed. In this region, V-shaped CV curves are observed in quasi-static CV measurements. In the case of region D in Figure 3.15, the interaction zone of the probe is located in such a way that neither the $n$, nor the $p$ character of the semiconductor can be resolved. This mixed state results in a V-shaped CV curve with a flat bottom, where it cannot be distinguished between inversion and accumulation. In our case, due to the asymmetry of the doping ($n$-type concentration higher than $p$-type concentration), region D is not placed symmetrically across the metallurgic junction, but it is slightly shifted towards the $p$ region, as it is shown in Figure 3.11. This condition defines the interval where the electrical junction is located. The symmetric characteristics obtained within region D indicate that the concentration of holes and electrons within the interaction zone is almost equal.

The width of region D can be assumed as the theoretical lateral resolution of the technique for the delineation of the electrical junction. The results obtained by simulation are validated by comparison to the experimental SCM measurements. Therefore, the derivative of the simulated CV curves has to be calculated. Such $\Delta CV/\Delta V_{AC}$ curves are represented in Table 3.3.

The analysis of the experimental data reveals four regions, only. This is due to the reduced sensitivity of the SCM detector. The maximum theoretical lateral resolution of the technique has been defined basing on the width of region D derived from simulated data. Since region D cannot be resolved in the SCM measurements, the maximum experimental lateral resolution is defined by the boundaries of region C. This value of about 200 nm is very large. Therefore, it cannot be assumed as a resolution indicator. Hence, it can be concluded that the analysis of the $dC/dV$ vs. $V$ curves ($dC/dV-V$) is not very suitable to delineate the electrical junction.
Figure 3.12: Simulated CV curves in the regions A and E.

Figure 3.13: Simulated CV curves in region B.
Figure 3.14: Simulated CV curves in region C.

Figure 3.15: Simulated CV curves in region D.
Highly-doped epitaxial sample

The response of highly-doped samples has been investigated by 2D simulation of an asymmetrically doped epitaxial junction (boron 1 \(10^{18}\), arsenic 5 \(10^{18}\) cm\(^{-3}\)) by assuming a parabolic probe (bending radius: 20 nm) to model the shape of a virgin probe, and a truncated probe (radius of the contacting surface: 10 nm) to model a worn probe. In this case, the interaction volume of the tip consists of a region in the close vicinity of the probe tip and by a larger depleted outer ring induced by the effect of the fringing fields. Due to these circumstances new effects arise. As soon as the probe approaches the junction, the depletion capacitance in the CV characteristic increases (partial inversion), while the accumulation capacitance decreases and does not reach the \(C_{ox}\) value anymore. Despite the reduced amplitude, the CV curve becomes symmetrical in proximity of the electrical junction, like in the case of the low-doped sample. It has to be mentioned that the region C is not observed here. In fact, the observed CV curves are either symmetrical or asymmetrical V-shaped. The CV curves are symmetrical in a very narrow region located within the \(p\)-side, 10 nm away from the metallurgical junction. Outside this interval the transition to asymmetric CV curves is abrupt. Basing on the previous discussion, it can be concluded that in this case, the electrical junction can be located with an accuracy of 10 nm. The simulation with a truncated probe shows that the resolution is not affected by the flattening of the probe.

3.5.2 Electrical junction delineation methods

So far, three methods have been applied in the SCM community to detect the electrical junction position:

- location of the locus within the depletion region where the \(dC/dV\) versus depth (\(dC/dV\)-x) signal is zero (node of the curve),
- determination of the most symmetrical CV curve with SCS,
- measurements of beveled samples.

The first method is limited by the fact that the location of the node of \(dC/dV\)-x is strongly affected by the unavoidable shift of the curve
Table 3.3: Measured (bottom row) vs. simulated 2D (top row) \( \frac{dC}{dV} - V \) curves in the different regions for the low-doped epitaxial pn junction.
at different biases. In the ideal case, the electrical junction is located by the node of the $dC/dV$-x curve measured at such a bias $V_{ZF}$, condition for which the electric field in the oxide is zero (zero-field condition). The zero-field voltage $V_{ZF}$ does not match exactly the flat band voltage $V_{FB}$. The concept of $V_{FB}$ applies only to the case of one-dimensional homogeneous doped unipolar samples. In a real case, the electrical junction is located within the depleted region of the $pn$ junction. According to the usual procedures, $V_{ZF}$ cannot be extracted from independent measurements. Therefore, it is often assumed arbitrarily that $V_{ZF}=0$ V. Investigating in more detail how the $dC/dV$ signal changes as the tip intercepts the electrical junction, it is possible to explain how the junction shifts at different biases. Figure 3.16 represents the simulated $dC/dV$-V curves at four random locations with the tip in close vicinity of the electrical junction, in the $p$-doped region (solid) and in the $n$-doped (dashed).

All curves exhibit a double peak. Thus each curve carries both the positive and the negative sign. At a fixed tip voltage, the four curves
3.5 pn junctions

Figure 3.17: 2D simulated $dC/dV$ vs depth curves.

define the $dC/dV$ ($X_1$) acquired at a particular tip position. Hence, after setting the bias voltage, it is possible to investigate how the $dC/dV$ ($X_1$) curve changes its sign within a given spatial interval. The only voltage interval, where the $dC/dV$-$V$ signal changes from positive to negative (when the semiconductor changes from the $p$-type to the $n$-type), is included between the vertical lines drawn in Figure 3.16. Outside this interval the SCM ($X_1$) response of the sample is either purely positive, or purely negative, i.e. it never crosses the $x$-axis. This means that outside the depicted voltage interval the $dC/dV$-$x$ signal cannot be used to detect the presence of a junction. When working within the interval of interest, the position of the node of the $X_1$ signal strongly depends on the net bias applied to the sample. This is the tip bias corrected by the $V_{ZF}$ shift due to the non-ideality factors of the sample. Non-idealities are for instance interface states, which cause at the same time a doping- and a bias-dependent shift and a distortion of the $dC/dV$-$V$ signal. In Figure 3.17, the position of the node is simulated for the sample II2 ($n^{++}p$). A bias $V_{DC}=0$ V and a $\Delta V_{ZF}$ in the range between -400 and 400 mV have been
assumed. It can be observed that the location of the node moves from 0.51 \( \mu \text{m} \) for \( \Delta V_{ZF} = -400 \text{ mV} \), to 1.35 \( \mu \text{m} \) for \( \Delta V_{ZF} = 400 \text{ mV} \). This indicates that the observed discrepancy between the theoretical and the measured electrical junction can be attributed to a \( V_{ZF} \) shift.

The second method, which bases on the use of SCS, has the inconvenience to require a dedicated experimental set up to acquire, in a reasonable time the CV curves. Its accuracy depends strongly both on the sensitivity and on the noise level of the detector.

Beveled samples can also be used as an alternative, even if they are impacted by the effects of the charge spilling and of the junction punch-through. The advantages to use beveled samples are discussed later on the chapter on the base of device simulations.

The procedure proposed in the following combines the ease of use of the dC/dV analysis, with the immunity of the SCS against the bias related shift and the presence of non-idealities in the ultra thin oxide layer. The starting point is the study of the rate of the dC/dV-x curve while the probe is scanned at different biases, in the close vicinity of the electrical junction. In particular, it is shown that the nodes of the dC/dV-x curves plotted as a function of the probe bias exhibit an accumulation point around the electrical junction position at the zero-field condition (Figure 3.18).

### 3.5.3 The new procedure

The proposed procedure [76] consists of two phases. In the first, the dC/dV-x signal is acquired across the junction at different tip biases. In the second, the positions of the nodes P of the dC/dV-x curves are plotted as a function of the applied bias V. As it will be shown later, as soon as the probe approaches the location where the CV curves exhibit the best symmetry (electrical junction position in the ideal case), the PV curve shows a decrease in the slope (the P points accumulate). Once the tip leaves the location of best symmetry the slope increases again. Therefore, the procedure aims to locate the 2D interval in the PV space, where the nodes accumulate. The centroid \( (V_0,P_0) \) of this interval defines the location of the electrical junction \( (P_0) \) and the \( V_{ZF} \) \( (V_0) \), respectively.

The physics behind this procedure is based on the analysis of the
3.5 pn junctions

Figure 3.18: Correlation between the regions of the PV curve and the related CV curves in different regimes. The nodes of the dC/dV-x curves accumulate in region D.

evolution of the shape of the CV curve when the probe is scanned across the electrical junction. Particularly in region D, in spite of the probe averaging effect, the carrier distribution changes very fast the larger the distance from the electrical junction. Scanning the probe exactly across the electrical junction location, the V-shaped CV curves keep the same degree of symmetry at each point. The curves differ from each other by their minimum capacitance value that is liable to large changes. In fact, the concave shape of the bottom of the CV curves represents a trade-off between the $C_d$ due to the most concentrated carriers (less deep) and that of the less concentrated carriers (deeper). This equilibrium situation strongly depends on the probe location. As a consequence, for a small spatial deviation from the electrical junction, a large increment (decrement) of the bias voltage has to be applied in order to reach the minimum of this concavity (node of $dC/dV$-x curves). Thus, within an interval corresponding to the size of the interaction zone, it is expected a strong dependence of $P_0$ over $V_0$ as shown in Figure 3.18. The radius (cm) of the hemispherical interaction zone depends on the doping $N$
(cm$^{-3}$) and on the density of the intrinsic carriers $n_i$ according to [77]

$$L = 12 \left( \frac{1}{N} \ln \left( \frac{n_i}{N} \right) \right)^{0.39}. \quad (3.2)$$

For a doping interval from $10^{15}$ up to $10^{18}$ cm$^{-3}$, L ranges from 100 down to 10 nm. Once the electrical junction leaves the interaction zone of the probe, the CV curves resume their unipolar behavior, keeping their V-shape. In this case, the CV curves are no longer symmetric and their bottom clearly shows whether the sample is swept from the accumulation to the inversion or vice versa. In this regime their minimum is always located at the voltage where the inversion sets on, i.e. basically the local threshold voltage (which depends on the doping only). Thus, it is observed just a weak dependence of $P_0$ on $V_0$. As soon as the distance of the probe from the electrical junction increases, the transport of minority carriers through the junction becomes more and more difficult. The set on of the inversion requires bias voltages to be applied that are higher than the local threshold voltage. This explains the increasing dependence of $P_0$ on $V_0$.

The new procedure is applied experimentally to cross-sectioned and beveled samples. The resulting data are illustrated in the following section.

**Cross-sectioned samples**

In this case, the layers to be profiled are considered to be parallel to the probe axis. The samples characterized here are the $n^{++}p$ (I12) and $p^{++}n$ (I10) junctions already profiled with SIMS in chapter 2. The location of the electrical junction at the equilibrium is extracted by device simulation of the doping profile acquired by SIMS. The exact location of the electrical junction is defined by the crossing point of the electrons and of the holes distributions. This yields 0.78 µm for the $n^{++}p$ and 1.06 µm for the $p^{++}n$ sample.

The carrier distribution, the width of the depleted region, and the simulated dC/dV-x curve at a probe bias of 0 V are represented in Figures 3.19 and 3.20. It is worth to notice that both SCM simulations yield exactly the electrical junction depth as calculated theoretically for each sample. Figure 3.21 shows the SCM measurements on the
3.5 pn junctions

Figure 3.19: Simulated SCM signal (circles) and free carriers distribution of the I12 n++p junction.

Figure 3.20: Simulated SCM signal (circles) and free carriers distribution of the I10 p++n junction.
$n^{++}p$ sample. Each curve is measured at different biases in order to identify the nodes of dC/dV-x. Figure 3.22 represents three different PV curves, as measured by the author and by two other independent laboratories (the samples have been prepared with different cross-sectioning and oxidation procedures). All curves exhibit an interval, where the nodes of the dC/dV-x curves accumulate. Furthermore, for increasing probe voltages, the nodes of dC/dV-x is shifted towards the p-type side of the junction. That is, for a positive probe voltage (above the zero-field condition), the depth of the EJ is underestimated for $n$-substrates, while it is overestimated for $p$-substrates [78].

![Figure 3.21: dC/dV-x plot of the $n^{++}p$ sample measured at different probe biases. The accumulation of the nodes is indicated by the circle.](image)

The accumulation interval of the nodes is delimited by a rectangle and $(V_S, P_S)$ is located by its centroid. These criteria applied to the curves in Figure 3.22 (top), yield the data summarized in Table 3.4.

The measurements at ETH and IMETEM laboratories deliver the same depth of the apparent junction. This represents an overestimation of the value obtained by simulation (0.78 μm). The experimental $V_{ZF}$ in the case of the best approximation of the electrical junction
Figure 3.22: Measured PV plots of the n++p (top) and p++n (bottom) samples. Three different measures are shown, acquired in three independent laboratories.
Table 3.4: Results of the new method applied to the I12 sample.

<table>
<thead>
<tr>
<th></th>
<th>$V_S$ [V]</th>
<th>$P_S$ [µm]</th>
<th>EJ at $V=0$ V [µm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETH</td>
<td>0.3</td>
<td>1.17</td>
<td>1.13</td>
</tr>
<tr>
<td>IMETEM</td>
<td>-0.13</td>
<td>1.17</td>
<td>1.24</td>
</tr>
<tr>
<td>IMEC</td>
<td>0</td>
<td>1.4</td>
<td>1.09</td>
</tr>
</tbody>
</table>

(Curve ETH) is positive, while in case of the IMETEM curve is slightly negative. The same procedure applied to the curve in Figure 3.22 (bottom) delivers the data in Table 3.5. In this case, all measurements underestimate the theoretical depth of the electrical junction computed by simulation (1.06 µm), while the experimental $V_{ZF}$ ranges from 100 mV to 300 mV (instead of 0V). Figure 3.23 represents the simulation of the PV plot for both the $n^{++}p$ and $p^{++}n$ samples.

