UNIVERSITY OF SOUTHAMPTON

FABRICATION AND CHARACTERIZATION OF P-MOSFETS WITH A STRAINED SIGE CHANNEL

Urs Norman Straube

A thesis submitted for the degree of
Doctor of Philosophy

Department of Electronics and Computer Science
Faculty of Engineering and Applied Science

March 2002

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

Faculty of Engineering and Applied Science

Department of Electronics and Computer Science

Doctor of Philosophy

FABRICATION AND CHARACTERIZATION OF P-MOSFETS WITH A STRAINED SIGE CHANNEL

by Urs Norman Straube

Strained SiGe heterostructures possess transport properties superior to Si. Their integration in the channel of MOSFETs can lead to an enhanced speed performance and is therefore an alternative to the scaling of the channel length.

This thesis reports on the fabrication and characterization of p-MOSFETs with a strained SiGe channel. As both design and material properties determine the performance of MOSFETs, the mobility in the SiGe heterostructure must be seen in context with the MOSFET design. SiGe heterostructures fabricated by molecular beam (MBE) and solid phase (SPE) epitaxial growth have been successfully integrated in the channel of p-MOSFETs.

A manufacturable SiGe CMOS process based on SPE has been demonstrated. A slightly enhanced oxidation rate and an increased interface state density have been observed for Si oxidized after Ge^+ implantation. The low performance of the SiGe p-MOSFETs is explained by the scattering and trapping of carriers at the degraded interface. This suggests that changing the annealing conditions could improve the SiGe p-MOSFETs.

Devices have been fabricated on a locally grown MBE layer. A considerable mobility enhancement over Si references has been shown for devices with a SiGe layer grown by MBE - more than or nearly a factor of two for devices with an LTO or thermal gate oxide, respectively. The devices fabricated on pillars display considerably better transport properties than the devices not fabricated on pillars. Altering the standard cleaning procedures has enabled us to achieve the required process control of one nanometer. Otherwise, the carriers would not be confined in the buried SiGe channel and would not profit from its enhanced transport properties. Low source and drain resistances have been obtained for SiGe p-MOSFETs subject to a maximum processing temperature of 800^oC.

Acknowledgments

I would like to thank my supervisors Prof. A. G. R. Evans and Prof. H. Kemhadjian for their support. The composure, patience and excellent guidance of Prof. A. G. R. Evans was the prerequisite for the enjoyable working environment.

Special thanks goes to the coworkers from other universities. Prof. E. H. C. Parker, Prof. T. E. Whall, Dr. G. Braithwaite, Dr. Tim Grasby, Dr. M. Palmer, Dr. C. Parry, Dr. P. Phillips, and A. Capewell from the University of Warwick were involved in the MBE growth for my SiGe devices and low energy SIMS analysis. Prof. P. L. F. Hemment, Dr. H. Graoui, Dr. C. Jeynes, and Dr. A. Nejim from the University of Surrey did critical implants, TEM or RBS analysis work of my samples. Prof. A. G. Cullis, Dr. G. Hill and Dr. D. J. Norris from the University of Sheffield analyzed my samples by SEM and TEM. Dr. S. Kaya and Dr. J. Watling from the University of Glasgow provided information about the SiGe heterostructure theory. Dr. K. Michelakis from Imperial College did s-parameter measurements on my devices.

Many thanks goes as well to my coworkers Dr. R. Sidek, Dr. Y. Tang and Dr. A. Waite as to helpful fellow students (Dr. I. Anteney, Dr. N. D'Halleweyn, J. Benson, Dr. T. Niblock, Dr. M. Routley, and S. Stefanou) from the University of Southampton.

My sincere appreciation is expressed to all staff of the Microelectronics Group. I especially would like to thank Dr. N. Afshar-Hanaii, Dr. J. Bonar, Dr. N. Lloyd, Dr. P. Routley, B. Ault, R. Baily, T. Blackburn, J. Clarke, S. Croucher, J. Humphrey, M. Josey, C. Kratovila, A. MacManus, I. McNally and A. Purdy.

Besides, I would like to thank Dr. A. Holmes-Siedle for his advice to do a Ph.D. at the University of Southampton and my parents for their continual support.

Urs N. Straube

Southampton

March 2002

Glossary of Symbols

Symbol	Description
a:	lattice constant
A:	frequency factor
A:	Richardson's constant
A_g:	gate area
a_o:	lattice constant of an unstrained $Si_{1-x}Ge_x$ alloy
a_c:	hydrostatic deformation potential of the conduction band
a_c^{dir}:	hydrostatic deformation potential of the direct conduction band
a_c^{ind}:	hydrostatic deformation potential of the indirect conduction band
a_v:	hydrostatic deformation potential of the valence band
$a_{\parallel}$:	lattice constant of an strained $Si_{1-x}Ge_x$ epilayer parallel to the growth plane
$a_{\perp}$:	lattice constant of an strained $Si_{1-x}Ge_x$ epilayer perpendicular to the growth plane
b:	amount of the Burgers vector
b_0:	amount of the Burgers vector of a perfect dislocation ($a/2 < 101 >$)
b_1:	amount of the Burgers vector of a $a/6 < 211 >$ dislocation
b_2:	amount of the Burgers vector of a $a/6 < 1\bar{1}2 >$ dislocation
b_e:	edge component of the Burgers vector
b_s:	screw component of the Burgers vector
C:	series capacitance of C_{ox} and C_{cap}
C_b:	capacitance due to the bulk charge
C_d:	depletion region capacitance
C_g:	gate capacitance
C_g:	oxygen concentration in the gas stream
C_{gc}:	gate-channel capacitance
C_i:	capacitance due to the inversion charge
C_i:	oxygen concentration at the SiO_2 interface
C_o:	oxygen concentration in the oxide
C_{ox}:	gate oxide capacitance
C_{cap}:	cap layer capacitance
C_s:	space region capacitance
C_s:	oxygen concentration in the gas stream at the wafer surface
$c_{ij,Ge}$:	elastic constant of pure Ge
$c_{ij,Si}$:	elastic constant of pure Si
d:	atomic spacing along the crystallographic direction
d:	deformation potential
$\overline{D_{it}}$:	mean surface state density
D_n:	diffusion constant for electrons
D_{O_2}:	diffusion coefficient of oxygen in SiO_2

Symbol	Description
D_s:	surface diffusivity
E:	electrical field
E:	incident energy
E_d:	energy barrier for the desorption
E_a:	activation energy
E_b:	binding energy
E_b:	potential barrier
E_B:	kinetic energy associated with boron
E_{BF_2}:	kinetic energy associated with BF_2
E_c:	conduction band energy
E_o:	vacuum energy
E_f:	Fermi energy
E_{eff}:	effective vertical field
E_{fm}:	Fermi energy of metal
E_g:	energy gap
E_i:	intrinsic Fermi energy
E_{ox}:	Fermi energy of silicon dioxide
E_{si}:	electric field at the oxide interface in Si
E_{ox}:	electric field at the oxide interface in SiO_2
E_T:	total energy
E_v:	valence band energy
$E_{v,1}$:	light hole valence band
$E_{v,2}$:	heavy hole valence band
$E_{v,3}$:	split-off band
E_δ:	energy accommodated by the dislocations
E_ϵ:	energy associated with the elastic strain
f:	frequency
f:	lattice mismatch
f_c:	cutoff frequency
f':	pre-exponential factor
F_o:	desorbing flux
F_δ:	force due to the selfenergy of the interfacial dislocations
F_ϵ:	elastic force
g:	shear strain potential
g_d:	conductance
g_m:	transconductance
G:	epilayer shear modulus
Ge_{max}:	critical Ge dose
h:	epilayer thickness
h:	Planck's constant
H:	Henry's gas constant
h_g:	mass transport coefficient
H_{21}:	current gain
I_{cp}	charge pumping current
I_{ds}:	current between source and drain
I_{dsat}:	saturation current
I_1:	source and drain current
I_1:	current at port 1

Symbol	Description
I_2:	substrate current
I_2:	current at port 2
j_{inc}:	incident flux of adatoms
J_n:	electron current density
j_s:	surface current density of adatoms
j_v:	net flux of adatoms
J_1:	oxygen flow through the stagnant gas layer
J_2:	oxygen flow through SiO_2 layer
J_3:	oxygen flow reacting with Si
J_S:	saturation current density
k_s:	chemical rate constant
K_p:	Preston coefficient
l_d:	length of the pinch-off region
L:	effective channel length
L_o:	drawn channel length
L_D:	Debye length for holes
L_m:	drawn channel length
M_B:	mass of boron
M_{BF_2}:	mass of BF_2
m_{hh}:	heavy hole mass
m_{lh}:	light hole mass
m^*:	effective mass
n:	free electron density
n_o:	electron density at thermal equilibrium
n_i:	intrinsic carrier concentration
n_s:	density of adsorbed atoms
n_s:	electron surface concentration
N_s:	density of surface positions
n_{seq}:	equilibrium density of adsorbed atoms
N_a:	acceptor density
N:	impurity concentration
N_{asd}:	acceptor density in source and drain
N_b:	net bulk doping
N_d:	donator density
p:	free hole density
p:	pressure
P_g:	partial pressure of oxygen in the gas flow
p_o:	hole density at thermal equilibrium
q:	elementary charge
Q_o:	effective interface charge
Q_b:	bulk charge
Q_g:	gate charge
Q_i:	inversion charge
Q_{id}:	inversion charge at the drain end
Q_{is}:	inversion charge at the source end
Q_s:	space charge
r_o:	dislocation core parameter
R:	deposition rate