Table 3.5: Results of the new method applied to the I10 sample.

<table>
<thead>
<tr>
<th></th>
<th>$V_S$ [V]</th>
<th>$P_S$ [µm]</th>
<th>EJ at $V=0$ V [µm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETH</td>
<td>0.2</td>
<td>0.73</td>
<td>0.64</td>
</tr>
<tr>
<td>IMETEM</td>
<td>0.1</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>IMEC</td>
<td>0.18</td>
<td>0.93</td>
<td>0.93</td>
</tr>
</tbody>
</table>

The behavior of these curves confirms the theory exposed in the previous section, and also enables to extract the same values of the depth of EJ and of $V_{ZF}$ (0 V) as obtained by different simulation approaches (cross-point of the electrons and holes distribution, simulation of the dC/dV-x curve). Figure 3.24 shows the simulation of the PV plots for the low-doped and for a high doped epitaxial junction. As expected from the previous theoretical considerations, an accumulation interval is observed only for the low-doped sample.

This is due to the fact that the interaction zone of the probe is very small in the high-doped sample, i.e. in the 10 nm range. Thus, there is no interval in the PV plot, where the nodes of dC/dV-x accumulate. In this case, the electrical junction can be assumed to be located at the point of inflection of the PV curve.

Experimental data may show deviations from the expected elec-
Figure 3.23: Simulated PV plots of the $n^{++}p$ (I12) and $p^{++}n$ (I10) sample. The electrical junction and the $V_{ZF}$ are located by the centroid of the rectangles.

Figure 3.24: Simulated PV plots of the epitaxial low- and high-doped samples. The accumulation of the nodes can be observed only for the low-doped sample.
trical junction that can range from 20 to 40% of the theoretical value. In the experiment, these discrepancies cannot be corrected just by adjusting the probe bias voltage to realize the zero-field condition. This indicates that the expected symmetry of the measured CV curves is strongly affected by the presence of interface states resulting into a non-linear distortion. This assumption is also supported by the fact that the observed discrepancies strongly depend on the sample preparation procedure applied by the different laboratories. The fact that an accumulation interval of the nodes of dC/dV-x curves is observed experimentally, just means that at this location the CV curves reach the maximum degree of symmetry also including the distortion effect due to the interface state.

The new procedure has some advantages over SCS. Basing on the acquisition of the dC/dV-x signal only, it does not require the use neither of high-speed hardware to collect the local CV curves nor data processing algorithms to extract quantitatively their degree of symmetry. The analysis of the PV curves can be performed automatically off-line providing also important information about the quality of the ultrathin oxide and about the noise level of the SCM sensor. Moreover, this procedure enables to replace an arbitrary junction location criteria with a well-defined methodology based on the physics of the signal generation.

**Beveled samples**

One way to reduce the effect of $\Delta V_{ZF}$ in the junction delineation, is the use of samples beveled at very sharp aperture angles (typically $5.73^\circ$, i.e. at a magnification factor of 10) [79]. The principle behind this technique is that the uncertainty in the delineation of the electrical junction due to $\Delta V_{ZF}$ (few tenths of a micrometer) is the same in cross-sectioned and in beveled samples. Thus, when the distances measured in beveled samples are scaled back to the cross-section case, the uncertainty in the delineation of the electrical junction is divided by the geometrical magnification factor, leading to an error of just some hundreds of Angstroms. Beveling of samples results in the carriers spilling effect, which produces the characteristic distortion of the junction close to the surface, and thus to a systematic error in the location of the electrical junction. Moreover, additional detrimental
Figure 3.25: Tip scanning on the bevel of the I12 n++ p sample in the vicinity of the metallurgical junction and CV curves simulated at each position of the tip, respectively at distance from junction 0 μm, 1 μm, 2 μm, 5 μm.
effects are observed in SCM measurements of beveled junctions. This is the case of the $n^{++}p$ junction where the low concentration of the $p$ substrate and the slow tail of the $n^{++}$ side result in an interaction region of the SCM tip that is in the 0.3 $\mu$m range.

Figure 3.26: Simulated electrical junction position and SCM measurements of the beveled H12 ($n^{++}p$) from two different laboratories.

Thus, in a wide interval on the left side of the metallurgic junction, the interaction region is much wider than the thickness of the wedge (punch-through). This yields a SCM signal showing an apparent doping-type inversion of the $n$-silicon, because of the injection of carriers from the substrate across the metallurgic junction. The width of this interval, which depends on several parameters as the probe bias, the aperture angle, and the dopant concentration, can be easily estimated by simulation of the SCM signal.

In the case of the $n^{++}p$ sample prepared with a beveling angle of 5.73° and biased at 0 V, the SCM simulation also demonstrates that the tip leaves this interval without intercepting the electrical junction.
3.5 pn junctions

<table>
<thead>
<tr>
<th>EJ cross-section</th>
</tr>
</thead>
<tbody>
<tr>
<td>(simulation) 0.78 μm</td>
</tr>
<tr>
<td>(measurements) 1.17 μm</td>
</tr>
</tbody>
</table>

Table 3.6: Experimental and simulated depth of the electrical junction in the I12 sample. The discrepancy is due to $\Delta V_{FB}$.

<table>
<thead>
<tr>
<th>Bevel demagnified</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ (simulation) 0.65 μm</td>
</tr>
<tr>
<td>$\Delta x$ carriers spilling 0.2 μm</td>
</tr>
<tr>
<td>EJ (simulation) 0.45 μm</td>
</tr>
<tr>
<td>(measurements) 0.40-0.48 μm</td>
</tr>
</tbody>
</table>

Table 3.7: Experimental and simulated data of the I12 ($n^{++}p$) beveled sample (MJ metallurgical junction, EJ electrical junction). The agreement between simulation and measurements is excellent. The discrepancies are due to the carrier spilling effect.

In fact, as demonstrated by further simulations, the electrical junction is shifted by the carriers spilling effect in the $n^{++}$ region at a distance of 2.0 μm from the metallurgical junction. This effect is shown in Figure 3.25. Since the magnification factor of the bevel is 10, this shift implies an underestimation of the depth of the electrical junction by 0.2 μm. Assuming the metallurgical junction position at 0.65 μm from the edge (as computed in chapter 2), this results in an apparent electrical junction depth of about 0.45 μm. It should be noted that the simulated depth of the electrical junction is in excellent agreement with the values measured in two independent laboratories (both 0.45 μm, Figure 3.26) [4]. Thus, it can be concluded that the junction delineation by means of SCM measurements on beveled samples is as accurate as the procedure used to correct the junction shift caused by the carrier spilling effect. The results are summarized in Table 3.6 and 3.7.
3.6 NPN bipolar transistor

The microscopy capabilities of SCM have been assessed by imaging real devices and by determining the lateral resolution that can be achieved. In this case the sample under study is the vertical npn bipolar transistor, already described in chapter 2. The SCM measurement shows the variation of the local free carrier concentration. The more intense is the signal (bright regions), the lower is the doping concentration. Darker regions represent both higher doping concentrations and space charge zones.

![Figure 3.27: Overall SCM image of the bipolar transistor showing the amplitude (top) and the phase (bottom).](image)

The probe used for this measurements is a metal-coated cantilever with a tip radius of about 20 nm, as it is the case for most of the metal probes. Scanning over abrupt edges has been avoided. In fact, the presence of different materials with different hardness than the silicon easily results into a premature wear out of the probe. First step the device of interest is located in the chip among the several
other different transistors. Due to the limitation in the scanning area, this operation requires some time. The transistor under investigation is imaged in Figure 3.27 in amplitude (dC/dV) and phase (X1). The structure of the transistor is quite similar to the layout provided by the manufacturer and shown in chapter 2. Measurements has been carried out at V_{DC}=0 V and V_{AC}=3 V (peak-to-peak). The best contrast is achieved under this bias condition. As depicted in Figure 3.28, the measurement reveals the thin dark line separating collector and base, which delineates the location of the space charge at the junction.

3.7 IGBT characterization

Scanning capacitance microscopy is also used as a microscopy tool in order to extract technology parameters of interest. In this second example three IGBT transistors are investigated. In this case, the low-temperature oxidation process had to be optimized to avoid possible artifacts due to the redeposition of the front and backside etched metallization. For this reason the step of the removal of the native oxide by hydrofluoric acid has been suppressed. Due to the large area to be scanned, this measurement is performed with a diamond-coated probe. Once the devices are located on the micro sectioned chip, a first characterization becomes possible. It consists of the definition of the transistor structure and the doping type. Amplitude and phase of
Figure 3.29: Phase images of the three IGBTs showing the different shape of the p-body and of the source diffusions.
the SCM signal are acquired to obtain the information about the concentration and the type of dopants, respectively. Figure 3.29 shows

![Image](image1.png)

**Figure 3.30:** Detail of the channel region of the three IGBTs in $dC/dV$ phase mode. The device #3 is represented also in amplitude mode.

the phase signal map. The coding is dark for $p$-type and bright for $n$-type regions. In the color scale, the more intense is the signal (bright or dark regions), the less doped is the semiconductor. The similar layout of IGBT #1 and #2 indicates that they are produced by the same manufacturer. On the contrary, IGBT #3 exhibits a completely different design. A zoom in the area of interest reveals several details of interest (Figure 3.30). The resolution of the image in Figure 3.30
is high enough to recognize the differences in the $n^{++}$ source region of IGBT #1 #2 and #3. In fact, they belong to a technology split. This information is evident both in amplitude and in phase mode.

3.8 Summary

SCM has been demonstrated to have good dopant profiling and electrical junction delineation capabilities in spite of the strong sensitivity of the technique on sample non-idealities, as fixed charges and interface states. For dopant profiling applications, the signal provided by the SCM detector can be calibrated by dedicated structures and/or by simulation data. Moreover, the use of simulation has been proposed as a valid support to the measurements. Scanning capacitance microscopy applied to image small-features in semiconductor devices has been shown to be very efficient, especially on extended structures, like IGBTs.
Chapter 4

Scanning Spreading Resistance Microscopy

This Chapter provides a systematic study of the characteristics related to SSRM, with particular emphasis on the differences between $p$- and $n$-type semiconductor. Where no distinction is necessary, the results are presented as a general outcome. The tips used are commercial boron-doped diamond-coated probes especially intended for SSRM. They are characterized by a short and stiff cantilever and a spring constant ranging from 20 to 80 N/m. As specified in the manufacturer datasheet, the bending radius of the tip is on the order of 100 nm, what is a typical value for diamond coated cantilevers.

4.1 Calibration of the $p$- and $n$-type staircase

The characterization of $p$- and $n$-type staircases by SSRM involves a thorough investigation of the dependence of the measured resistance of each epitaxial layer on the applied voltage and load. A typical calibration curve obtained with these data relates the doping concentration to the resistivity. The nominal profile of the $p$- and the $n$-type staircase calibration samples is represented in chapter 2. A prelimi-
nary investigation of the $p$-doped staircase showed that the resistance profile strongly depends on the applied bias, especially for the low doped layers.

### 4.1.1 p-type samples and voltage dependence

Figure 4.1 shows the raw resistance profile (before being converted in Ohm) of the $p$-staircase for a tip bias ranging from -0.2 V up to -1.28 V, in steps of about 200 mV. For positive tip biases, the measured resistance values do not increase monotonically with the decreasing doping. Under this bias condition, the profile exhibits a contrast reversal effect that is explained in detail later on this chapter. In order to process the measured resistance profile, it is important to understand how the measured resistance value is rendered by the SSRM detector.

![Figure 4.1: SSRM measurements on a p-type staircase sample at different negative bias.](image)

As already mentioned in chapter 2, the core circuitry of the detector is a logarithmic amplifier, which delivers an output signal $V_{out}$, according to the coding scheme in Figure 4.2. As it can be observed,
4.1 Calibration of the p- and n-type staircases

\[ V_{\text{out}} \text{ vs. } V_{\text{bias}} \text{ graph according to the logarithmic amplifier operating curve.} \]

\[ V_{\text{out}} \text{ is not a unique function neither of the doping type, nor of the doping concentration. In fact, it depends on the sign of the applied bias } V_{\text{bias}}. \] In other words, a current entering the probe \( (V_{\text{bias}} < 0) \) yields

\[ V_{\text{out}} = -\log \left( \frac{I \times 10^6}{V_{\text{bias}}} \right) \] \hspace{1cm} (4.1)

therefore

\[ I = V_{\text{bias}} \left( \frac{10^{\text{out}}}{10^6} \right) \] \hspace{1cm} (4.2)

and the resistance

\[ R = \left( \frac{10^6}{10^{-V_{\text{out}}}} \right). \] \hspace{1cm} (4.3)

Thus, if \( V_{\text{out}} \) is positive, then the measured resistance is above the reference value of 1 M\( \Omega \). If \( V_{\text{out}} \) is negative, the related resistance value is below 1 M\( \Omega \). The amplifier is calibrated in such a way that one decade variation of the doping concentration yields 1 V change
in \( V_{out} \) within an operating range extending from \( 10^4 - 10^{11} \) Ω. Vice versa, for a current exiting the tip (\( V_{bias} > 0 \))

\[
V_{out} = \log \left( \frac{I \times 10^6}{V_{bias}} \right)
\]  

(4.4)

the current is then

\[
I = V_{bias} \left( \frac{10V_{out}}{10^6} \right)
\]  

(4.5)

and the resistance

\[
R = \left( \frac{10^6}{10V_{out}} \right).
\]  

(4.6)

Then, if \( V_{out} \) is positive the resistance is below 1 MΩ and if \( V_{out} \) is negative the resistance is above 1 MΩ.

According to this scheme, the \( V_{out} \) profiles in Figure 4.1 can be immediately converted into resistance values. Because of the negative \( V_{bias} \), the working condition of the detector is described by the third and fourth quadrant of the operating area in Figure 4.2. The dependence of the signal on the doping is treated separately in the following sections.

<table>
<thead>
<tr>
<th>Doping Conc. ([\text{cm}^{-3}])</th>
<th>Resistivity ([\Omega \text{ cm}])</th>
<th>Resistance=(\rho/4a) ([\Omega])</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5 E+19</td>
<td>1.49 E-3</td>
<td>3.73 E+3</td>
</tr>
<tr>
<td>1.3 E+19</td>
<td>6.78 E-3</td>
<td>1.70 E+4</td>
</tr>
<tr>
<td>3.3 E+18</td>
<td>1.89 E-2</td>
<td>4.73 E+4</td>
</tr>
<tr>
<td>1.4 E+17</td>
<td>1.17 E-1</td>
<td>4.28 E+5</td>
</tr>
<tr>
<td>3.0 E+16</td>
<td>5.70 E-1</td>
<td>1.43 E+6</td>
</tr>
<tr>
<td>3.8 E+15</td>
<td>3.62 E0</td>
<td>9.5 E+6</td>
</tr>
</tbody>
</table>

**Table 4.1:** Doping concentrations, resistivity and calculated spreading resistance of the p-type staircase for a tip radius \( a=1 \text{ nm} \).

**Low doping concentrations**

At low tip bias (\( V_{bias}=-192 \text{ mV} \)) the profile contrast is very high, even saturating to positive values at positions corresponding to the lowest dopings.
4.1 Calibration of the p- and n-type staircases

Figure 4.3: IV curves as measured for two layers of low doping concentrations. The deflection setpoint is 6 V.

Figure 4.4: IV curves as measured for the first five layers of the p-staircase.
According to Figure 4.1, this means that a very high resistance value is measured. This does not correspond to the expected spreading resistance, as indicated in Table 4.1. Thus, it indicates that low tip biases are not suitable to quantify low doping concentrations. The IV characteristics of two low-doped epitaxial layers plotted in Figure 4.3, show that the conduction between diamond coated tip and silicon sample is governed by Schottky transport. The series resistance of the heterojunction at biases around 0 V is much higher than the spreading resistance value expected for the given geometry. Besides this unwanted behavior at low tip biases, all the IV characteristics are strongly non-linear as shown in Figure 4.4. The measured current is not proportional to the applied voltage, as one would expect if the spreading resistance were the dominating component.

**High doping concentrations**

In Figure 4.1, very little contrast is observed for the high doped layers, therefore for certain biases they can be hardly distinguished. In this case, one has to take into account further resistance components in addition to the bare spreading resistance. First because the small resistivity associated with high dopings leads to small spreading resistance values (few hundreds of Ω). Second, the saturating behavior observed in Figure 4.5, indicates the presence of a dominating stray resistance component, which does not depend on the sample doping.

In Figure 4.5, the measurements for $V_{bias}=-1238$ mV has been fitted assuming a tip radius of 1 nm. Although this is in accordance with the nanometer spatial resolution of SSRM demonstrated elsewhere [57] [58] [80], this is a value, which highly underestimates the real tip radius of the diamond coated tip used here. Even in this case, the measured curve deviates from the expected behavior for the highest doped epitaxial layers saturating to a constant level. To describe such a behavior in the highly doped region, the contact resistance, already defined in chapter 2, is written in a more compact form

$$R_{cTotal} = R_c + R_m = \frac{\rho_c + \rho_m}{\pi a^2}. \quad (4.7)$$

When $R_{cTotal} = R_{sp}$

$$\frac{\rho_c + \rho_m}{\pi a^2} = \frac{\rho_s}{4a} \quad (4.8)$$
4.1 Calibration of the p- and n-type staircases

and the surface resistivity $\rho_{cTotal} = \rho_c + \rho_m$ is calculated as

$$\rho_{cTotal} = \frac{\pi a}{4} \rho_s.$$  \hfill(4.9)

Figure 4.5: Calibration curves obtained with the data in Figure 4.1, at $V_{bias} = -1283 \text{ mV}$, $V_{bias} = -497 \text{ mV}$, and $V_{bias} = -192 \text{ mV}$ (squares, triangles, and diamonds, resp). The black circles locate the theoretical calibration curve as calculated from $R = \rho/4a$, with $a=1 \text{ nm}$.

According to Equation 4.7, a larger tip radius could help in reducing the contact resistance, and, as consequence, to extend the sensitivity range of the SSRM towards higher doping concentrations. The effect of the contact resistance seems to be more or less present at each doping concentration. A more detailed investigation can be made in order to evaluate, at least qualitatively, the meaning of $R_{cTotal}$ on the overall resistance measurement. In Figure 4.5 a saturation is observed at the last point of calibration, corresponding to the seventh step of the staircase. The influence of the two higher doped buffers seems to be crucial at this position since the width of this step is too narrow due to a strong lateral diffusion of the carriers. The resistance value is affected by a non-negligible local deviation of the carrier concentration from the equilibrium value.
4.1.2 Qualitative evaluation of the series resistance

The non-linear behavior of the experimental calibration curves (Figure 4.5) suggests a dominating contribution given by a variable series resistance, so far attributed to the total contact resistance (Equation 4.7). Nevertheless, the resistance extracted by the system is simply the ratio $R = V/I$, which is an accurate estimate of the real value for linear IV characteristics, only. However, as described in Figure 4.6 the evaluation of resistances on non-linear IV curves (in our case Schottky-like) might bring to large differences whether $R$ is calculated as an instantaneous or a differential value.

![Figure 4.6: Plot of a current-voltage curve (solid line), the resistance calculated as $R = V'/I'$ returns the inverse of the slope of the dotted line A, whereas the differential resistance $R = \partial V/\partial I$ is related to the tangent line B.](image)

The resistance calculated as

$$R = \frac{V'}{I'}$$  \hspace{1cm} (4.10)

considers the system as an ohmic load and the resulting characteristic is a straight line (A in Figure 4.6). The resistance obtained by Equa-
4.1 Calibration of the p- and n-type staircases

4.10 overestimates the actual value. A more accurate resistance value should be instead calculated as

\[ R = \frac{\partial V}{\partial I} \left( V = V' \right). \]  

(4.11)

It represents the tangent line to the IV curve at the bias \( V' \).

The measurement system is therefore affected by a systematic error due to the assumption of a too simplistic definition of the resistance, leading to an overestimation of the measured parameter. This error is larger for low doping concentrations and in those regions of the IV curve that deviate more from the linear behavior (i.e. for bias voltages higher than the Schottky barrier). On the contrary, the contribution of this systematic error is negligible for doping concentrations higher than \( 10^{18} \text{ cm}^{-3} \), since in this case the IV curves are almost ohmic. If the latter resistive parasitic contributions are suppressed, the experimental calibration curves \( R(\rho) \) in Figure 4.5 are expected to decrease in amplitude. The lower the overall resistance values the more the spreading resistance is likely to be measured. Thus, the experimental calibration curve tends to be closer to a theoretical calibration curve computed for realistic tip sizes (>1 nm).

In the case of a pure dependence of \( R \) on the spreading resistance, the slope of an ideal calibration curve \( R(\rho) \) is 1, represented in a log-log plot, while the intercept is the logarithm of four times the tip radius. The deviation from the ideal behavior can be expressed by the ideality factor \( \varepsilon \) in the general expression

\[ R = \frac{k\rho^\varepsilon}{4a}. \]  

(4.12)

Where \( k \) is a generic constant and \( \varepsilon \) is the slope in a log-log scale. The theoretical \( R(\rho) \) calibration curve has slope \( \varepsilon = 1 \). In fact, applying the logarithm to Equation 4.12, it becomes

\[ \log R = \log(\rho^\varepsilon) + \log \left( \frac{k}{4a} \right) = \varepsilon \log \rho + \log \left( \frac{k}{4a} \right). \]  

(4.13)

The experimental curves in Figure 4.5 exhibit a slope \( \varepsilon < 1 \) in the low and in the high doping range, and a slope \( \varepsilon > 1 \) in the intermediate
doping range of the staircase sample. For $\varepsilon < 1$, the sensitivity to local changes of resistivity is lower than expected from the theory. Thus, a constant parasitic resistance is in series with the spreading resistance. An exponent $\varepsilon >> 1$ indicates that the dependence on the doping is much stronger than expected for pure spreading resistance.

In summary, at a given voltage, the parasitic resistance consists of two contributions: the contact resistance and the systematic reading error of the instrument. The stray resistance also depends on the local doping concentration and, under given circumstances, its value can exceed the spreading resistance. This would lead to a mixed global resistance, which is still doping dependent, although it is no longer correlated to the local resistivity by the ideal relation $R(\rho)$. This concept is further explained in the next section upon the analysis of the measures at different probe loads.

### 4.1.3 Force calibration and load dependence

The load is applied to the probe through the piezoelectric unit, which controls the vertical position of the cantilever after the tip/surface engagement. The parameter delivered by the instrument is the so-called deflection set point, which indicates how much the tip is pushed against the surface of the sample. In order to evaluate the force exerted by the probe on the sample, the system has to be calibrated at least for each set of featured probes. Figure 4.7 represents the force calibration curve for $p$- and $n$-type low-doped silicon, related to the probes used in this work for all the SSRM measurements. It has to be noticed that the statistical spread of the spring constants of the cantilever is quite large, since it is in the 20-80 N/m range. This fact limits the reproducibility of the force when replacing the cantilever. The dependence of the measured SSRM signal on different loads is represented in Figure 4.8 for the $p$-staircase. The experiment shows that the contact resistance is reduced down to reasonable values for forces in the order of 20-40 $\mu$N. This represents a good trade-off to minimize the friction with the sample, while still leading to the formation of the high conducting $\beta$-tin. This turns also into the lowering of the electrostatic barrier between tip and sample.

The application of a low force (e.g. 3.3 $\mu$N) results in a very high local resistivity, which exceeds by at least two orders of magnitude or
4.1 Calibration of the p- and n-type staircases

Figure 4.7: Force calibration plot for the p- and n-type silicon assuming the same value for the spring constant (50 N/m).

Figure 4.8: P-type staircase profiles measured at different probe loads and voltages.
more the value expected for the spreading resistance. This points out that the contact resistance dominates at such low forces and rules the total resistance. Moreover, the contact resistance strongly depends on the doping concentrations. Thus, the technique provides a signal that is modulated by the local doping concentration. However, it is not related to the spreading resistance (Figure 4.8). A further increase of the load (from 11.3 to 34.1 μN) evidences how the output profile signal decreases toward lower values, first for high doping concentrations and then for low doping concentrations. In the practical use, the set point is increased until the signal settles down to an almost force-independent value over the whole doping concentration range (e.g. at a force of 22.7 and 34.1 μN the two profiles are almost coincident). At this point the voltage bias is increased from -600 mV to -1062 mV. The system provides a lower \( V_{out} \) profile, thus a smaller contact resistance. The voltage is not increased any further to avoid to damage the tip.

The resistance profile, measured under these conditions, can be considered as a combination of a spreading resistance signal and of a component of the contact resistance. This latter cannot be completely eliminated. This situation leads, nonetheless, to a strong dependence of the signal on the doping concentration. The quantification of the SSRM should therefore include, if possible, the additional effects of the series resistances.

### 4.1.4 n-type calibration sample

The doping profile of the \( n \)-type calibration sample is shown in chapter 2. Table 4.2 summarizes the basic sample characteristics, i.e. the impurity concentration of every epitaxial layer, the theoretical resistivity, and the theoretical spreading resistance calculated for a tip radius of 1 nm. The IV characteristics measured at different locations of the \( n \)-type staircase sample are very similar to the curve measured for the \( p \)-type sample. Hence, also in this case, the highly doped diamond exhibits a non-negligible conduction both for positive and for negative voltages. The IV curves for low doping show a Schottky characteristic with a positive threshold voltage (Figure 4.9). At high doping concentrations the IV curves show a conducting behavior for positive and negative voltages with a very low threshold. The series resistance of the Schottky diode is extracted from the slope of the IV
4.1 Calibration of the p- and n-type staircases

characteristics for biases beyond the threshold voltage. This is also the case of the spreading resistance.

<table>
<thead>
<tr>
<th>Doping Conc. [cm⁻³]</th>
<th>Resistivity [Ω cm]</th>
<th>Resistance=ρ/4a [Ω]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4 E+19</td>
<td>4.12 E-3</td>
<td>1.03 E+4</td>
</tr>
<tr>
<td>1.2 E+18</td>
<td>2.17 E-2</td>
<td>5.43 E+4</td>
</tr>
<tr>
<td>1.0 E+17</td>
<td>8.64 E-2</td>
<td>2.16 E+5</td>
</tr>
<tr>
<td>5.0 E+15</td>
<td>9.81 E-1</td>
<td>2.45 E+6</td>
</tr>
<tr>
<td>4.2 E+14 (substrate)</td>
<td>1.05 E1</td>
<td>2.63 E+7</td>
</tr>
</tbody>
</table>

Table 4.2: Doping concentrations, resistivity and calculated spreading resistance of the n-type staircase for a tip radius a=1 nm.

The similar behavior observed for impurity concentration p- and n-type, allows to define a similar series resistance to be accounted to the theoretical spreading resistance. In Figure 4.10 the staircase profile is measured at different tip loads, the largest variation is again observed for the low doped layers. In order to see accurately the

Figure 4.9: IV characteristics at each epitaxial layer of the n-type staircase.
Figure 4.10: Resistance profile of the n-type staircase at different probe loads and at constant bias.