Symbol	Description
R:	removal rate
R_c:	contact resistance
r_{ch}:	channel sheet resistance
R_{ch}:	channel resistance
R_d:	drain resistance
$\bar{R}_p$:	projected range
R_s:	source resistance
R_{sd}:	combined source and drain resistance
$s_{11}, s_{12}, s_{21}, s_{22}$:	s-parameters
T:	absolute temperature
t_{ch}:	thickness of the channel
t_f:	fall time
t_h:	time at high voltage
t_l:	time at low voltage
t_{ox}:	gate oxide thickness
t_r:	rise time
T:	absolute temperature
T_p:	period
U_s:	energy barrier
v:	relative velocity
V:	voltage
v_o:	superimposed signal
V_b:	potential between bulk and source
V_{base}:	base voltage
V_{cb}:	potential between channel surface and source
V_{ds}:	potential between source and drain
V_{dsat}:	drain saturation voltage
V_{fb}:	flat band voltage
V_g:	potential between gate and substrate
V_{gd}:	potential between gate and drain
V_{ox}:	potential dropped across the oxide
V_{pt}:	punch-through voltage
V_{rev}:	reverse bias
V_t:	thermal voltage
V_{th}:	threshold voltage
V_{thd}:	threshold voltage for a short channel device
W:	effective channel width
W:	width of depletion region
W_{disl}:	selfenergy of misfit dislocations
W_{fault}:	selfenergy of stacking faults
W_{layer}:	strain energy of the epilayer
W_s:	activation energy
$y_{11}, y_{12}, y_{21}, y_{22},$:	y-parameters
x:	distance from the Si/SiO_2-interface into the silicon
x:	Ge content
x_d:	end of the space charge region
X_{Ge}^{peak}:	peak Ge concentration
y:	distance from the edge of the source along the channel

Symbol	Description
y_c:	position of the amorphous/crystalline interface
Z_i:	proton number of ion
Z_t:	proton number of target
α:	factor describing the energy of the dislocation core
β:	gain factor
β:	kurtosis
γ:	body factor
γ:	skewness
γ:	stacking fault energy
δ:	lattice strain mismatch relieved through the formation of misfit dislocations
Δ_o:	spin orbit splitting parameter
ΔE_c:	conduction band offset
ΔE_v:	valence band offset
ΔE_c^i:	shift of the conduction band valley i
$\Delta E_{c,av}$:	shift of the position of the average conduction band
$\Delta E_{v,av}$:	shift of the position of the average valence band
$\Delta E_{v,1}$:	shift of the light hole valence band
$\Delta E_{v,2}$:	shift of the heavy hole valence band
$\Delta E_{v,3}$:	shift of the split-off band
ΔG:	free energy change
ΔG_n:	free energy per unit volume
ΔL:	difference between drawn and electrical channel length
$\Delta \bar{R}_p$:	standard deviation of the projected range
ΔV_a:	amplitude of superimposed signal
$\Delta\Delta L$:	error of Δ L
ϵ:	dielectric constant
ϵ:	lattice mismatch strain between epitaxial layer and substrate
ϵ^*:	equilibrium strain
ϵ_{ch}:	permittivity of the SiGe channel
ϵ_o:	initial lattice mismatch strain
ϵ_o:	vacuum permittivity
ϵ_{ox}:	dielectric constant of SiO_2
ϵ_s:	initial density of dislocations
ϵ_{si}:	dielectric constant of silicon
$\epsilon_{\parallel}$:	lattice mismatch strain between epitaxial layer and substrate parallel to the plane of the interface
$\epsilon_{\perp}$:	lattice mismatch strain between epitaxial layer and substrate perpendicular to the plane of the interface
Θ:	angle between the dislocation line and its Burgers vector
Θ:	sum of Θ_o and R_{sd}
Θ_b:	mobility degradation factor due to the substrate voltage
Θ_c:	mobility degradation due to the voltage between source and drain
Θ_o:	mobility degradation factor
λ:	angle between the misfit dislocation Burgers vector and a line in the interface drawn perpendicular to the dislocation line
λ:	channel length modulation factor
λ:	propagation constant
λ_s:	mean migration length

Symbol	Description
μ:	mobility
μ_{eff}:	effective mobility
μ_{effv}:	effective mobility taking into account the variation of C_{gc}
μ_{FE}:	field-effect mobility
μ_{Ge}:	mobility in Ge
μ_l:	mobility due to the longitudinal field
μ_n:	electron mobility
μ_o:	low-field mobility
μ_s:	surface mobility
μ_{Si}:	mobility in Si
μ_t:	mobility due to the transversal field
ν:	epilayer Poisson ratio
ρ:	charge density
σ:	DIBL parameter
σ:	supersaturation parameter
σ:	surface energy
σ_n:	capture cross section of electrons
σ_p:	capture cross section of holes
$\sigma_{R_{sd}}$:	standard deviation of R_{sd}
$\sigma_{r_{ch}}$:	standard deviation of r_{ch}
$\sigma_{\Delta L}$:	standard deviation of ΔL
σ_{μ_o}:	standard deviation of μ_o
σ_{Θ_o}:	standard deviation of Θ_o
τ_t:	transport scattering time
υ:	carrier velocity
ϕ:	electrical potential
ϕ:	implantation dose
ϕ_b:	barrier height
ϕ_{bi}:	built-in potential of a pn-junction
ϕ_f:	Fermi potential
ϕ_{fi}:	intrinsic Fermi potential
ϕ_{fn}:	quasi-Fermi potential for n-type silicon
ϕ_{fp}:	quasi-Fermi potential for p-type silicon
ϕ_s:	surface potential
ϕ_{ss}:	constant value for the surface potential
:	in the subthreshold region model
Φ_m:	metal work function
Φ_{max}:	critical implant dose
Φ_{ms}:	gate to substrate work function difference
Φ_s:	semiconductor work function
χ_s:	electron affinity of silicon
χ_{ox}:	electron affinity of silicon dioxide
Ψ:	critical angel for channelling
ω_c:	cutoff frequency

Contents

Chapter 1

Introduction

From the Dawn of the IC Era to SiGe MOS

The MOSFET (Metal-Oxide-Semiconductor-Field-Effect-Transistor) [1-4] was proposed in the early 1930s by Lilienfeld and Heil. After Pearson and Shockley had studied the MOSFET in the late 1940s, Kahng and Atalla fabricated the first MOSFET using a thermally oxidized silicon structure in 1960. At the beginning of the IC era in 1959, ICs based on bipolar transistors occupied the broadest market share. Because of the advantages in device miniaturization, low power consumption, and high yield, the sales volume of CMOS (complementary MOS) based ICs that employs the pairing of n- and p-channel MOSFETs has steadily soared. In 1993, the MOSFET market share amounted to 75 % of the total IC market and is predicted to capture 88% by the year 2000. Apart from Si based MOSFETs and bipolar transistors, III-IV compound semiconductors for high speed devices form the third major IC group.

The speed performance of the MOSFET has steadily been improved by shrinking the channel length and gate oxide thickness. Today, scaling is close to reaching its performance and physical limits. Thin gate oxides result in the tunneling of carriers through the gate oxide degrading continuously the oxide quality. The negative effect of scaling the gate length on the subthreshold characteristics has to be compensated by increasing the bulk doping concentration. However, the increase of the bulk doping concentration leads to the reduction of the mobility and an increase of the electrical field at the source and drain junctions favoring hot carrier generation.

Semiconductor heterostructures were suggested by Shockley in a patent for the first time in 1951. Since the beginning, Ge has been an obvious candidate for the fabrication of silicon based heterostructures because it is chemically compatible with Si. However, the large lattice mismatch between Si and Ge makes it difficult to grow Ge epilayers on a Si substrate. In the late 1960s and early 1970s, the first unsuccessful attempts to grow Ge epilayers on a Si substrate were performed.

The first strained SiGe heterostructure devices were realized by Erich Kasper based on the

concept of Frank and Van der Merwe. The first SiGe FET was reported in 1985 [5]. SiGe can be used in gate, source and drain or substrate. In the gate, SiGe can reduce the gate resistance and adjust the threshold voltage. In source and drain, SiGe can improve the short channel characteristics and prevent parasitic bipolar transistor action. The interest in strained SiGe layers for the application in the channel arose from the discovery that strained SiGe layers possess a higher mobility than silicon and was soon adopted as an alternative approach to enhance the speed performance of MOSFETs.

Objective of this Project

This work focuses on the use of strained SiGe layers in the channel of p-MOSFETs. A considerable performance improvement of SiGe MOSFETs in comparison to Si reference MOSFETs was found [6]. However, the SiGe MOSFETs presented in ref. [6], do not have a performance comparable to state-of-the-art MOSFETs because the process was tailored for the requirements imposed by the incorporation of strained SiGe layers.