Figure 4.11: Resistance profile of the n-type staircase at different biases, for a constant tip load of 34.1 μN.
contrast between substrate and the last layer, a load of about 15 μN (Figure 4.7) has to be applied to the probe. Higher forces could cause an earlier wear out of the tip. At lower loads the lowest doped layers cannot be defined properly. The successive step is to set the bias working point. In Figure 4.11 the profile is acquired at different tip voltages. The response of the low-doped layers is more sensitive to the bias changes. At a bias of 1200 mV, the resistance decreases by a factor of two. Figure 4.12 shows the comparison of the theoretical calibration curve for the spreading resistance (tip radius of 1 nm) with the calibration curves as extracted from the profiles of Figure 4.11.

Figure 4.12: Theoretical calibration curve for the spreading resistance (black circles, tip radius 1 nm) and experimental calibration curves (white symbols) for a tip voltage of 500 mV (white circles), 700 mV (white diamonds), 900 mV (white triangles), and 1200 mV (white squares). The measured resistance is approximated by V/I.

The local slope of the calibration curves is less than that expected for the theoretical curve over the whole resistivity range. In particular, the higher is the bias the lower is the dependence of the resistance on the doping concentration. As shown in Figure 4.13, this trend is even
more evident when the differential resistance \( \partial V/\partial I \) is represented as a function of the resistivity. Furthermore, it has to be noticed that for the low-doped epitaxial layers, at voltages largely above the Schottky threshold, the measured resistance values appear less doping dependent. This effect can be explained by the fact that the bias dependent component of the stray resistance clearly dominates over the contribution expected from the spreading resistance. The extrapolation of

![Graph showing theoretical calibration curves compared to measured instantaneous and differential experimental resistances calculated as V/I (white circles) and as ∂V/∂I (white squares).](image)

**Figure 4.13:** Theoretical calibration curves compared to the measured instantaneous and differential experimental resistances calculated as V/I (white circles) and as ∂V/∂I (white squares).

the tip radius from the measurements on low-doped epitaxial layers, yields a tip radius \( a \) that exceeds by far any realistic value (e.g. \( a=10 \) nm). This confirms in addition that the doping dependence of the measured resistance is no longer related to the spreading resistance contribution. On the contrary, under these experimental conditions, the response of the SSRM on the doping concentration seems to depend dominantly on the doping dependence of the contact resistance.
4.2 Quantification of experimental results

The electrical model of the SSRM has been introduced in chapter 2 under the assumption that all the contributions due to the parasitic effects are negligible compared to the spreading resistance. Nevertheless, the experimental evidence demonstrated that these contributions can be on the same order of magnitude, or they can even exceed the spreading resistance.

\[ R = \frac{\rho}{4a} \]  

\( (4.14) \)

cannot be used for the accurate conversion of the resistance into re-
sitivity values. The more general equation

\[ R = \frac{\rho}{4a} + R_{cTotal} \] (4.15)

which takes into consideration the contact resistance, delivers more realistic estimates of the measured resistance. Since \( R_{cTotal} \) cannot be deduced accurately from theoretical considerations, several experimental procedures based on correction factors or iterative loops have been proposed in the past to minimize the quantification error. These procedures are mainly based on experimental calibration curves, similar to those plotted in Figure 4.5. Even using these procedures, the repeatability of the results still remains an issue. The calibration has to be carried out for every measurement keeping unchanged the operating conditions. The scarce repeatability of the SSRM is demonstrated in Figure 4.14, where two calibration curves corresponding to different sets of measurements acquired under the same experimental conditions.

Figure 4.15: Reproducibility test performed on the n-type staircase in two different laboratories. Comparison between two measurements of IMEC (white circles) and ETH (white squares) and two theoretical calibration curves for tip radii \( a=1 \) nm and \( a=10 \) nm.
conditions are shown. In both cases, the sample has been prepared according to the procedure described in chapter 2, the tips used are from the same lot, and the same pressure and bias have been applied. The comparison of the experimental curves with the theoretical calibration curves for a tip radius of 1 and 10 nm, evidences the large discrepancies over the whole doping concentration range.

**Figure 4.16:** Reproducibility test performed on the p-type staircase sample in two different laboratories. Comparison between two measurements of IMEC (white triangles) and ETH (white circles first measure, white squares second measure) and two theoretical calibration curves for tip radii $a=1$ nm (black circles) and $a=10$ nm (black squares).

The result of a survey made recently in a different laboratory (IMEC, Leuven, Belgium) [4] is represented in Figure 4.15 for the n-type staircase. This plot is a good indication of the reproducibility of the SSRM measurements carried out in different environments, without standard measurement conditions. It has to be noticed that the bias applied for the measurement performed at ETH (1.2 V) has been chosen above the Schottky threshold measured from the IV curves.

The results of the same reproducibility test performed on the p-
type staircase sample is shown in Figure 4.16. The ETH curve has been obtained at \( V = -1 \) V. For biases below the Schottky threshold, the resistance variations with the doping concentration become also larger. This effect is more evident for the low-doped layers and it is due to the increase of the impact of the contact resistance. The main consequence is the stretching out of the measured curves over the resistance axis. In summary, the quantification of SSRM data has been observed to depend strongly on the applied load, on the bias, on the roughness of the sample, on the probe wear out, on the quality of the mechanical contact, on the formation of the \( \beta - \text{tin} \) phase, and possibly on the density of the interface states \([54]\). All these factors produce a severe deviation of the measured curves from the theoretical behavior expected by the spreading resistance.

A semi-quantitative study of these spurious effects (especially at low tip bias) can be performed taking into account the specific contact resistance \([71]\), which is defined as

\[
\rho_c \equiv \left( \frac{\partial J}{\partial V} \right)_{(V=0)}^{-1}
\]

where \( J \) is the current density. The contact resistance \( R_c \) is obtained by

\[
R_c = \frac{\rho_c}{A} \tag{4.16}
\]

where \( A \) is the tip contact area. Under the assumption of a circular contact of radius \( a \), the contact area is described as \( A = \pi a^2 \).

Figure 4.17 shows the comparison of the contact resistance \( R_c \) with the spreading resistance \( R_{sp} \) for an aluminum contact with a tip radius \( a = 10 \) nm on a \( n \)-type silicon sample with a doping concentration ranging from \( 10^{14} \) to \( 10^{20} \) cm\(^{-3} \). In the whole doping concentration range, \( R_c \) exceeds \( R_{sp} \) by at least four orders of magnitude. As expected the deviation is much larger and poorly dependent on the doping in the range from \( 10^{14} \) to \( 10^{17} \) cm\(^{-3} \). This situation represents the case of an ideal mechanical contact (the contacting surface are perfectly flat and conformal, \( R_m = 0 \) V) with no pressure applied. For smaller contact radii the discrepancy is even larger. \( R_c \) is calculated at a bias \( V = 0 \) V, and corresponds to the maximum contribution to the voltage dependent component of the total contact resistance \( R_{c\text{Total}} \). In SSRM
4.2 Quantification of experimental results

![Graph showing theoretical contact vs. spreading resistance for a tip radius $a=10$ nm, aluminum contact on n-type silicon.]

**Figure 4.17:** Theoretical contact vs. spreading resistance for a tip radius $a=10$ nm, aluminum contact on n-type silicon.

measurements, the presence of the pressure-dependent $\beta-tin$ phase, and a mechanical resistance $R_m \neq 0$, produces respectively

- a decrease of $R_c$ that is particularly large for low doping concentrations due to the onset of lower barriers and different conduction mechanisms.

- an offset resistance $R_m$ depending on the quality of the mechanical contact between tip and sample surface.

Therefore

$$R_{cTotal} = R_c + R_m$$

(4.17)

where $R_c$ slightly depends on the doping concentration and $R_m$ accounts for a slight shift of the curve towards higher resistance values. Considering the bias-dependent components of $R_{cTotal}$, one can observe that it decreases non-linearly as the applied bias increases. Therefore, under the usual working conditions of the SSRM, the discrepancy between $R_{cTotal}$ and the spreading resistance is expected to be smaller. However, the contact resistance component still remains dominant.
This produces the non-linear behavior of the experimental calibration curves as a function of the doping concentration. Thus, in case of a doping concentration lower than $10^{18} \text{ cm}^{-3}$, $R_{cTotal}$ represents the upper limit of the measured resistance. On the contrary, at high doping concentrations, the contact is almost ohmic and the spreading resistance itself just accounts for few hundreds of ohms. In this case, the lower limit of the measured resistance is represented by the resistance of the diamond-coated tip, by the components related to the mechanical contact, and by the backside contact ($R_{probe}$, $R_{bc}$, $R_{m}$), which accounts for an overall resistance in the kΩ range. This effect explains the saturation of the measured resistance value at high doping concentrations.

![Resistance-voltage characteristic of the diamond coated probe on gold.](image)

**Figure 4.18:** Resistance-voltage characteristic of the diamond coated probe on gold.

In order to quantify this contribution, the probe resistance $R_{probe}$ is measured in soft contact using the SSRM system on an aluminum substrate first etched and then sputtered with 30 nm gold. The variation of the differential resistance measured applying a voltage ramp to the tip is plotted in Figure 4.18. The resistance of the probe reaches a minimum value of about 1.5 kΩ for a bias of $V=\pm 2$ V, while the max-
4.3 Contrast reversal

The intrinsic resistance of diamond-coated probes limits the range of applicability of the SSRM to doping concentrations that are not too high. As an example, for a reasonable tip radius of 10 nm and a tip resistance of 2 kΩ, the minimum resistivity that can be resolved is about $10^{-2}$ Ω cm. This value is in very good agreement with the experimental observations reported in this work. Even though a better sensitivity towards high doping concentrations has been reported for full diamond probes [83], the reliability of absolute measurements is still questionable due to the difficulty to estimate the contribution of the back contact resistance $R_{bc}$.

4.3 Contrast reversal

In the previous measurements, the bias has been chosen to operate the tip-sample Schottky contact in forward mode. In that case, the SSRM current at a given tip bias has been observed to increase monotonously with the increasing doping concentration, as predicted by the theory for Schottky contacts. If the tip/sample Schottky contact is operated in reverse bias, a non-monotonic response of the SSRM system as a function of the doping concentration is observed. This phenomenon of contrast reversal is shown in Figure 4.19, in case of a reverse tip bias (positive) applied to the $p$-type staircase. In Figure 4.19, the high-doped layers are not affected by the change in the tip bias polarity and still exhibit the regular behavior. On the contrary, the SSRM resistance measured for the low-doped layers decreases with decreasing doping concentration. This anomalous behavior at reverse bias is also reproduced by the IV characteristics of the epitaxial layers represented in Figure 4.4. The monotonic response of the system with decreasing doping concentration is observed for the first four steps, while the last three steps yield the highest current levels. This dependence is evidenced in Figure 4.20, where the reverse bias characteristics of the lower-doped layer shows a stronger voltage dependence than the layers with higher doping concentration.

The explanation of the observed doping dependence of the SSRM signal is not obvious and one should go back to the modeling of the carrier transport mechanism occurring in heterojunctions (highly-doped
Figure 4.19: Resistance profile of the p-type staircase for a reverse bias of 500 mV showing contrast reversal.

Figure 4.20: Current-voltage curves for the p-type staircase in reverse bias voltage at low doping concentration.
diamond probe on a silicon sample) in reverse bias operation. In the case of moderate doping concentrations in the $10^{17}$ cm$^{-3}$ range, the observed reverse current seems to be reasonably described by a reverse-biased Schottky contact. On the contrary, for lower doping concentrations, the contact does not longer exhibit any blocking state and the reverse current depends on the applied bias. The lower the doping concentration, the lower the observed resistance. Multiple current transport mechanisms may be involved in this regime showing a strong dependence on the degree of ideality of the Schottky contact (including the presence of the $\beta$–tin phase, traps, and interface states) and on possible surface contaminations (e.g. alkali ions) due to the sample preparation procedure.

\begin{figure}[h]
    \centering
    \includegraphics[width=0.5\textwidth]{resistance_profile.png}
    \caption{Resistance profile of the n-type staircase for a reverse bias voltage $V=-500$ mV showing contrast reversal.}
    \label{fig:resistance_profile}
\end{figure}

The n-type sample in Figure 4.21 exhibits a behavior similar to the p-type sample. However, in this case, the contrast reversal occurs at a lower doping concentration than for the p-type sample, since it has just been observed for the substrate. The turning points in the calibration curves of the p-type and n-type samples are shown in Figure 4.22. It must be noticed that the interpretation and the quantification
of SSRM profiles can be also affected by contrast reversal effects. This problem might arise for example in a \( pn \) junction, where one or both sides are low doped.

### 4.4 \( pn \) junctions

\( pn \) junctions are characterized on the base of the previous knowledge acquired by the investigation of unipolar staircase samples. The interpretation of the results is carried out by device simulation. First, the results are used to calculate the distribution of the injected current and the depletion zone. Second, they are studied in order to understand how the technique interacts with the sample and how it modifies the electrical properties of the surface under study. A thorough description of the simulation model and the obtained results is presented in the next section.
4.4 pn junctions

Results for I10 junction

The quantitative characterization of a \textit{pn} junction and the junction delineation impose a difficult challenge to the SSRM. In fact, since a positive voltage represents a forward bias for the \textit{n}-type and the negative bias is forward bias for the \textit{p}-type silicon, it is almost impossible to select a tip voltage that enables the tip Schottky contact to be operated in forward polarization at the same time on both regions. Furthermore, in presence of large doping concentration gradients the contact of the tip with the sample changes from an ohmic contact for high doping concentrations to a Schottky contact for the lowest concentrations (especially when approaching the metallurgical junction). These experimental circumstances make the interpretation of the resistance curve not straightforward. The characterization of \textit{pn} junctions includes the evaluation of the carrier concentration on both sides, the delineation of the metallurgical and of the electrical junction, if possible. Referring to the delineation of the electrical junction, one among the main assumptions is that the carrier profile within the depletion region should not be modified that much by the tip bias.

Past publications [54] described that, when the probe crosses the electrical junction, a resistance peak should appear in the resistance profile. This interpretation is based on the assumption that, being the electrical junction the location where the density of free carriers reaches its minimum, the spreading resistance should reach its maximum. The sample considered here is the I10 \textit{p++n} junction, described in chapter 2. In Figure 4.23 two curves are plotted, which represent the resistance profile of the I10 at two opposite tip biases. The curve obtained for the positive bias is different from that acquired at negative bias. The high resistivity region within the first 60 nm is associated with the glue layer between the I10 and the dummy sample. After this region, the resistance decreases very fast and saturates down to the levels related to the very high \textit{p}-type doping. Approaching the depletion layer, the resistance increases and displays a peak before settling down to the resistance value associated with the low-doped \textit{n}-substrate. On the contrary, the curve acquired for the negative tip bias does not show any peak and exhibits a smooth and monotonic transition from the \textit{p++} to the \textit{n}-type region. The mechanism that leads to the formation of the peak observed for the positive tip bias
Figure 4.23: Resistance profile of the I10 p$^{++}$n junction for a positive and a negative tip bias.

is still not completely understood. In fact, if it would be related to the electrical junction, the peak should be located at about 0.8 μm from the sample surface, as predicted by the simulation in chapter 2. Although, it occurs well before the location of the metallurgical junction, within the p$^{++}$ doped side. The experimental observation that the location of the peak does not shift, as a consequence of a tip bias variation of several hundreds of millivolts [84], is a further argument against the supposed correlation between the peak and the electrical junction. Both the theory and the simulations predict the opposite behavior. In fact simulations highlighted that the electrical junction really experiences large shifts, depending on the tip bias and on the tip position. As a final argument it can be mentioned that, surprisingly, the peak does not appear applying a negative bias. The occurrence of the observed peak can be tentatively explained by the presence and by the influence of interface states. They may dramatically impact the conduction mechanisms at the contact between the sample and the tip. Thus, it is very likely that the peak is just produced by an artifact of the tip/sample interaction. All these arguments lead to the
conclusion that the observed peak is not in relation with the electrical junction at the equilibrium.

Results for I12 sample

The same procedure of measurements used for I10 is applied to the measure of I12 $n^{++}p$ samples, which are described in chapter 2. Figure 4.24 shows the resistance profiles for the I12 sample measured at opposite biases, which shows a different behavior. In particular, the profile measured at negative biases, which is forward bias to $p$-type, shows a monotonic behavior. The high resistance region measured from depth 0 μm to about 0.25 μm locates the non-conductive layer of glue between the sample and the protective dummy silicon. Immediately after, the high doped region is characterized by a low resistance, which increases monotonically settling to the resistance values of the low-doped bulk level. At positive biases the profile is again monotonic, but characterized by a larger and less steep transition separating the high-doped region from the low-doped bulk region. No
peak is observed in either the two cases. This behavior is reported also in another laboratory [54], where the disappearance of the peak is described to be due to the inversion layer produced by the interface states. The inversion hinders the electrical junction to cross the cross-sectioned surface, such that it cannot be detected.

The characterization of highly doped samples like I10 or I12 is simplified by the fact that the contact is almost ohmic over a large extent of the profile. On the contrary, more problems arise when imaging low-doped junctions, where the measured resistance is affected by large parasitic components.

Results for low-doped sample

The existence of a peak correlating with electrical junction has been investigated also by means of the low-doped epitaxial junction described in chapter 2. The low doping concentrations of both the $p$- and the $n$-region result into a very large depletion region with a well-defined location of the electrical junction. From an experimental point of view, this sample represents the worst case for the SSRM, since the low doping (in the $10^{15}$ cm$^{-3}$ range) results into high-resistive Schottky contacts and thus the structure is prone to contrast reversal effects. The SSRM profile of this sample acquired at different biases is shown in Figure 4.25. Due to the large Schottky barrier, a bias around 1 V is required to obtain a good contrast. No signal has been detected for negative biases. As already observed in the previous measurements with the staircase samples, and as expected from tip contacts with a dominating Schottky component, the higher is the bias, the lower is the measured resistance. Furthermore, the contrast of the SSRM signal improves after successive acquisitions. The common point of all the measured curves is the location where the resistance starts to increase. This point is located close to the metallurgical junction of the sample.

4.5 Simulation details

The SSRM measurement process is simulated with DESSIS. The optimization and the assessment of the accuracy of the simulation models
4.5 Simulation details

Figure 4.25: Low-doped epitaxial junction resistance profile at different voltages and times.

are achieved by comparison of the numerical data with the results of the analytical solutions in 2D and 3D. The simulation problem is first solved in 2D. This allows investigating the impact of the model size, model boundaries, and of the grid density on the accuracy of the results, keeping under control the computational time. The sample is modeled as a rectangle of silicon. The backside and the tip contacts are considered as two ohmic electrical boundary conditions. This simple structure with a uniform doping concentration is shown in Figure 4.26. The mesh is done with the classical box-method and it is adapted to the physics of the problem. A coarse mesh is defined for the whole rectangle. In a square region below the tip the mesh is refined within a box with a side measuring three times the tip diameter. During the simulation a voltage ramp is applied to the tip. The backside is kept to ground. As a result, a current flows in the silicon model spreading from the tip contact. The IV characteristic for this simple model is linear. The total resistance is calculated as the ratio V/I. The mesh is optimized comparing the accuracy of the simulated resistance with the analytical value. Using the meshing rules listed
in Table 4.3, 99% of the analytical value of the resistance can be obtained within a simulation time of 2 minutes on a Sun Ultra 333 MHz.

<table>
<thead>
<tr>
<th>Placements</th>
<th>MaxSize (x y)</th>
<th>MinSize (x y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>global mesh</td>
<td>L/2</td>
<td>L/5</td>
</tr>
<tr>
<td>refinement under tip</td>
<td>d/15</td>
<td>d/30</td>
</tr>
</tbody>
</table>

Table 4.3: Optimum meshing conditions for the 2D model (L is the width of the rectangle and d is the tip radius).

Typical sizes of the model are on the order of millimeter for the silicon bulk (L, 2×1 mm²), and 30 nm for the tip radius (d). In the case of homogeneous doping concentration, the simulation time is reduced by simulating just half of the model. The conversion of the simulated resistance into the corresponding doping concentration is performed by extracting the resistivity value through the analytical formula and using the Irvin’s curves already described in chapter 2. The analytical expression in 2D is not the same as in 3D. In fact, in the 2D model the conduction and the spreading just occur into a plane and the tip is blade-shaped. The simulator delivers a resistance related to a default thickness of the model of 1 μm. The analytical expression used to validate the simulation and to convert the simulated resistance...
4.5 Simulation details

Figure 4.27: Structure for the Hall’s resistance (solid lines), and entire 2D structure (solid+dashed).

into the resistivity is derived from the Hall’s formula [85]

\[
R = \frac{\rho}{\varepsilon} \left[ N + \frac{2}{\pi} \ln \left( \frac{1}{\cos \left( \frac{a\pi}{2} \right)} \right) \right]
\]  

(4.18)

where \( \rho/\varepsilon \) is the sheet resistance [\( \Omega \)], \( N=h/L \) is the number of vertical squares composing the structure, and \( a=(L-d)/L \). This refers to the geometry drawn in Figure 4.27. Under the simulation conditions reported in Table 4.3, the simulator delivers the analytical resistance value within 2% error.

4.5.1 Current confinement

The confinement of the current through geometry-related causes produces an increase of the measured SSRM resistance. This happens, for instance, when the tip approaches a physical edge or a layer of doping with higher resistivity. The usual expression for the spreading resistance assumes that the current can spread isotropically from the tip into the material. The current flow towards the backside contact is not influenced neither by the proximity of the boundaries of the model nor by the change in the local resistivity. If these conditions are not realized, the current distribution becomes asymmetric and the
apparent local resistance appears to increase. Of course, if the measured resistance value is converted by means of the usual formulas, this results into an overestimation of the local resistivity, i.e. in an underestimation of the local doping concentration. Similarly, if the tip approaches a region with a lower resistivity, the resulting SSRM resistance leads to an overestimation of the local doping concentration. It can be easily demonstrated basing on simple symmetry considerations, that the resistance for a tip located in the center of a 2D model is a factor of two smaller than the resistance for the same tip placed at a boundary of the same simulation model. The same occurs in a 3D model. In 2D, the increase of the SSRM resistance for the tip located in an intermediate location between the center and the boundary can be estimated analytically by the use of the Hall’s formula, Equation 4.18. The simulation model is composed by a rectangle of silicon with arbitrary constant doping concentration, terminated by a rectangle of oxide much larger than the tip. The tip is moved from the center of the model towards the oxide, as shown in Figure 4.28. At each simulation step, the spreading current is stored in a data file and the resistance is calculated. The approximated analytical computation of the confinement effect assumes that the tip can be subdivided in two equal and independent parts. Each part injects current into a prismatic region whose width is $L_{left}$ and $L_{right}$ respectively. The resulting model is represented in Figure 4.29. The tip scan is modeled with an increase of $L_{right}$ and an equal decrease of $L_{left}$, such that $L_{right} + L_{left} = L$. The resistance of the whole asymmetrical system is then computed as the parallel of the resistances

$$R_{left} = \frac{\rho}{\varepsilon} \left[ N_{left} + \frac{2}{\pi} \ln \left( \frac{1}{\cos \left( \frac{a_{left} \pi}{2} \right)} \right) \right]$$ (4.19)

$$R_{right} = \frac{\rho}{\varepsilon} \left[ N_{right} + \frac{2}{\pi} \ln \left( \frac{1}{\cos \left( \frac{a_{right} \pi}{2} \right)} \right) \right]$$ (4.20)

The exact solution, computed by device simulation, is compared in Figure 4.30 to the approximation obtained through the analytical model. Figure 4.30 shows that the analytical formula overestimates the effect of the current confinement. This is due to the fact that the analytical model considers separately the effect of both parts. As shown in Figure 4.30, the discrepancy between the simulation and
Figure 4.28: Current distribution if the probe approaches the edge. The tip size is enlarged to show its location.
Figure 4.29: Model for the current confinement using Hall’s formula.

Figure 4.30: Quantification of the current confinement in 2D by simulation and by the analytical formula. $R_0$ is the resistance with the tip located at the center of the model.
the analytical computation is about a factor of two within the interval $L_{\text{right}}/L_{\text{left}} \leq 4$. As soon as the tip approaches asymptotically the boundary of the simulation model the resistance reaches twice the central value.

The resistance begins to change consistently at 300 μm from the edge, this strong interaction is related to the 2D nature of the model, a much smaller distance is expected in 3D.

### 4.5.2 2D SSRM simulation on silicon with varying doping concentration

The same procedure can be applied with some precautions also to non-homogeneous doping distributions. The structure considered here is an abrupt 1D homojunction formed by two $p$-type epitaxial layers with a doping concentration of $10^{17}$ and $10^{18}$ cm$^{-3}$, respectively. In this case, the symmetry criteria are not applicable and the full model has to be simulated leading to longer computational times.

The tip scanning sequence is realized moving the doping concentration profile. The whole process is controlled by a position parameter entered in GENESISE. The simulation starts with the tip at position 0 μm (right boundary of the model), i.e. within the epitaxial layer doped $10^{18}$ cm$^{-3}$, and ends at 2000 μm on the second epitaxial layer (left boundary of the model). The tip step length is adapted to the local gradient of the carrier concentration. In order to define accurately the transition region, the mesh is refined around the metallurgical junction over a width on the order of the local Debye length. The calculation of the resistance profile in Figure 4.31 by using the Hall’s formula, and the Irvin’s curves, leads to the distribution shown in Figure 4.32. The extracted distribution is much smoother than the carrier distribution simulated by DESSIS. In fact, a continuous decrease of the carrier concentration is observed as the tip approaches the metallurgical junction from the high-doped side as a consequence of the current confinement effect. When the tip approaches the metallurgical junction from the low-doped side, an apparent increase of the carrier concentration is noticed. This is due to the fact that a considerable amount of the current reaches the backside contact through the highly-doped layer. The error, induced by this effect, in the quantification of the resistance profile reaches its maximum at the metallur-
Figure 4.31: (Top) Resistance profile of the homojunction vs. the position of the doping step, on the left (step position between 0 and 1000 μm) the tip is on the higher doped side $1 \times 10^{18}$ cm$^{-3}$, on the right (step position between 1000 and 2000 μm) the tip is on the lower doped side $1 \times 10^{17}$ cm$^{-3}$. (Bottom) Relative error calculated as $R/R_{\text{bulk}}$ of the two sides.
gical junction, as shown in Figure 4.31. The lateral interaction range predicted by the 2D simulation is quite large and may reach several hundreds of micrometers, especially in low-doped layers. It has to be noticed that the interaction range observed on real measurements and in 3D simulations is much shorter. This is due to the blade-shaped tip of the 2D simulation model, which leads to a dramatic overestimate of this effect.