The integration of an MBE grown SiGe layer into a competitive CMOS process is one of the three major objectives of this project. Fundamental problems are the thermal budget and the control of the Si cap layer. A high thermal budget can lead to the relaxation of the SiGe layer and, in consequence, a deteriorated performance. A thin Si cap layer is essential for the confinement of the carriers in the SiGe channel.

A thick Si buffer layer is supposed to be the prerequisite for a high hole mobility in the SiGe channel. However, small geometry devices require a sufficient dopant concentration in the bulk in order to prevent short channel effects. Implantation may damage the integrity of the heterostructure and little is known about the incorporation of dopants during MBE. The incorporation of an insitu doped MBE layers poses further problems in CMOS as MBE does not allow selective growth. A thick, insitu doped MBE layer must either be etched off or counterdoped for the n-MOSFETs.

The fabrication of MOSFETs with a Ge^+ implanted channel [7-10] is a further objective of this work. MBE is the ideal technique for the fabrication of high quality heterostructures. But, its major drawback is that heterostructures grown by MBE cannot be easily integrated into CMOS. Besides, MBE does not have a high throughput and is expensive. Although strained heterostructures with only low Ge concentration can be achieved by ion implantation compared to MBE, the low cost for implantation and the capability to implant Ge^+ into a selected area may render ion implantation an attractive alternative to MBE. In solid phase epitaxial growth (SPEG), strained SiGe layers can be grown in a recrystallization step after the implantation of Ge^+. Ge^+ implantation does not allow the fabrication of a Si cap layer on top of the SiGe channel which reduces carrier scattering at the Si/SiO_2-interface. As proposed in ref. [7], the performance of Ge^+ implanted p-MOSFETs might be enhanced due to Ge segregation during

gate oxide growth resulting in a higher Ge peak concentration. The EPIFAB process (epitaxial regrowth across phase boundaries) developed at the University of Surrey enables the fabrication of SiGe/Si heterojunctions without extended defects [11-15].

The third and last objective of this project is the investigation of p-HMOSFETs displaying a considerable mobility enhancement. A high Ge content in the strained SiGe channel resulting in a considerable mobility enhancement can be achieved only if grown on top of a virtual substrate [16-20]. As the virtual substrate has to be relaxed, a thick MBE layer is in general required.

The local MBE growth on microscopic pillars can reduce the required thickness of the virtual substrate layer, crosshatch, and threading dislocations [21, 22]. Trenches etched into the Si substrate interrupt the MBE layer and are suitable sites for the nucleation of dislocations. SiGe MOSFETs fabricated on pillars should have a better performance than those on a flat surface.

Organization of the Thesis

The thesis contains 10 chapters and 4 appendices. All symbols are explained in the table at the beginning of the thesis.

The first chapter introduces the reader into the topic of the thesis. Chapters 2-6 give an overview over issues related to this work. Chapter 2 addresses the physical properties of SiGe heterojunctions. Strain is inherent in SiGe heterojunctions due to the lattice mismatch between Si and Ge and improves the electrical characteristics of SiGe alloys.

The third chapter describes the theory of conventional MOSFETs. The drain current characteristics of HMOSFETs with a strained SiGe layer differ slightly from those of conventional MOSFETs. The gate capacitance and the mobility can be a strong function of the gate bias. The constant capacitance approximation and a simple mobility degradation model are usually not justified. No analytical expression for the HMOSFET is given, but the different device behavior can be explained qualitatively.

In the fourth chapter, parameter extraction routines for conventional and HMOSFETs are discussed. The mobility extraction for HMOSFETs is investigated. It is shown that split-CV measurements developed for conventional MOSFETs with an ultrathin oxide are the appropriate methods for mobility extraction.

Chapter 5 deals with vapor phase deposition. Vapor-phase methods are predominant in the deposition of thin films. UHVCVD and MBE - both vapor phase deposition methods - are the most common methods for the fabrication of strained SiGe heterostructures. Chapter 6 reviews further standard processing steps that are also used in SiGe technology. The adaptations which may be imposed by the presence of strained SiGe layers (layer control, thermal budget, integrity of the heterojunction) and the potential benefits arising from the use of Ge are discussed.

Chapters 7-9 present SiGe batches processed in the course of this work. A SiGe CMOS process with a thin thermal gate oxide, an elevated thermal budget, a thin Si cap, and an insitu

doped MBE layer is investigated in chapter 7. The MBE layer is counterdoped for the fabrication of n-MOSFETs. The initial cap layer is thinned by several cleans and the growth of the thermal gate oxide. The calculation of the final cap layer thickness is based on an internal experiment on the etch rate of standard cleans. The control of the Si cap layer is even more critical than previously thought.

Chapter 8 presents a SiGe CMOS/Ge^+ implantation batch. This batch is based on a conventional CMOS batch and the EPIFAB process developed at the University of Surrey and is subject to the thermal budget of a conventional CMOS batch. The amorphous layer is recrystallized during gate oxide growth. Emphasis is put on the characterization of the Si/SiO_2-interface which may be modified due to the segregation of Ge during gate oxidation.

A SiGe virtual substrate/limited area PMOS batch is described in chapter 9. A virtual substrate layer is grown in order to achieve a high Ge content in the SiGe channel. The devices are fabricated on tiny pillars surrounded by trenches reducing the thickness of the virtual substrate and improving its quality. Such devices are compared to Si reference p-MOSFETs and p-HMOSFETs which are not fabricated on pillars.

In chapter 10, a conclusion of this project and a proposal for further work are given. In appendix A, a formula used in error calculation is derived. The detailed process listings of all three fabrication processes are given in appendix B. Appendix C presents electrical results for the SiGe CMOS/ Ge^+ implantation batch extracted with an alternative parameter extraction routine. Appendix D lists the complete results for the limited area/virtual substrate batch.

Bibliography

[1] C. Y. Chang and S. M. Sze, "ULSI Technology", McGraw-Hill, 1996

[2] S. M. Sze, "Physics of Semiconductor Devices", John Wiley & Sons, 1981

[3] N. Afshar-Hanaii, "Some Aspects of Submicron CMOS Technology", Ph.D. thesis, University of Southampton, 1993

[4] S. C. Jain, "Germanium-Silicon Strained Layers and Heterostructures", Academic Press, 1994

[5] T. P. Pearsall, J. C. Bean, R. People, and A. T. Fiory, "Ge_xSi_{1-x} modulation-doped p-channel field-effect transistors", in Proc. 1st Int. Symp. Silicon Molecular Beam Epitaxy, J. C. Bean Ed., Pennington, NJ, Electrochemical Society, pp. 400-405, 1985

[6] R. M. Sidek, U. N. Straube, A. M. Waite, A. G. R. Evans, C. Parry, P. Phillips, T. E. Whall, and E. H. C. Parker, "SiGe CMOS fabrication using SiGe MBE and anodic/LTO gate oxide", Semicond. Sci. Technol., Vol. 15, pp. 135-138, 2000

[7] H. Jiang and R. G. Elliman, "Electrical Properties of GeSi Surface- and Buried-Channel p-MOSFET's Fabricated by Ge Implantation", IEEE Transactions on Electron Devices, Vol. 43, No. 1, pp. 97-103, 1996

[8] K. C. Liu, S. K. Ray, S. K. Oswal, and S. K. Banerjee, "A deep submicron $Si_{1-x}Ge_x/Si$ Vertical PMOSFET Fabricated by Ge Ion Implantation", IEEE Electron Dev. Lett., Vol. 19, No. 1, pp. 13-15, 1998

[9] S. Kal, N and B. Chakrabarti, "Strained Silicon-SiGe Devices using Germanium Implantation", IETE J. Res., Vol. 43, No. 2 & 3, pp. 185-192, 1997

[10] Z. Xia, E. O. Ristolainen and P. H. Holloway, "$SiGe/Si$ heterostructures produced by double-energy Si^+ and Ge^+, and Ge^+ and Ge^{2+} ion implantations", J. Vac. Sci. Technol. B, Vol. 16, No. 5, pp. 2629-2632, 1998

[11] Y. Suprun-Belevich, F. Christiano, A. Nejim, P. L. F. Hemment, and B. J. Sealy, "Evolution of mechanical strain and extended defects in annealed (1 0 0) silicon samples implanted with Ge^+ ions ", Nucl. Instr. and Meth. in Phys. Res. B 140, pp. 91-98, 1998

[12] A. Nejim, N. P. Barradas, C. Jeynes, F. Cristiano, E. Wendler, K. Gaertner, and B. J. Sealy, "Residual post anneal damage of Ge and C co-implantation of Si determined by quantitative RBS-channeling", Nucl. Instr. and Meth. in Phys. Res. B 139, pp. 244-248, 1998

[13] F. Christiano, A. Nejim, B. de Maudit, A. Claverie, and P. L. F. Hemment, "Characterization of extended defects in SiGe alloys formed by high dose Ge^+ implantation into Si", Nucl. Instr. and Meth. in Phys. Res. B 120, pp. 156-160, 1996

[14] F. Christiano, A. Nejim, Y. Suprun-Belevich, A. Claverie, and P. L. F. Hemment, "Formation of extended defects and strain relaxation in ion beam synthesised SiGe alloys", Nucl. Instr. and Meth. in Phys. Res. B 147, pp. 35-42, 1999

[15] F. Christiano, C. Bonafos, A. Nejim, S. Lombardo, M. Omri, D. Alquier, A. Martinez, S. U. Campisano, P. L. F. Hemment, and A. Claverie, "Interstitial trapping efficiency of C^+ implanted into preamorphised silicon - control of EOR defects", Nucl. Instr. and Meth. in Phys. Res. B 127 / 128, pp. 22-26, 1997

[16] A. J. Pidduck, D. J. Robbinson, D. Wallis, G. M. Williams, A. C. Churchill, J. P. Newey, C. Crumpton, and P. W. Smith, "On the growth of high quality relaxed $Si_{1-x}Ge_x$ layers on Si by vapour phase epitaxy", Inst. Phys. Conf. Ser., No. 157, pp. 135-144, 1997

[17] M. Kummer, B. Voegli, and H. Von Kaenel, "STM study of step graded $Si_{1-x}Ge_x/Si(100)$ buffers", Thin Solid Films, Vol. 336, pp. 100-103, 1998

[18] E. Muller, R. Hartmann, C. David, and D. Grutzmacher, "Structural characterization of SiGe step graded buffer layers grown on prestructured Si [100] substrates by molecular beam epitaxy", Thin Solid Films, Vol. 336, pp. 92-95, 1998

[19] K. Lyutovich, E. Kasper, F. Ernst, M. Bauer, and M. Oehme, "Relaxed SiGe buffer layer growth with point defect injection", Materials Science and Engineering, Vol. B71, pp. 14-19, 2000

[20] E. Kasper, K. Lyutovich, M. Bauer, and Michael Oehme, "New virtual substrate concept for the vertical MOS transistor", Thin Solid Films, pp. 319-322, 1998

[21] T. Rupp, J. Messarosch, and I. Eisele, "Silicon-germanium molecular beam epitaxy system for high-quality nanostructures and devices", Journal of Crystal Growth, Vol. 183, pp. 99-108, 1998

[22] T. Rupp, F. Kaesen, W. Hansch, E. Hemmerl, D. J. Gravesteijn. R. Schorer, E. Silveira, G. Abstreiter, and I. Eisele, "Defect-free strain relaxation in locally MBE-grown SiGe heterostructures", Thin Solid Films, Vol. 294, pp. 27-32, 1997

Chapter 2

SiGe Heterostructures

The performance of CMOS circuits has steadily been improved by aggressively scaling the device dimensions to the submicron regime. The scaling becomes more and more difficult due to technological and physical limitations. But, the performance of CMOS devices could be improved by applying new materials. Today, silicon dominates microelectronics, whereas other semiconductor materials such as GaAs only occupy niche markets. In the near future, it is unlikely that other materials will substitute silicon whose $Si - SiO_2$ structure offers unique physical and chemical properties.

However, the development of new compound materials based on silicon could further enhance the performance of integrated circuits. The Si/SiGe system is a strong candidate for a new technology that would be compatible with the existing Si technology and might considerably enhance the speed of microelectronics devices because of higher carrier velocities.

The purpose of this chapter is to review the fundamental physical properties of SiGe heterostructures. SiGe heterojunctions without defects are necessarily strained due to the lattice mismatch between Si and Ge. Strain plays an integral role in the concept of the SiGe heterojunction. Strained SiGe heterostructures have better electrical properties than Si such as a higher carrier mobility. But, the same properties can be degraded by strain relief. Defects are formed if a certain energy barrier which is lowered with increasing temperature is surpassed.

2.1 Crystal Structure

Binary alloys of silicon and germanium are continuously miscible and crystallize in a diamond lattice. The diamond lattice consists of two face-centered cubic lattices displaced a quarter of the space diagonal. Its unit cell contains eight atoms (see fig. 2.1; note: in the 45^o-perspective, the two atoms at the front shown in red obscure two atoms at the back).

SiGe [1-13] bulk alloys are never ordered. However, order [1,14] exists at the interfaces of SiGe layers deposited by MBE on Si substrates of (100) orientation at growth temperatures between

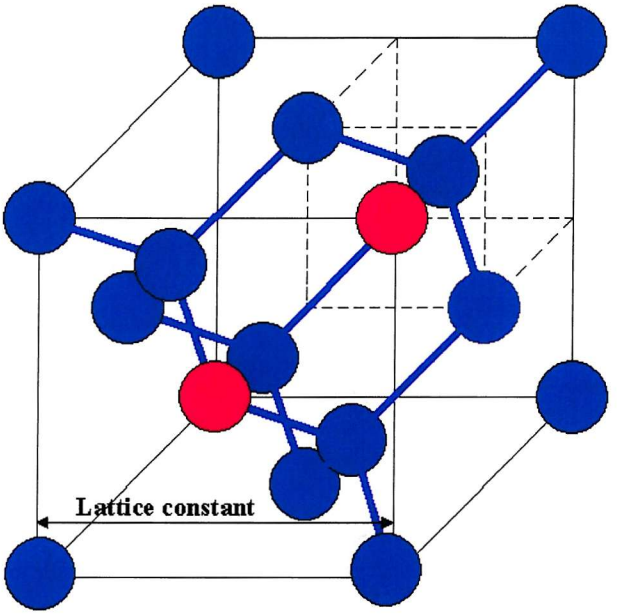


Figure 2.1: Diamond structure

Parameter	Si	Ge
c_{11}	165.8	128.5
c_{12}	63.9	48.3
c_{44}	79.6	66.8

Table 2.1: Elastic constants for Si and Ge

300^o C and 800^oC. An energetically more favorable configuration is certainly responsible for the order of thin SiGe layers and not yet fully explained.

The lattice parameter of Si and Ge differ by 4.2% at room temperature. The misfit increases slightly with temperature. The lattice parameter $a_0(x)$ of $Si_{1-x}Ge_x$ alloys deviates slightly from Vegard's rule and depends parabolically on the Ge composition [1, chap. 2]:

$$a_0(x) = 0.002733x^2 + 0.01992x + 0.5431(nm) \tag{2.1}$$

In an epitaxial layer with lattice constant a_1 grown on a substrate with different lattice constant a_0, the mismatch f can be accommodated by the tetragonal distortion and the formation of misfit dislocations. In heterojunctions, misfit is exclusively accommodated by misfit strain within the critical thickness. Then, its lattice parameter in the growth plane $a_{||}$ is completely determined by the underlying substrate. In the strained epilayer, the lattice parameter perpendicular to the growth plane $a_\perp$ is 5.82 $\AA$ for strained Ge on relaxed Si and 5.26 $\AA$ for strained Si on relaxed Ge.

The elastic constants of SiGe alloys c_{ij} can be determined from linear interpolation [1, chap. 3.1; 16; 17]:

$$c_{ij} = c_{ij,Ge}x + c_{ij,Si}(1-x) \tag{2.2}$$

Where $c_{ij,Ge}$ and $c_{ij,Si}$ are the elastic constants of pure Si and Ge given in tab. 2.1.

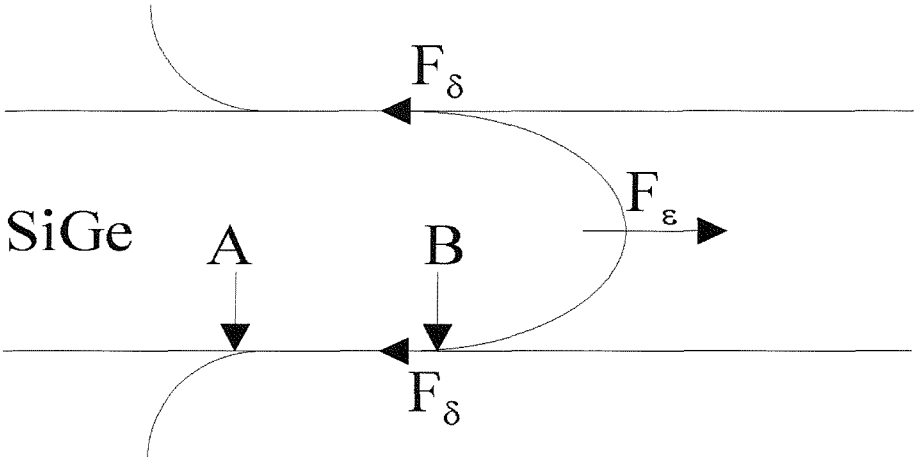


Figure 2.2: Threading dislocation acting as a misfit dislocation source in a capped SiGe layer

The strain parallel to the plane of the interface $\epsilon_{\|}$ is given by [17]:

$$\epsilon_{\|} = \frac{a_{\|}}{a_0} - 1 \tag{2.3}$$

Where $a_{\|}$ and a_0 are respectively the lattice constant of the substrate and the lattice constant of the relaxed alloy given by equation 2.1. The strain perpendicular to the plane of the interface $\epsilon_{\perp}$ is given by [17]:

$$\epsilon_{\perp} = \frac{a_{\perp}}{a_0} - 1 \tag{2.4}$$

Where $a_{\perp}$ is given by [17]:

$$a_{\perp} = a_0[1 - D(\frac{a_{\|}}{a_0} - 1)] \tag{2.5}$$

D depends on the interface orientation defined by Miller indices [17]:

$$D^{001} = 2\frac{c_{12}}{c_{11}} \tag{2.6}$$

$$D^{110} = \frac{c_{11} + 3c_{12} - 2c_{44}}{c_{11} + c_{12} + 2c_{44}} \tag{2.7}$$

$$D^{111} = 2\frac{c_{11} + 2c_{12} - 2c_{44}}{c_{11} + 2c_{12} + 4c_{44}} \tag{2.8}$$