### 4.5.3 2D simulation of SSRM on doping staircases

The procedure described in the previous section is used to simulate also large staircase structures (see chapter 2). The width of the silicon rectangle is 1000 μm, the height is 500 μm, and the tip is represented as an ohmic contact with 30 nm radius, fixed in the center of the sample. The 1D doping profile consists in seven 5 μm-wide epitaxial layers. Each layer is separated from the neighbor by a 1 μm-wide buffer layer, whose doping concentration is at least a factor of two higher than the higher doped layer of the couple. In order to avoid
artifacts due to the current confinement at the boundaries of the simulation model, the widths of the first and of the last epitaxial layers are extended to reach the boundaries of the structure. During two subsequent simulations, the doping profile is shifted horizontally, by a distance corresponding to the scanning step. The carrier profile of

Figure 4.33: Carrier profile of the p-type staircase at the equilibrium as calculated by DESSIS and as extracted from the SSRM simulation.

the p-type staircase sample as extracted by 2D SSRM simulation is compared in Figure 4.33 to the carrier profile at the equilibrium computed by the DESSIS simulation. It is evident that in none of the steps the carrier profile is homogeneous along the width of the epitaxial layers. Furthermore, the carrier density is overestimated in six out of seven layers. These results are due to the tip interaction with layers having higher doping concentration (the buffers and the adjacent layer). The opposite behavior observed in the layer with the highest doping concentration is due to an overestimation of the length of the simulation model in the Hall’s formula. The error in the worst case reaches 300% of the doping concentration. In a real measurement and in 3D simulations, this effect is not as evident as in the 2D simulation of Figure 4.33, since similarly to the case of the homojunction, the
lateral interaction is exacerbated by the blade-shaped tip.

4.5.4 2D simulation of SSRM on pn junctions

In the following the SSRM response of a bipolar diode is investigated by device simulation in order to study both the dependency of the measured curves on the experimental parameters, and the junction delineation capabilities of the technique.

For sake of simplicity the simulations are performed first in 2D, being conscious that, in this case, the lateral interaction of the blade-shaped tip increases the contribution of the bulk resistance in the direction of the tip scan. Furthermore, the contact is assumed as ohmic to reduce the computational time. This represents a good approximation in the case of the occurrence of the $\beta - tin$ phase at the interface with the sample, or in any case, if the transport mechanism is not limited by the contact resistance. It has to be mentioned that the assumption of a Schottky-like tip contact would also provide a quasi-quantitative information. In fact, it does not take into account the large changes in the local barrier, which are due to the fact that the tip is scanned on a silicon sample whose doping concentration ranges over several decades. The simulation model is a rectangular structure reproducing the real size of the sample (2 mm by 1 mm). Such a large simulation domain is necessary to cope with the macroscopic diffusion length of the carriers. The simulated sample is the low-doped epitaxial junction described in chapter 2. The low doping concentration is chosen to produce a wide depletion region, while the asymmetric doping concentration ($n$ side $5 \times 10^{15}$ cm$^{-3}$, $p$ side $10^{15}$ cm$^{-3}$) results into two distinct locations of the metallurgical and of the electrical junction.

The tip is modeled as a flat ohmic boundary condition, whose position is scanned across the junction at constant simulation step. A voltage ramp ranging from -1 to 1 V is applied to the tip at each position to enable the extraction of both the instantaneous and the differential resistance. The ohmic rate of the IV curves, obtained for the bulk, changes to a Schottky-like rate in proximity of the metallurgical junction if some bias and doping conditions occur. The strong effect of the junction built-in potential on the spreading of the current is the cause of this alteration. In order to simplify the discussion,
Figure 4.34: Simulated IV curve for a tip centered at 50 nm from the metallurgical junction in p side.

Figure 4.35: P-type side: current distribution at tip bias 1 V (A)(top-left), 0.4 V (B)(top-right), -0.4 V (C)(bottom-left), and -1 V (D)(bottom-right).
the results obtained for the tip scanning over the \textit{p}-doped side are considered first. Figure 4.34 represents the IV characteristics at different tip locations. When the tip is far away from the metallurgical junction, the IV curve shows the expected linear behavior. As soon the tip approaches the depletion region, the IV curve turns suddenly into a rectifying characteristic. The diode-like characteristic observed in the close vicinity of the metallurgical junction can be explained by considering Figure 4.35, i.e. the current distributions in the \textit{pn} structure in the regimes A, B, C, and D as defined in Figure 4.34. In the regime B, the bias of the tip is not sufficient to completely overcome the potential barrier due to the depletion zone, such that the current injection occurs laterally towards the \textit{p}-doped region. This results into a very large SSRM resistance value, which increases as soon as the tip approaches the metallurgical junction.

In the regime A, the tip voltage is high enough to forward bias the \textit{pn} junction, such that the main component of the injected current flows through the highly-doped \textit{n}-side. This turns into a reduced SSRM resistance value. Finally, in the regime C and D the \textit{pn} junction is reverse-biased, such that the current injection can only occur through the region in the \textit{p}-side, resulting into high values of the SSRM resistance. Figure 4.36 represents the IV characteristics simulated on the \textit{n}-doped side at two different distances from the metallurgical junction. As expected, when the tip is far away from the metallurgical junction the IV curve exhibits a linear behavior with resistance values lower than in the \textit{p}-doped side due to the higher doping concentration. However, as soon as the tip approaches the metallurgical junction, the IV curve exhibits a slight decrease of the slope. This behavior can be explained by considering the different bias regimes in Figure 4.37, as defined in Figure 4.36.

In the regimes A’ and B’, the \textit{pn} junction is reverse-biased such that the current conduction occurs mainly within the higher-doped \textit{n}-side. In these regimes, the resistance value increases slightly when the tip approaches the metallurgical junction.

This is caused by the current confinement effect. When the tip is within the depletion zone, the lateral injection towards the \textit{n}-doped region is much easier than in the regimes C and D. In this case the depletion region is thinner than in the \textit{p}-side due to the higher doping concentration. In the regime C’ the tip bias is still too low to forward
**Figure 4.36:** Simulated IV curve for a tip centered at 50 nm from the metallurgical junction in the n-type side.

**Figure 4.37:** N-type side: current distribution at tip bias 1 V (A')(top-left), 0.4 V (B')(top-right), -0.4 V (C')(bottom-left), and -1 V (D')(bottom-right).
bias the junction, therefore the current spreads mainly towards the 
$n$-doped region and the resistance value is similar as for $A'$ and $B'$. In 
the regime $D'$, as soon as the junction is forward-biased, the current 
starts to flow within the $p$-doped region. However, due to the lower 
doping concentration of the $p$-doped region, the current mainly stays 
confined within the $n$-doped side and the SSRM resistance does not 
decrease. This systematic investigation provides an explanation of the 
resistance profiles simulated at different tip biases.

Figure 4.38 represents the resistance profiles obtained at positive 
tip biases below, close, and above the turn on voltage of the $pn$ junc-
tion shown in Figure 4.34. At tip biases lower than the turn on voltage, 
the resistance profile ranges over one or more order of magnitudes and 
exhibits a sharp peak within the depletion region whose maximum is 
located in the close vicinity of the metallurgical junction.

At tip biases higher than the turn on voltage, the profile does 
not show any peak. The resistance changes monotonously from a 
low value in the $n$-doped side to a higher value in the $p$-doped side. 
Furthermore, it has to be noticed that on the $p$-doped side, away from 
the metallurgical junction, the resistance value is less than expected. 
This is due to the fact that in this regime the highly-doped $n$-side 
still participates to the conduction reducing the overall resistance. Of 
course, this effect is mainly caused by the 2D nature of the simulation 
model and it is not expected to occur in this extent in real 3D samples.

Figure 4.39 shows the resistance profile at different negative tip 
biases. In this case, all curves exhibit a peak. This effect is due to 
the asymmetry in the doping concentration. In fact, when the tip 
is located on the $n$-type side, the $pn$ junction is forward biased. In 
spite of this, the adjacent $p$-doped side does not represent a competi-
tive conduction path, due to the lower doping concentration. On the 
contrary, as soon as the tip reaches the $p$-doped region the $pn$ junc-
tion becomes immediately reverse-biased, such that the only available 
conduction path is the low-doped $p$-side (regimes C and D in Figure 
4.34).

When the tip is located away from the metallurgical junction the 
resistance settles on both sides down to almost the theoretical re-
sistance value. Although this simple 2D simulation has been per-
formed by assuming ideal conditions, it already enables to explain 
semi-quantitatively the observed experimental resistance profiles. Nev-
Figure 4.38: Resistance profile at different positive biases.

Figure 4.39: Resistance profile at different negative biases.
4.5 Simulation details

Nevertheless, real samples can be affected by several non-ideality factors (including fixed oxide charges, interface states, the $\beta-tin$ phase, the Schottky characteristic of the contact, and the tip wear out), which may impact on the response of an SSRM system especially when the tip is scanned across the depletion region.

Junction delineation with SSRM

As already mentioned in chapter 3, the electrical junction definition is delineated by the crossing point of electron and holes distributions. If a current is injected across the junction, the distribution of the free carriers is altered.

The shift of the electrical junction as a function of the bias is shown in this paragraph basing on simulations. Figure 4.40 represents the electron and the hole distribution along the low doped $pn$ junction, respectively. The electrical junction at the equilibrium is located at a depth of 1000.18 $\mu$m from the sample left edge. The application of a

![Figure 4.40: Electrical junction position defined as crossing point of holes and electrons, the vertical line locates the metallurgical junction.](image)

Figure 4.40: Electrical junction position defined as crossing point of holes and electrons, the vertical line locates the metallurgical junction.
Scanning Spreading Resistance Microscopy

(EJ) involves the injection of a current through the scanning probe. The distribution of the free carriers is affected by the tip bias (if it is large or small and/or if it is positive or negative) and by the tip position respect to the metallurgical junction (MJ). The sample under investigation is the low-doped pn junction, with the n-side more doped than the p-side. When the tip is on the p-side, a positive bias means a forward polarization of the junction, whereas a negative bias corresponds to a reverse polarization. It is considered now the case in which the tip is located at 50 nm from the MJ in the p-side, and positive and negative biases are applied. Figure 4.41 shows the shift of the location of the electrical junction at a given tip position and at different biases. Low positive biases causes a shift of EJ that can be considered not local. On the contrary a high positive bias modifies the EJ location locally under the tip. In the first case, at low bias, the region underneath the tip is totally full of electrons (majority carriers) and the EJ is far away from the tip. In the second case, at high bias,
both types of carriers are present underneath the tip, where the holes are majority carriers. The application of a negative bias produces a shift of the electrical junction. The semiconductor underneath the tip is of $p$-type, such that the electrical junction is moved into the $n$-side, which has a higher doping concentration. A negative bias always produces a reverse bias condition at this tip position.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure4.42.png}
\caption{Electrical junction shift at different tip positions: 50 nm from MJ in $p$-type (green), on MJ (red), 50 nm from MJ in $n$-type (black). The tip bias is 1 V.}
\end{figure}

If the tip bias is kept fixed and the tip moved across the $pn$ junction, the location of the electrical junction also changes according to Figure 4.42. Here the tip is located at three different positions, at 50 nm from MJ on the $n$-side, just above the MJ, and at 50 nm from MJ on the $p$-side. A bias of 1 V is applied to the tip, which is supposed to be located first on the $p$-side in a condition of forward polarization. The electrical junction follows the tip movement till it reaches the MJ. Afterwards, as the tip is completely on the $n$-side, the reverse polarization occurs and the EJ is moved away into the $p$-side, far from the MJ. According to these three cases, the EJ position dramatically changes depending on the bias polarity, amplitude, and to
the tip position. Therefore, the location of the electrical junction is not constant.

The concept of electrical junction at the equilibrium loses its meaning once a current is injected into the semiconductor through the probe. In order to demonstrate the behavior mentioned above, the scan of the tip across the location of the electrical junction at the equilibrium has been simulated. The tip is assumed to be at low-bias conditions. The shift of the electrical junction is represented in Figure 4.43 at different tip location across the equilibrium location of the EJ.

\[ \begin{array}{c}
\text{Figure 4.43: Electrical junction line at different tip positions nearby the EJ position at the equilibrium. The tip bias is 0.2 V.} \\
\end{array} \]

Once the tip is biased, the equilibrium situation is destroyed. The location of the electrical junction depends on several parameters, including the polarity of the bias, the position of the tip, and the doping type of the contacted semiconductor. For this reason, the system does not detect anything particular around the position of the EJ at the equilibrium. This happens even at low bias, due to the occurrence of many other factors. The electrical junction only represents the locus where the electron density equals the hole density.
4.5.5 SSRM 3D simulations

The need of 3D simulations for SSRM is imposed by the need of easily quantify the measurements, by direct comparison. The real system is a 3D structure and its resistance is given by the Holm and Störmer equation introduced in chapter 2. As it has been shown 2D simulations do not reflect the reality, such that full 3D simulations are required. The doping profiles used are the same as for the 2D simulations and the results can be directly compared in order to highlight the differences between a 2D and a 3D simulation model. The first application is the simulation of the SSRM measurement of a homogeneously-doped silicon, a cuboids, whose dimensions are $2000 \, \mu m \times 2000 \, \mu m \times 1000 \, \mu m$ (Figure 4.44). The tip is located on the top surface center and has an octagonal shape. The distance between two opposite sides of the octagon is 40 nm. For the generation of 3D Cartesian boundaries and grids, the same tool (MESH) adopted for 2D simulations is used. The octagonal shape of the tip has been chosen since the definition of a circular object could require too many points. The Holm and Störmer formula has been corrected by means of a multiplicative coefficient to be used for the perimeter of the octagon.

Figure 4.44: Full 3D model for SSRM simulations, the tip is enlarged to show its position.
\[ Perimeter_{circle} = Perimeter_{octagon} \]  
\[ 2\pi a = 16 \tan \left( \frac{\pi}{8} \right) a_{oct} \]  
\[ R = \frac{\rho}{4a} = \frac{\rho}{4 \times 1.0547 a_{oct}} \]

where \( a \) is the radius of the equivalent circular tip and \( a_{oct} \) is the half the side-to-side distance of the octagon. The grid is defined basing on the 2D grid with the addition of the third dimension, as in Table 4.4.