2.2 Strain Relaxation of $Si_{1-x}Ge_x/Si$ Heterostructures

The critical thickness [18-23] is the minimum layer thickness at which the generation of misfit dislocations at the interface is energetically favored and can be found from equilibrium theory. It depends on the growth conditions (temperature, pressure...), the design of the final transistor (ie capped layer or not), the nucleation mechanism, the substrate and the germanium fraction. Capped SiGe layers [1, chap. 1.2] are slightly more stable than uncapped SiGe layers. Misfit strain relief may lag behind equilibrium predictions despite the presence of dislocations as it is often retarded by barriers to nucleation and motion of dislocations (Peierls-Nabarro barrier,

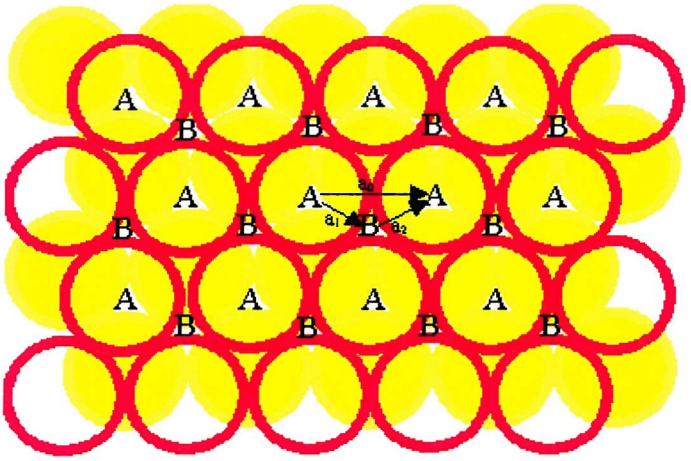


Figure 2.3: Plane (red hollow circles) slips over another plane (yellow solid circles)

[18, p. 215]). The discrepancy between theoretical predictions and experimental results becomes obvious for low growth temperatures. The regime between experimental measurement and theoretical values is often referred to as metastable regime.

In general, dislocations cannot terminate within the bulk of a crystal, but must either form a closed loop or at a node with another defect, eg other dislocation or surface (image forces, [18, p. 88]). A dislocation segment within a stressed region of a crystal experiences a force that has both climb and glide components. Gliding is opposed by the so-called Peierls barrier that is due to the discrete nature of the crystal, but is usually much faster than climbing that requires material transport. When the stress is too small to overcome the Peierls barrier, dislocations can still migrate by thermal creation, migration of kinks and diffusion.

Threading dislocations are the most favorable sources of misfit dislocations. As displayed in fig. 2.2, an existing threading dislocation is bowed under the elastic force F_ϵ of the strained $Si_{1-x}Ge_x$ layer and elongated to form a misfit dislocation line between A and B.

In crystals with diamond lattice, lattice constant a and covalent bonds like Si and Ge, perfect dislocations whose Burgers vector is a lattice translation vector have Burgers vector $a/2 < 101 >$ and slip on $\{111\}$ glide planes. Fig. 2.3 illustrates how the perfect dislocations dissociate into Shockley partial dislocations [1, chap. 1.2; 18, p. 93]:

$$\begin{aligned} a_o &= a_1 + a_2 \\ a/2 < 101 > &= a/6 < 211 > + a/6 < 1\bar{1}2 > \end{aligned} \tag{2.9}$$

Instead of moving directly from a lattice site A to another, it can be energetically favorable to move first to a B site and then to an A site because the energy of a dislocation is proportional to the square of its Burgers vector. The perfect dislocation with Burgers vector $a/2 < 101 >$ is stable for low Ge concentration and thick epilayers. It splits up otherwise into two dislocations with Burgers vectors $a/6 < 211 >$ and $a/6 < 1\bar{1}2 >$. The angle between the Burgers vector and the misfit dislocation line is 60^o for the parent $a/2 < 101 >$, 30^o and 90^o respectively for

the $a/6 < 211 >$ and $a/6 < 1\bar{1}2 >$ dislocations. These two Shockley partial dislocations [18, p. 91] are mutually repelled and glide apart. It results a ribbon of stacking fault between them on the [111] glide planes. The balance of their repulsive interactions and the intervening stacking fault energy determine the equilibrium partial separation which is typically in the order of a few nm for unstrained layers. The lattice mismatch stress is resolved differently onto the two partials (much higher on the 90^o than on the 30^o partial). Compressive stress reduces the partial separation whereas tensile stress increases the partial separation.

Results of general validity for strain relaxation can be obtained by the application of the common harmonic models [19] although they are quite crude. Harmonic models assign isotropic harmonic interactions to the the atoms within the crystal and on both sides of an interface where the atoms are not out of step and unharmonic interactions to the atoms which are out of step. Essentially, harmonic models treat the crystal as an elastic continuum. The two main harmonic approaches differ essentially in the representation of the crystal around the dislocations.

The Frenkel-Kontorowa model supposes that the interfacial force is limited to the atoms at the interface and varies periodically with disregistry (relative tangential displacement). Subsequently, the model was extended for two seminfinite crystals (Nabarro and van der Merwe) and for a monolayer with anisotropic interfacial misfit (Frank and van der Merwe). Unfortunately, the governing equations are mathematically complex and are not solvable for intermediate and thick epilayers.

The Volterra model assumes that two crystal halves that can be treated as elastic continua and are "glued" together all over the interface except for a hole of radius r_o along the dislocation line (see fig. 2.4). r_o eliminates a singularity in the mathematical description. It is of atomic dimension and provides a "core" energy for the dislocation. An outer radius R cancels stress fields by other dislocations and by the presence of surfaces. R is limited by the thickness of the eiplayer. The Volterra model is therefore not suitable for very thin epilayers. Besides, it does not account for interactions between individual dislocations. The Volterra model is the basis of the Matthews-Blakeslee theory [20], the most frequently quoted model for the calculation of the critical layer thickness.

In the Matthews-Blakeslee theory, the critical thickness of a uniform and uncapped $Si_{1-x}Ge_x$-layer grown on a Si substrate is found by equating the elastic force of the strained $Si_{1-x}Ge_x$ layer F_ϵ and the force due to the selfenergy of the interfacial dislocations F_δ (see fig. 2.2, [1, chap. 1.2]):

$$F_\epsilon = F_\delta \tag{2.10}$$

According to the isotropic linear elasticity theory, the elastic force of the strained SiGe layer is given by [1, chap. 1.2]:

$$F_\epsilon = 2Gbh\epsilon \cos\lambda \frac{1+\nu}{1-\nu} \tag{2.11}$$

The dislocation theory gives the following expression for the force due to the the selfenergy of