<table>
<thead>
<tr>
<th>Placements</th>
<th>MaxSize (x y z)</th>
<th>MinSize (x y z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>global mesh</td>
<td>( x/2 ) ( y/2 ) ( z/2 )</td>
<td>( x/5 ) ( y/5 ) ( z/5 )</td>
</tr>
<tr>
<td>intermediate</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>refinement tip</td>
<td>0.0025</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 4.4: Dimensions and locations of the 2D model box-grid.

The resulting CPU time is on the order of 15 minutes. Although it is much longer than for 2D simulation, it is still acceptable for practical applications. The computational time is longer in the case of a voltage ramp and of a non-homogeneous doping distribution, which usually requires some additional mesh refinements. Since ohmic contacts are used, the simulation of a single bias value is sufficient to determine the resistance of the system. Under these assumptions the discrepancy between the simulated resistances and the values computed by the Holm and Störmer equation is within 4%. The accuracy can be improved by an additional refinement of the mesh.

### 4.5.6 3D SSRM simulation of a doping step

SSRM measurements of 1D doping profiles (\( p \)-type doping step) are simulated by using a Cartesian model. The tip vertex are parameterized and introduced in GENESISE in order to simulate the motion of the tip along the profile. The considered resistance profile is shown in
Figure 4.45. Comparing this profile to the resistance profile obtained by 2D simulation it can be noticed that the transition from the 3D simulation is much steeper than in the 2D case. This indicates that in 3D the interference effect observed in 2D is much less relevant. As usual the range of the interference effect can be expressed by taking as a reference the radius of the tip. In the 3D case this range has been demonstrated to be within few tip radii.

![Resistance profile of the p-type homojunction as extracted from 3D simulation of SSRM.](image)

**Figure 4.45:** Resistance profile of the p-type homojunction as extracted from 3D simulation of SSRM.

The doping profile in Figure 4.46 has been obtained by direct inversion of the simulated resistance values. The conversion from resistance to doping concentration has been done analytically, by means of the conversion curve obtained fitting the Irvin’s curve. The error introduced by this conversion procedure has been estimated in about 20%, up to a maximum of 200% in the close vicinity of the homojunction.

### 4.5.7 3D SSRM simulation of p-staircase

The $p$-staircase has been simulated by the same model used for the homojunction. An additional refinement box has been added to the
**Figure 4.46:** Doping profile of the p-type homojunction as extracted from a 3D simulation of SSRM (white circles) and the theoretical doping profile (black circles).

**Figure 4.47:** Comparison between the doping profile (solid line) and the carrier concentration extracted with SSRM simulation (diamonds).
grid to cover the six steps (about 45 µm) of doping concentrations. In order to save time, the simulation has been performed in the center of each doping step. The agreement of the 3D simulation with the expected doping values is very good, with a maximum error of 21% at the central step of the staircase (Figure 4.47). As already observed in the case of the homojunction, the tip interference effect is much less relevant than for 2D simulation.

4.5.8 3D simulation of SSRM on a pn junction

The 3D simulation of the low-doped pn junction requires a large number of meshing points in the depletion region and a voltage to be ramped at the tip contact. This results in a computational time and effort that is much heavier than in the 2D case or than in any other 3D case dealing with homojunctions. Therefore, the system has been modeled only for crucial tip positions. Figure 4.48 represents a cross-section of the 3D structure, whose boundaries have been already described. The doping profile represents the low-doped junction introduced in chapter 2. The section is obtained performing a cut along the plane of symmetry. The tip position is at 50 nm from the metallurgical junction on the p-side. The current distribution is shown for two different tip biases.

![Figure 4.48: Current distribution in the cross-section of the 3D structure of the low-doped pn junction for a bias of 1 V (left) and -1 V (right).](image)
**Figure 4.49:** IV curve traced with the tip centered at 50 nm from the metallurgical junction on the p-side of the low-doped junction.

**Figure 4.50:** Resistance profile from the 2D and 3D simulations. In the 3D the tip radius is 5 nm, while in the 2D the tip radius is 20 nm. The bias is 0.1 V.
The results of the 2D simulation and the current distribution along the symmetry axis of the 3D simulation are in good qualitative agreement. A quantitative comparison is not possible due to the different geometry conditions. The IV curve simulated at the location shown in Figure 4.49 (p-side) presents a diode-like characteristic. The threshold voltage is lower than the usual 0.6 V and the reverse current is very small. For biases lower than the threshold voltage the n-type region does not participate to the conduction. For voltages above the threshold voltage the conduction takes place in both the p-doped and the n-doped sides.

The resistance profile simulated for this junction is shown in Figure 4.50 for a tip bias of 0.1 V. The 2D and the 3D profiles show a similar shape. The peaks of both structures are nearby the MJ. This is the typical response for a bias below the threshold voltage. The surface conditions considered here are ideal. A ionic contamination has been demonstrated by simulation to produce a shift of the resistance peak away from the MJ.

4.6 SSRM imaging of a bipolar transistor

The two-dimensional imaging capabilities of SSRM are described in this paragraph taking into consideration real devices. The npn bipolar transistor, which has been already described in chapter 2, and that has been successfully characterized with SCM (chapter 3), is investigated. The following images are examples of 2D SSRM measurements on such a structure. Figure 4.51 shows in the left side the deflection map, which evidences the topography of the cross-sectioned device. The topography is visible due to the selective etching action of the colloidal silica during the final step of the sample preparation (polishing). The three tungsten plugs and the metal lines associated with the collector, emitter, and base contacts are clearly visible. This image is acquired at a very low pressure (deflection setpoint 0.6 V) to avoid the smearing of the topography through the scanning tip.

The related $V_{out}$ map of the SSRM image is shown on the right side of Figure 4.51. In spite to the fact that the tip bias is 0 V, the buried layer (collector) can be clearly distinguished from the surrounding silicon substrate. A contrast is also observed to arise in the
Figure 4.51: Deflection and Vout map of the bipolar transistor as measured with the SSRM. The scan size is 6.3 µm.

Figure 4.52: SSRM map of the bipolar transistor. Detail of the region of polysilicon emitter/collector/base region. The scan size is 1.6 µm.
metallization (tungsten plugs and metal lines). This effect is very likely due to the collection through the tip of carriers photogenerated in the junctions either by the laser beam or by the microscope illumination. Figure 4.52 shows the SSRM map of a detail of the polysilicon emitter/base region. The horizontal scan size is 1.6 µm, the tip bias is 0.5 V, and the applied deflection setpoint is 5 V. The color scale from dark to light indicates different resistances, ranging from high to low values, respectively. The base diffusion corresponds to the dark region between the emitter and the collector. This image also evidences the doping modulation within the collector area. Due to the particular tip bias, the different metallurgic junctions are not delineated. Special attention is paid to minimize the load applied to the cantilever. In fact, the application of an excessive pressure while scanning a sample with different hardness could produce a premature wear out of the diamond coated tip, resulting into degradation both of the lateral resolution and of the signal contrast.

4.7 SSRM imaging of DMOS and MOS transistors

The characterization of relevant IC features like the channel length of transistors is one of the main microscopy applications of techniques like SSRM. A DMOS transistor is investigated here by SSRM. The dedicated test structure, shown in Figure 4.53, has been purposely designed to simplify the sample preparation procedure for SSRM and SCM. Two vertical DMOS transistors and two lateral n-MOS transistors with a channel width of 640 µm have been integrated in such a way to constitute a 2D structure that can be easily cross-sectioned. All terminals of the transistors have been grouped along the sides in order to avoid interferences during the sample polishing phase. In this preparation all device terminals have been shortcircuited. All bond pads are coated with a conductive silver paint, the whole surface is sputtered with a gold layer, and finally everything is contacted to the sample holder with conductive glue to realize a backside contact with low resistance. The first measurement is carried out on the first DMOS. The resulting resistance map is shown in Figure 4.54. Dark
colors refer to high resistance while light colors refer to low resistance values. As it can be seen, the technique cannot clearly discriminate between doping types. This makes the interpretation of the SSRM images not straightforward. The contrast which arises between differently doped regions is due to the fact that the tip contact resistance for the same doping level is not the same for $n$- and $p$-doped silicon. Unfortunately, this characteristic does not allow to assess which, between two adjacent differently doped regions, has the higher doping concentration. The polysilicon gate also appears to be high-resistive. This can be attributed to the series resistance due to the geometry of the gate, which is a thin and long isolated structure.

Figure 4.55 shows an overall SSRM image of the $n$-MOS transistor while two details of the source region are represented in Figure 4.56, one at higher magnification. All images are acquired at a tip bias of 500 mV and a very high tip load corresponding to a deflection setpoint of 6 V. The highly-doped source and drain regions can be easily identified, including the lightly-doped areas, which penetrate into the channel region. Surprisingly, the channel area appears in Figure 4.55 as a high-doped region. This effect is no longer observed at higher magnification (Figure 4.56). The bright strip on the top of the source, which is interrupted immediately before reaching the spacer, is the silicided area.
4.7 SSRM imaging of DMOS and MOS transistors

**Figure 4.54:** SSRM map of the channel region of a DMOS, the voltage applied is 1.2 V with a deflection setpoint of 5 V.

**Figure 4.55:** SSRM image of a lateral n-MOS transistor.

**Figure 4.56:** SSRM details of the source in the lateral n-MOS transistor shown in Figure 4.55 (left). Magnification of the lightly doped drain area (right).
4.8 Summary

Scanning Spreading Resistance Microscopy has been evaluated as quantitative tool for dopant profiling and junction delineation of semiconductor devices. The first experimental analysis on calibration structures has revealed the presence of several parasitic resistance components. In fact, the spreading resistance appears to be just one of the various contributions to the total measured SSRM resistance value. Therefore, the use of the simple conversion formula presented in chapter 2 is not straightforward. It follows that, the quantitative use of SSRM for doping profiling requires a thorough calibration by the use of dedicated test structures upon standardized experimental conditions. Moreover, the current injected during SSRM measurements does not allow to use this technique for the delineation of the electrical junction in $pn$ samples [86]. The imaging capabilities of SSRM are satisfying when considering very small structures (few microns). Excellent results can be reached in case of measurements performed with full-diamond cantilevers.
Chapter 5

Conclusions

In this work, Scanning Capacitance (SCM) and Scanning Spreading Resistance Microscopy (SSRM) have been characterized as dopant profiling and imaging tools. The focus of the research has been a comprehensive investigation of the operating limits and of the technical capabilities of both techniques, supported by extensive experimental results. Special attention has been paid to the effect produced by the probe scanning on the surface of the sample, and to its influence on the measured parameters.

The reproducibility and the repeatability of both techniques (especially of SCM) have been obtained with the development of appropriate sample preparation procedures, by the design of dedicated test structures, and by the definition of the optimum working point of the instrumentation. This has been done by searching a trade off between the accuracy of the results and the easiness of use of the techniques. More complex procedures could be envisaged (e.g. for the sample preparation), however they would have a larger impact in terms of processing costs and time and they would require a less user-friendly apparatus.

In general, where the techniques did not deliver quantitative results, the qualitative data provided already represented excellent boundary conditions for the characterization of sub-micron semiconductor devices and structures.
5.1 Dopant profiling application of SCM and SSRM

Present investigations pointed out that SCM is prone to the effects due to sample non-idealities (like fixed charges or interface states), which limit the immediate dopant profiling capabilities of the technique. In spite of this, the signal provided by the SCM detector can be accurately calibrated by dedicated calibration structures and/or by simulation data.

The quantitative doping profiling capabilities of SSRM have been investigated in this work by the same test structures used for SCM. The limits of SSRM as a quantitative dopant profiling tool in terms of lack of reproducibility have been clearly evidenced. In particular, the simple spreading resistance model, which correlates the measured resistance value with the local resistivity has been shown to be very inaccurate. For practical purposes, it needs to be replaced by a more sophisticated model, which takes into account the different components of the measured resistance, as derived from the experimental observations. The simple spreading resistance model has been corrected by the addition of different series resistances, which account for different experimental conditions, as the applied tip bias and the tip pressure.

The analysis of the experimental SSRM data shows that the resistance component due to spreading is only a fraction of the overall measured value. Furthermore, the dependence of the overall resistance on the local resistivity strongly deviates from the linear behavior, which is predicted by the spreading resistance theory and which is postulated in several publications. In addition, the tip radius that has to be known for the conversion, is not a well-defined measurable, since at the usual tip pressures, the formation of highly-conductive silicon phases can be expected to arise, whose lateral extension differs considerably from the physical bending radius of the diamond coated tip. Moreover, the tip radius also depends on the wear out of the probe during scanning. Anyhow, under suitable experimental conditions the SSRM signal depends monotonously on the local doping concentration.

Thus, since the spreading resistance is just one among the compo-
nents of the measured SSRM resistance value, the use of the simple conversion formula $R = \rho/(4a)$ is not straightforward. In fact, the quantitative use of SSRM for doping profiling requires a thorough calibration procedure based both on dedicated test structures and on standardized experimental conditions. This applies in particular to two-dimensional doping profiles and both $p$ and $n$ doping types. In fact, the measured SSRM resistance in 2D or 3D doping profiles depends in a very complex way on the local doping distribution.

5.2 Electrical junction delineation

In this work an original and efficient measurement procedure has been developed to delineate quantitatively the electrical junction by SCM. This new scheme is proposed as a valid alternative for the traditional spectroscopy approach that requires a much more complex experimental apparatus. Advantages and limitations of this procedure have been investigated both by simulation and experimental data.