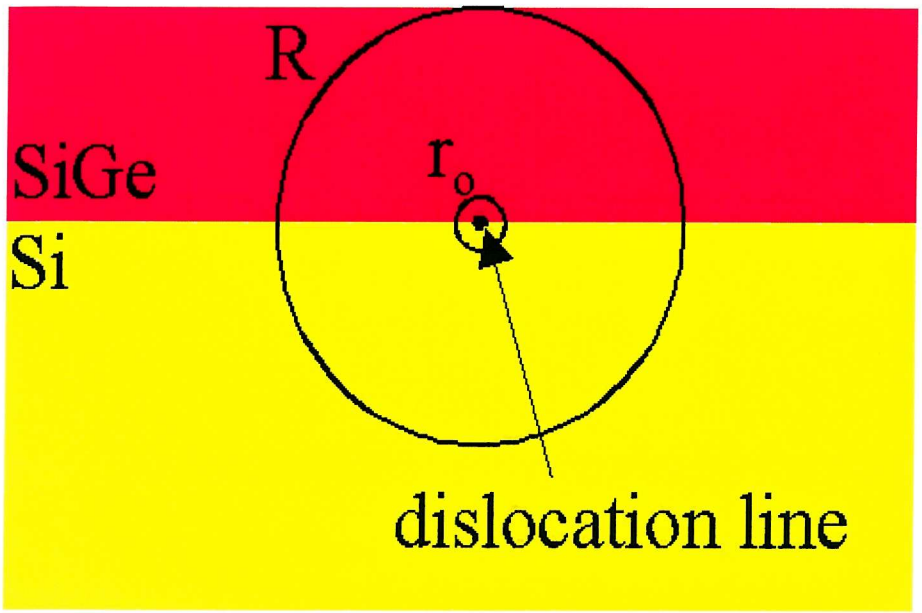


Figure 2.4: Cross section through a dislocation line in the Volterra model

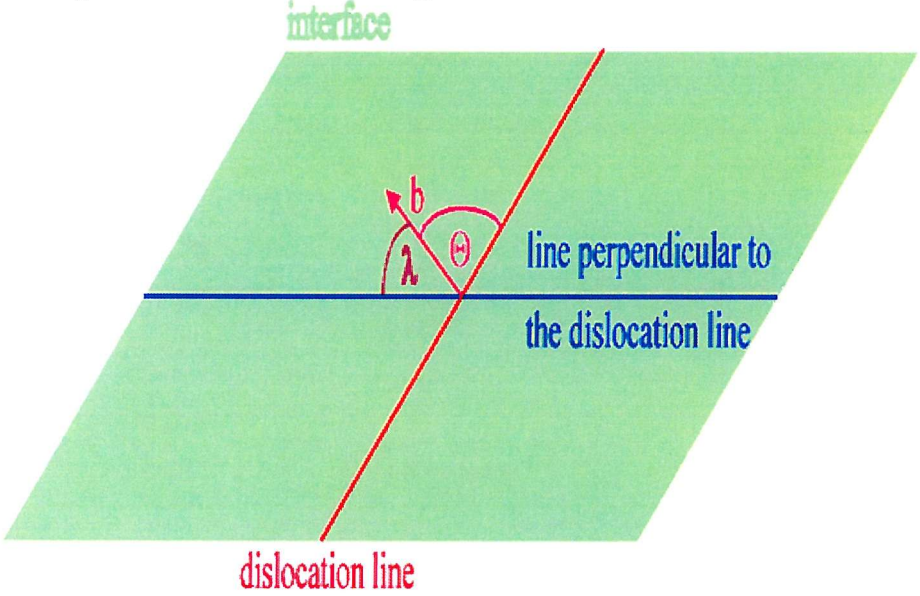


Figure 2.5: Position of λ and Θ

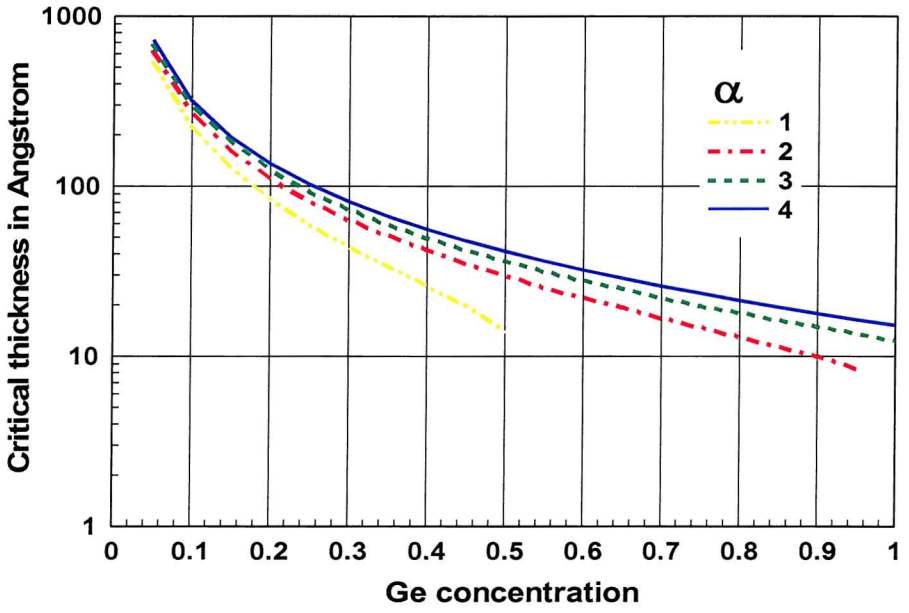


Figure 2.6: Theoretical critical thicknesses for different α

the interfacial dislocations [1, chap. 1.2]:

$$F_\delta = Gb^2 \frac{1 - \nu \cos^2 \Theta}{4\pi(1-\nu)} \ln \frac{\alpha h}{b} \tag{2.12}$$

Where h is the epilayer thickness, G is the shear modulus of the epilayer, ν is the Poisson ratio, b is the magnitude of the Burgers vector, λ is the angle between the misfit dislocation Burgers vector and a line in the interface drawn perpendicular to the dislocation line direction, Θ is the angle between the dislocation line and its Burgers vector and α is a factor describing the energy of the dislocation core (see fig. 2.5). Figure 2.6 shows the results for $G = 64\ GPa, \nu = 0.28, \cos\lambda = 0.5, \cos\Theta = 0.5, b = 3.9\text{Å}, \epsilon = 0.041\ x$. As the core energy of Si and Ge is not definitively known, different values of α are shown. Critical thicknesses can be calculated for any Ge content in the relaxed substrate and in the strained channel if the appropriate material parameters are chosen.

The Matthews-Blakeslee theory agrees well with experimental findings for $Si_{1-x}Ge_x$ structures grown at high temperatures. For low growth temperatures the critical layer thickness can exceed the predicted values considerably. In the metastable regime, strain relaxation is prevented by the presence of activation energy barriers to the nucleation and propagation of misfit dislocations. In the metastable regime and at a given temperature, The Dodson-Tsao model describes the time-dependent relaxation [15, p. 51]:

$$\frac{d\epsilon(t)}{dt} = CG^2(\epsilon_o - \epsilon(t) - \epsilon^*)^2(\epsilon(t) + \epsilon_s) \tag{2.13}$$

Where ϵ^*, ϵ_o, ϵ_s and C are respectively the equilibrium strain, the initial lattice mismatch strain, strain due to dislocations present at the beginning and a constant. As these values are most often not known, the Dodson-Tsao model seldom allows predictions concerning the strain relaxation.

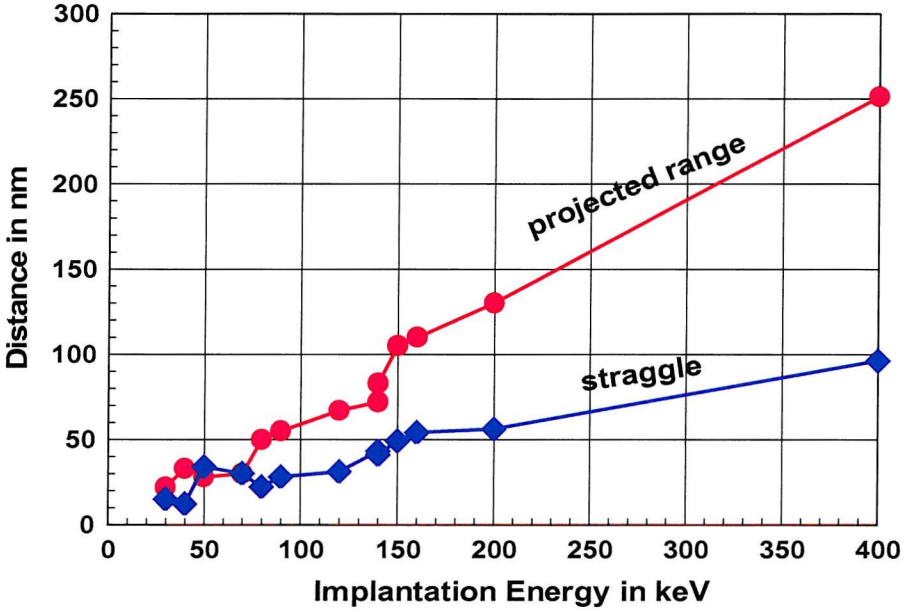


Figure 2.7: Projected range R_p and straggle ΔR_p for Ge [15, p. 160]