The impact of the carrier spilling on the accuracy of the junction delineation by SCM on beveled samples has been quantified both experimentally and by simulation. In particular, it has been shown that the geometrical magnification through the bevel has beneficial effects in reducing the impact of the unwanted junction shift, produced by sample non-idealities (e.g. interface states). After applying a model to correct the delineation error introduced by the carrier spilling an excellent accuracy and an excellent reproducibility among different laboratories has been observed. This establishes the most proficient method for quantitative junction delineation.

The capabilities of SSRM to delineate quantitatively the electrical junction have been investigated first by device simulation. 2D and 3D simulations have demonstrated that SSRM behaves differently respect to the Spreading Resistance Profilometry (SRP), since in SSRM no resistance peak is associated with the electrical junction. This explains the large deviations reported in the literature between the expected location of the electrical junction (the minimum of the free carrier concentration at the equilibrium) and the observed SSRM peak. This behavior is mainly due to the fact that the bias applied to the tip during the SSRM measurement strongly interferes with the electrostatic
potential within the semiconductor, resulting into a relevant modification of the local free carrier distribution at the equilibrium. On the contrary, the simulation of ideal $pn$ samples confirmed that a SSRM peak arises, under given experimental conditions, in conjunction with the metallurgical junction. It has also been shown that sample non-idealities, like fixed surface charges (e.g. $10^{11}$ cm$^{-2}$ Na ions), can result into a sensible shift of the SSRM resistance peak. In general, the presence of a peak in a real sample has been observed to depend on several factors, as for example the doping concentration of the $p$ side to the $n$ side, the tip bias and polarity, and the tip position with respect to the metallurgical junction.

5.3 Main differences between SCM and SSRM

From a practical point of view, the main differences between SSRM and SCM are that SSRM does not require the critical step of the low-temperature sample oxidation, and that the tip pressure applied during SSRM measurements is much higher than for SCM.

Thus, the very stringent requirements to prepare a good quality semiconductor/oxide interface do not apply for SSRM, with the considerable advantage to enable the characterization of semiconductors that cannot be easily oxidized. In converse, the major drawback of SSRM is the need to apply high tip pressures with the non-negligible tip wear out problems, arising especially when scanning large surface devices and hard semiconductors. These high levels of pressure also result into a smearing of the sample surface, which strongly reduces the definition of the topography images. In this respect, SCM seems to be a cleaner technique, since it just requires a soft tip contact. This results into an enhanced lifetime of the probes and into a reduced damage of the sample surface. On the contrary, SCM is much more sensitive than SSRM to the effects of charges trapped in the thin oxide, to the interface states, to the oxide quality and to the oxide contaminations.
5.3.1 Probes

In this work, the great advantages related to the use of diamond-coated probes over the traditional metal-coated probes have been confirmed. These are the enhanced tribologic wear out properties of the probe and of the sample due to the particularly hard diamond coating, the particularly long lifetime of a single probe, which enables the acquisition of dozens of large size SCM images, and the resulting reproducibility of the measurements. In converse, the major disadvantages are related to the relevant thickness of the conformal probe diamond coating, which turns into a relatively large bending radius of the tip, i.e. into a reduced lateral resolution.

Due to the high tip pressure involved with SSRM, the use of hard probes (e.g. diamond-coated probes) is mandatory. Especially in the case of soft semiconductors (e.g. GaAs and InP), this results into a considerable indentation that hinders multiple scans over the same sample area. Due to their robustness and the intrinsic sharpness of the vertex, full diamond probes are more suitable when both a long probe lifetime and a high lateral resolution are required. Unfortunately, up to now monolithic diamond probes are limited commercially available. This is the reason that in this work, all SSRM measurements have been carried out with dedicated diamond coated probes with extra stiff cantilevers in order to sustain the necessary pressure levels. Such high pressure levels and the related friction effects were responsible for the damages at the surface of the sample and for the reduced probe lifetime, especially when scanning large area samples. The lifetime of diamond-coated probes for SSRM has been estimated to be about a factor of five shorter than the diamond-coated probes for SCM.

5.4 SCM and SSRM simulation

Device simulation has been used in this work to investigate the basic physical response and the intrinsic limitations of SCM and SSRM. For this reason, higher order effects, which were impossible to be quantified experimentally, have not been included intentionally in the simulation model. This has been the case of the interface states in SCM and the Schottky characteristic of the tip-probe contact in SSRM.
Simulation has been a great support to the measurement; nevertheless, setting up the proper boundary and grid structure for each technique has been a challenging task. Several solutions have been attempted to create mesh grids conformal to the physics of the signal generation of the considered scanning probe techniques. The convergence behavior and the accuracy of the simulations have been optimized and verified on the base of analytical solutions.

In all SCM simulations the tip-to-substrate capacitance has been computed by the small signal approximation, under the assumption that the deep depletion effect was negligible. The minimum size of the 2D simulation model for SCM was at least twenty Debye lengths. While simulating the response of 1D doping profiles, unwanted boundary related artifacts have been avoided by dedicated models with a fixed probe and a moving doping profile.

Basically, the simulation model consisted of a rectangle of silicon with a grounded backside contact and with small rectangular refinement region centered on the tip axis. Five concentric semi spheric shells have been defined under the tip contact and the grid has been generated by boundary conforming meshing. Close to the tip contact, where the inversion layer is formed, the distance between adjacent grid points was in the sub-nanometric range. The distance between the points increases with the depth reaching hundreds of nanometers. The concentric semi spheres of the models allow to relax the grid step to accurately map the depletion region under the tip, which may reach a depth of some micrometers in the case of low doping concentrations and especially in the case of a depleted region of a \( pn \) junction. This meshing scheme has been demonstrated to be very efficient for 2D and cylindrical symmetric 3D models. In fact the resulting number of grid points was still reasonable and allowed to manage a whole set of simulations of a scanning tip within a reasonable computing time.

In order to cope with the finite diffusion length of the carriers, SSRM simulations have been carried out by simulation models with a realistic size in the millimeter range. In accordance with the physical models of the current spreading from macroscopic contacts, the size of the mesh refinement region extends up to ten times the tip radius. Because of the required accuracy and the macroscopic size of the simulation model, the number of grid points of the 3D model was about 70000, in spite of the small refinement area and in spite of the
fast relaxation rule used to define the distance between adjacent grid points.

Nevertheless, the CPU simulation time was still within reasonable limits (20 min). The necessity to verify the quantitative aspect of SSRM for dopant profiling, and the feasibility of three-dimensional simulation led to develop and investigate a 3D simulation model. Computing times for SSRM simulation became much larger with the use of Schottky instead of ohmic contacts, and with the use of models for the surface charges. Most of the simulations have been set with a tip being an ohmic boundary condition, with the intention to include the presence of the highly conductive $\beta-\text{Sn}$ phase.

5.5 Microscopy application of SCM and SSRM

SCM and SSRM have been shown to be definitely two excellent microscopy tools. They both offer the capability to image with sufficient contrast different doping concentration within a dynamic range covering up to six orders of magnitude, and to resolve non-conducting embedded structures (e.g. oxide layers) exhibiting nanometer features. If the doping type has to be discriminated ($p$- or $n$-type), a SCM system equipped with a phase-sensitive detector has to be preferred. This is also the case when dealing with large area samples or with devices needing multiple scans to be localized. In fact, SCM requires less pressure to be applied to the probe, thus it produces a reduced surface damage and a lower tip wear out.

SSRM, in particular in conjunction with a full diamond cantilever, has to be preferred when a very high lateral resolution is required, as it is the case for the imaging of nano-structured devices. Even at intermediate doping concentrations, some contrast arises between adjacent regions with different doping types, because the Schottky contact between the tip and the sample is operated either in direct or in reverse polarization depending on the position. However, due to the complex conduction mechanisms under the reverse tip polarization, contrast reversal effects have been observed to arise in some cases in the lower doping concentration range. Furthermore, the transition
from different doping types could not be always identified, since SSRM images the depletion zone under some particular conditions, only.

5.6 Quantification of SCM and SSRM

The quantification of doping profiles measured either by SCM or by SSRM has been done in the past with the use of procedures based on direct data conversion. This approach has been demonstrated to be very inaccurate if the extension of the structure to be quantified is much smaller than the local Debye length.

In this work, a new quantification procedure based on the numerical solution of the inverse modeling problem has been proposed. It combines a physical-based sampling strategy with the use of Artificial Neural Networks (ANN) trained with a wide set of 1D simulations. The 1D doping profiles returned by this novel conversion scheme, demonstrated an accuracy level that is comparable to the results obtained with exact analytical methods.

The main advantages of the application of ANN are the suitability for arbitrary carrier profiles, the fact that no iterations are needed, and the relative easiness of the extension of the method to the two-dimensional problem. Moreover, ANN can be applied without any pre-processing to noisy data as they occur in real measurements. In fact, profiles obtained by the novel conversion scheme turned out to be insensitive to the noise, whereas the analytical procedure delivered almost useless data, unless a smoothing algorithm with a very low bandwidth was used with a consequent relevant loss of information.

5.7 Outlook

In terms of 2D imaging of nano-scale devices, the future needs of metrology can be properly fulfilled by scanning probe techniques like SCM and SSRM.

Nowadays, the use of SCM and SSRM to extract quantitative 2D doping profiles is still debated, in particular when bipolar samples are concerned. The quantitative extraction of 1D doping profile by SCM and SSRM has been demonstrated to be feasible, provided that the
techniques are calibrated case by case, and that the raw experimental data is properly pre- and post-processed.

SCM represents a valid method for junction delineation under the assumption that the sample surface is properly prepared and that strategies are used, which exploit the information related to the shape of the local C-V curves. In general, the predictability of the response of the SCM systems could be largely improved if the amplitude of the UHF component in the capacitance sensor could be further reduced.

The large development potential of SSRM resides in the fact that the conversion of the measured resistance into resistivity can be made more reproducible and repeatable. This can be achieved by keeping under control the parasitic components of the resistance. Therefore, the immediate needs for SSRM include the optimization of the sample preparation to improve the surface conditions and minimize backside contact resistance, and the development of the new diamond probes for increased reliability, robustness, spatial resolution, and minimum resistance.

For the exploitation of the complementarities of both techniques, the development of multi-mode systems can be also developed to perform SCM and SSRM measurements by the same cantilever and integrated sensor.

In summary, although up to date SCM and SSRM cannot be yet considered as standard quantitative techniques for universal use, the aforementioned investigations have pointed out that the quantitative use of both techniques still have large improvement margins, especially if they are applied to a restricted variety of experimental cases. Furthermore, their suitability for the future semiconductor materials as silicon carbide and diamond has been already demonstrated.
Bibliography


[31] Digital instruments, Santa Barbara, USA.


[63] P. Eyben, N. Duhayon, T. Clarysse, W. Vandervorst, Bias-induced junction displacements in scanning spreading resistance


[72] P. Malberti, L. Ciampolini, M. Ciappa, W. Fichtner, Experimental investigation and 3D simulation of contrast reversal effect in
scanning capacitance microscopy, Characterization and Metrology for ULSI Technology, AIP conference proceedings, 2000.


Acknowledgements

First of all I would like to thank Prof. Wolfgang Fichtner, who gave me the opportunity to work under his supervision on this project. I also thank Dr. Vito Raineri for being the co-examiner of this work. I am thankful to Mauro Ciappa for the scientific supervision and for his careful and precious review of the manuscript. Special thanks to Prof. Massimo Vanzi from University of Cagliari who first believed in me and gave me the opportunity to study at ETH.

I am grateful to all the colleagues of IIS for the hints and support along these four years. Special thanks to Andreas Schenk, Jens Krause, Lorenzo Ciampolini, and Hidehiko Yabuhara who gave me a large help in several occasions.

I am also indebted to the Integrated Systems Engineering (ISE) people, in particular Andreas Wettstein, Gyorgy Kiralyfalvi, and Heinz-Olaf Muller, for the availability, help, and hints they gave me.

I would like also to thank Peter de Wolf and all the researcher of the HERCULAS network, coordinated by Prof. Wilfried Vandervorst. Within this project, I had the possibility to meet and collaborate with other young researchers and specialists who enlarged my scientific field of view with suggestions, new ideas and enthusiasm.

Of course I have to thank Christoph Wicki and Anja Boehm, for their care and patience in solving any small problem of my workstation. A special thank goes to Hanspeter Mathys for its great kindness and humanity. Many thanks also to the administration group: Margit, Verena, Christine, Hansjörg and Bruno.

I thank my colleagues of the Physical Characterization Laboratory for their friendship and collaboration: Chiara Corvasce, Davide Barlini, Yves Saad, Luca Sponton, and Marco Buzzo, and all my friends at
ETH: Chiara Martelli, Corrado Carta, Renato Negra, Claudia Cibin, Valerio Laino, Flavio Carbognani.

I am very thankful to the thesis students: Maurino Chiappini, Elisa Ricci and Camilla Coletti, for their important contributions to this work.

I owe gratitude to my friends Elisabetta Franceschi, Matteo Michelotto, Giovanna Mura, Kathrin Münscher, and finally, I thank my family for the support they constantly provide me.
Curriculum Vitae

Maria Virginia Stangoni was born in Sassari, Italy, in 1972. She received her M.Sc. degree in Electronic Engineering from the University of Cagliari, Italy, with a thesis on phase-resolved Scanning Capacitance Microscopy characterization of semiconductors. In 2001 she joined the Integrated Systems Laboratory of the Swiss Federal Institute of Technology (ETH) in Zurich as research and teaching assistant. Her work focused on high-resolution techniques for imaging and doping profiling of semiconductors. This research, including physical device simulation, experimental measurements, and data quantification, has been carried out in the framework of the European Network HERCULAS and of the European Network of Excellence SINANO. She received her Ph.D. degree in Electrical Engineering Sciences in 2005.