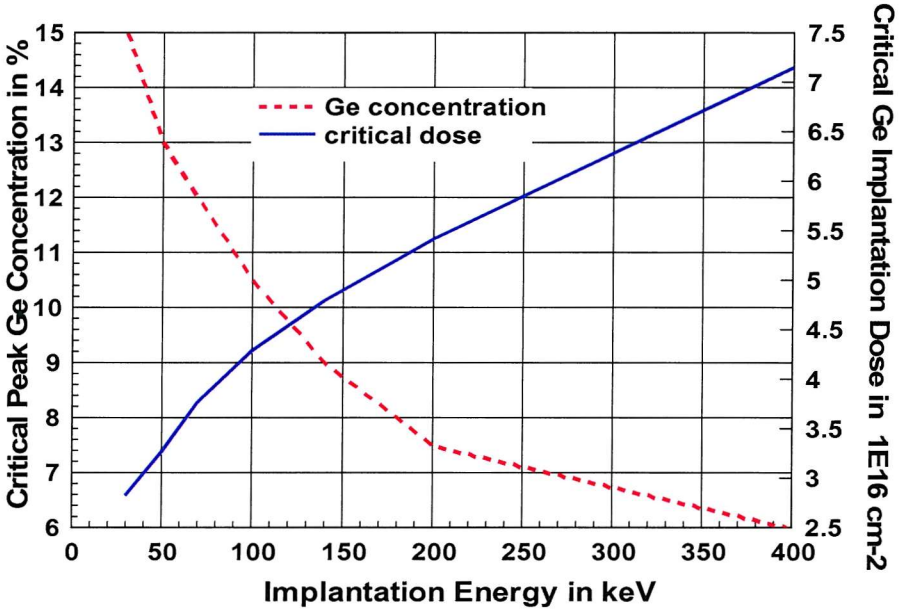


Figure 2.8: Critical peak Ge concentration and dose

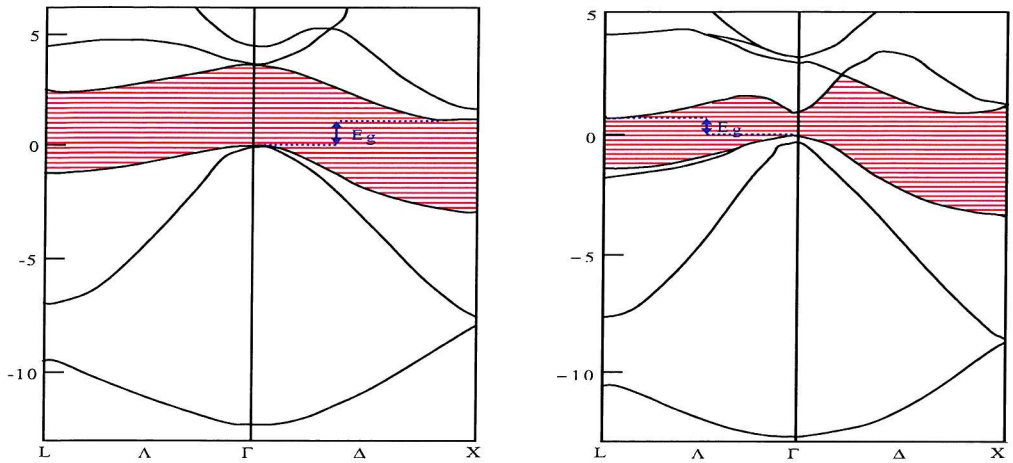


Figure 2.9: Band structure [25] of (a) Si (b) Ge

A linearly graded SiGe layer has the same critical thickness as a uniform SiGe layer that has the average Ge concentration of the linearly graded SiGe layer [2]. F. Christiano derived the critical Ge implantation dose [15, 23] from the minimization of the total energy $E_T = E_\epsilon + E_\delta$.

The energy E_ϵ associated with the elastic strain ϵ is given by [15, p. 152]:

$$E_\epsilon = 2G(\frac{1+\nu}{1-\nu}) \int_0^h \epsilon^2(y)dy \tag{2.14}$$

The energy E_δ associated with the misfit δ accommodated by dislocations is given by:

$$E_\delta = (f-\epsilon)\frac{Gb}{2\pi(1-\nu)}[\ln(\frac{h}{b})+1] \tag{2.15}$$

The model predicts a critical elastic strain energy of 310 mJ/m^2 that is independent of the implantation dose. Although the model does not account for stacking faults and is based on an energy balance approach instead of a force balance approach, it agrees with experimental findings [15, p. 160]. Assuming a Gaussian Ge profile, the following formula for the critical implant dose Φ_{max} can be derived [24] (see fig. 2.8):

$$\Phi_{max} = \sqrt{2\pi}\Delta R_p \cdot Ge_{max} \cdot \frac{5 \cdot 10^{20}}{\% \cdot cm^3} \tag{2.16}$$

The projected range R_p and straggle ΔR_p are given in fig. 2.7 [15, p. 155]. Φ_{max} and the critical Ge implantation dose are is shown in fig. 2.8 as a function of the implantation energy.

2.3 Band Structure of SiGe

Both Si and Ge possess an indirect fundamental band gap [25, 26]. Like for all semiconductors with a diamond structure, the maximum of the valence band of Si and Ge is situated in the

center of the Brillouin zone (Γ maximum) and consists of two converging subbands that are degenerated in the maximum.

In Si (see fig. 2.9 (a)), the minimum of the conduction band is situated in the (001) direction (Δ minimum). Due to the diamond structure of silicon, (100),($\bar{1}$00), (010), (0$\bar{1}$0),(001) and (00$\bar{1}$) are equivalent directions and the conduction band is sixfold degenerated. At 300 K, the indirect band gap of Si is 1.12 eV.

In Ge (see fig. 2.9 (b)), the minimum of the conduction band is situated in the (111) direction at the edge of the Brillouin zone (L minimum). Although there are 8 equivalent directions ($\overset{\pm\pm\pm}{1\,1\,1}$), the conduction band is fourfold degenerate because the eight semiellipses can be represented by 4 ellipses centered at the points (111), $\bar{1}$11), (1$\bar{1}$1) and (11$\bar{1}$). At 300 K, the indirect band gap of Ge is 0.66 eV.

The latest band structure calculations of unstrained SiGe alloys [1, chap. 4; 16; 17] use the molecular coherent potential approximation (MCPA) and nonlocal empirical pseudopotential calculations. The agreement between theory and experiment is excellent if the effect of alloy disorder is accounted for. In bulk alloys [1, chap. 4], the band gap decreases as a function of the Ge concentration. The conduction band structure is silicon-like for Ge concentrations up to 85 % and germanium-like for Ge concentrations over 85 % when the band gap decreases rapidly. The indirect energy gap of a $Si_{1-x}Ge_x$ alloy for $x < 0.85$ is given by [1, chap. 4.1] [1]:

$$E_g = 1.155 - 0.43x + 0.0206x^2 \ (eV) \tag{2.17}$$

And for $x > 0.85$:

$$E_g = 2.010 - 1.27x \ (eV) \tag{2.18}$$

The band gap of strained Si grown on a relaxed $Si_{1-x}Ge_x$ substrate is given by [29]:

$$E_g = 1.11 - 0.74x \ (eV) \tag{2.19}$$

The band gap of strained $Si_{1-x}Ge_x$ grown on a relaxed Si substrate for $x < 0.25$ can be described by [1, chap. 4.1]:

$$E_g = 1.17 - 0.896x + 0.396x^2 \ (eV) \tag{2.20}$$

The impact of strain on the band structure was first calculated by C. G. van de Walle and R. M. Martin [16, 17]. Their model-solid theory requires exclusively the knowledge of macroscopic parameters, such as strain. It is based on the the local-density-functional theory, the artificial introduction of a supercell that introduces translation invariance into the problem and nonlocal norm-conserving pseudopotentials. The local-density-functional method underestimates the band gaps of semiconductors severely, but is able to determine precisely the shift of the bands

[1] All values for the energy gaps, band offsets, and effective masses have been proposed by the Glasgow device modeling group. Estimated values for $Si_{1-y}Ge_y$ substrates can often be obtained by replacing x with $(x-y)$, e.g. $E_g = 1.17 - 0.896x - 0.396x^2$ becomes $E_g = 1.17 - 0.896(x-y) - 0.396(x-y)^2$.

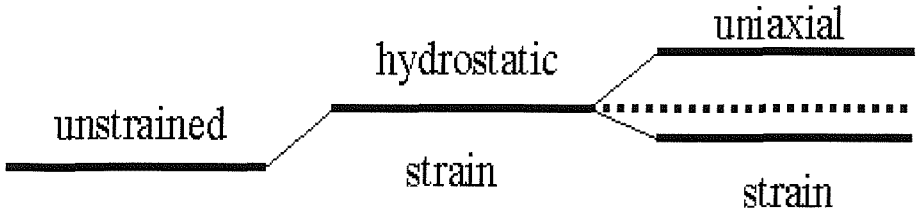


Figure 2.10: Schematic representation of the splitting of a triply degenerate band due to strain

due to strain. In combination with bulk band structure calculations, the calculations provide us with information about the valence and conduction band discontinuities without making any assumptions.

As strain has two main effects on the conduction and valence band structure of a SiGe alloy, the hydrostatic and uniaxial components of the strain are examined separately. The hydrostatic strain affects the relative position of the valence and conduction bands. The effect of uniaxial strain which splits the bands up is treated a posteriori. Fig. 2.10 illustrates the effect of strain on a triply degenerate band. The average position of the band is shifted by hydrostatic strain. In this particular case, the threefold degenerate band is split into a set of twofold degenerate bands at lower energy and a singly degenerate band at higher energy by uniaxial strain (like for the compressively strained conduction band of $Si_{1-x}Ge_x$). The average position of the band is not shifted by uniaxial strain (dotted line).

2.3.1 Effect of Strain on the Valence Band

Effect of Hydrostatic Strain The shift of the valence band due to hydrostatic strain is nearly independent on the interface orientation [1, chap. 4.2; 16, 17]. The shift of the average position of the valence band $\Delta E_{v,av}$ due to hydrostatic strain is given by:

$$\Delta E_{v,av} = a_v(2\epsilon_{\parallel} + \epsilon_{\perp}) \tag{2.21}$$

Where a_v is the hydrostatic deformation potential of the valence band. Material parameters for Si and Ge can be found in tab. 2.2. $\epsilon_{\parallel}$ is negative and $\epsilon_{\perp}$ positive for compressive strain (see eq.'s (2.3) and (2.4)). Appropriate values for $Si_{1-x}Ge_x$ can be obtained by linear interpolation [17].

Effect of Uniaxial Strain The shifts of the light hole valence band $E_{v,1}$, the heavy hole valence band $E_{v,2}$ and the split-off band $E_{v,3}$ under perpendicular uniaxial strain along [001] or [111] with respect to the average position of the valence band are given by:

$$\Delta E_{v,1} = -\frac{1}{6}\Delta_0 + \frac{1}{4}\delta E + \frac{1}{2}[\Delta_0^2 + \Delta_0\delta E + \frac{9}{4}(\delta E)^2]^{1/2} \tag{2.22}$$

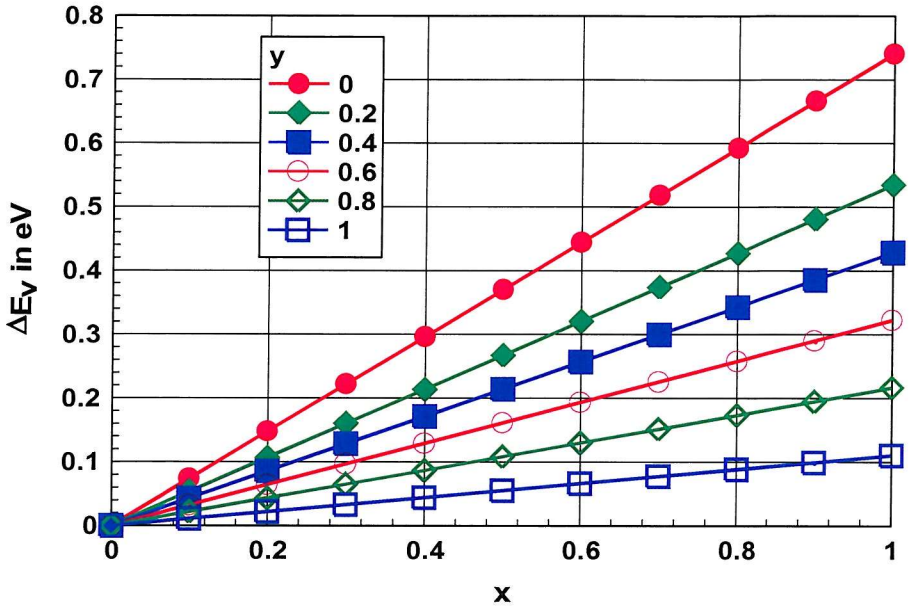


Figure 2.11: Valence band offset of strained $Si_{1-x}Ge_x$ grown on a relaxed $Si_{1-y}Ge_y$ substrate [28]

$$\Delta E_{v,2} = \frac{1}{3}\Delta_0 - \frac{1}{2}\delta E \tag{2.23}$$

$$\Delta E_{v,3} = \frac{1}{6}\Delta_0 + \frac{1}{4}\delta E - \frac{1}{2}[\Delta_0^2 + \Delta_0\delta E + \frac{9}{4}(\delta E)^2]^{1/2} \tag{2.24}$$

For perpendicular uniaxial strain along [001], δE is given by:

$$\delta E = 2g(\epsilon_\perp - \epsilon_\parallel) \tag{2.25}$$

For perpendicular uniaxial strain along [001], δE is given by:

$$\delta E = \frac{2}{3}\sqrt{3}d(\epsilon_\perp - \epsilon_\parallel) \tag{2.26}$$

Δ_o is the spin orbit splitting parameter, g the shear strain potential, and d the deformation potential. The case for uniaxial stress along [110] is more complicated and no analytical expression can be written down.

The valance band offset (ie the difference of the maximum of the valence band of the strained $Si_{1-x}Ge_x$ layer with respect to Si) of strained $Si_{1-x}Ge_x$ grown on a relaxed $Si_{1-y}Ge_y$ substrate can be described by [28]:

$$\Delta E_v = (0.74 - 0.53y)x \ (eV) \tag{2.27}$$

2.3.2 Effect of Strain on the Conduction Band

Effect of Hydrostatic Strain The shift of the conduction band due to hydrostatic strain is as well nearly independent on the interface orientation [1, chap. 4.3; 16, 17]. The shift of average

Parameter	in eV for Si	in eV for Ge
Δ_o	0.04	0.3
a_v	2.46	1.24
a_c^{dir}	1.98	-8.24
a_c^{ind}	4.18	-1.54
g	-2.35	-2.55
d	-5.32	-5.50
Ξ_u^{Δ}	9.16	9.42
Ξ_u^{L}	16.14	15.13

Table 2.2: Theoretical parameters for Si and Ge

position of the conduction band $\Delta E_{c,av}$ due to hydrostatic strain is given by:

$$\Delta E_{c,av} = a_c(2\epsilon_{\parallel} + \epsilon_{\perp}) \tag{2.28}$$

Different values for the the hydrostatic deformation potential a_c are used for the direct a_c^{dir} and indirect a_c^{ind} band gap.

Effect of Uniaxial Strain The energy shift of the conduction band valley i for a homogenous deformation can be written as:

$$\Delta E_c^i = [\Xi_d \overleftrightarrow{1} + \Xi_u\{\hat{a}_i\hat{a}_i\}] : \overleftrightarrow{\epsilon} \tag{2.29}$$

Where $\overleftrightarrow{1}$ is the unit tensor, $\hat{a}_i$ a unit vector parallel to the k-vector of valley i, and $\{\}$ denotes the dyadic product.

The shift of the mean energy of the conduction band extrema is given by:

$$\Delta E_{c,av}^o = (\Xi_d + \frac{1}{3}\Xi_u)\ \overleftrightarrow{1} : \overleftrightarrow{\epsilon} \tag{2.30}$$

Perpendicular uniaxial strain along [001] causes the bands along [100] and [010] to split off from the bands along [001]:

$$\Delta E_c^{001} = +\frac{2}{3}\Xi_u^{\Delta}(\epsilon_{\perp} - \epsilon_{\parallel}) \tag{2.31}$$

$$\Delta E_c^{100,010} = -\frac{1}{3}\Xi_u^{\Delta}(\epsilon_{\perp} - \epsilon_{\parallel}) \tag{2.32}$$

Under uniaxial strain along [110], the energy shifts are described by:

$$\Delta E_c^{001} = -\frac{1}{3}\Xi_u^{\Delta}(\epsilon_{\perp} - \epsilon_{\parallel}) \tag{2.33}$$

$$\Delta E_c^{100,010} = \frac{1}{6}\Xi_u^{\Delta}(\epsilon_{\perp} - \epsilon_{\parallel}) \tag{2.34}$$

The conduction bands at L are not affected by strain along [001], whereas strain along [111] leads to:

$$\Delta E_c^{111} = \frac{2}{3}\Xi_u^{L}(\epsilon_{\perp} - \epsilon_{\parallel}) \tag{2.35}$$

$$\Delta E_c^{\bar{1}11,1\bar{1}1,11\bar{1}} = -\frac{2}{9}\Xi_u^{L}(\epsilon_{\perp} - \epsilon_{\parallel}) \tag{2.36}$$

Parameter	Si	Ge
m_l/m_e	0.1905	0.082
m_t/m_e	0.9163	1.58

Table 2.3: Effective electron masses for bulk Si and bulk Ge [25, chap. 4.4]

Finally, the effect of strain along [110] is described by:

$$\Delta E_c^{111,11\bar{1}} = \frac{1}{3}\Xi_u^L(\epsilon_\perp - \epsilon_\parallel) \tag{2.37}$$

$$\Delta E_c^{\bar{1}11,1\bar{1}1} = -\frac{1}{3}\Xi_u^L(\epsilon_\perp - \epsilon_\parallel) \tag{2.38}$$

In ref. [29], the conduction band offset (i.e. the difference of the conduction band minimum of strained Si and relaxed $Si_{1-x}Ge_x$) for strained Si on a relaxed $Si_{1-x}Ge_x$ substrate is given by :

$$\Delta E_c = 0.67x \tag{2.39}$$

Valence bands and conduction bands in strained $Si_{1-x}Ge_x$ grown on Si(100) and in strained $Si_{1-x}Ge_x$ grown on Ge(100) are shown in fig. 2.12 and 2.13, respectively. The top of the valence band in Si is used as a reference point. Dashed lines indicate the average values for the valence bands and Δ conduction bands. Valence bands and conduction band values for any substrate can be found by interpolation.

2.4 Effective Masses of SiGe

The electron concentration n and the hole concentration p for nondegenerate semiconductors are given by [25, 26]:

$$n = 2(\frac{2\pi m_{e,DoS} kT}{h^2})^{3/2} \cdot \kappa \cdot \exp(-\frac{E_c - E_f}{kT}) \tag{2.40}$$

$$p = 2(\frac{2\pi m_{h,DoS} kT}{h^2})^{3/2} \exp(\frac{E_v - E_f}{kT}) \tag{2.41}$$

Where the density-of-state effective mass in the conduction band $m_{e,DoS}$ and the density-of-state effective mass in the valence band $m_{h,DoS}$ are given by [2]:

$$m_{e,DoS} = (m_l m_t^2)^{1/3} \tag{2.42}$$

$$m_{h,DoS} = (m_{lh}^{3/2} + m_{hh}^{3/2})^{2/3} \tag{2.43}$$

κ refers to the degeneracy of the conduction band ($\kappa = 6$ for Si and $\kappa = 8$ for Ge). The longitudinal effective electron mass and the transversal effective electron mass for Si and Ge are given in tab. 2.3. Only slight changes of the electron masses due to strain and Ge content are expected [1, chap. 4.4]. Appropriate values for $Si_{1-x}Ge_x$ may be obtained by interpolation.

The dependence of the heavy hole and light hole masses on the Ge fraction x has been investigated by the Glasgow device modeling group [30]. Their results are given in tab. 2.4 [3].

[2]The notation $m_{e,DoS} = (\kappa^2 m_l m_t^2)^{1/3}$ can also be found in literature. However, the resulting carrier concentrations are identical as equation (2.40) does not contain the factor 6 in this case.

[3]The values were proposed bu the Glasgow Modeling Group. They differ considerably from the values in

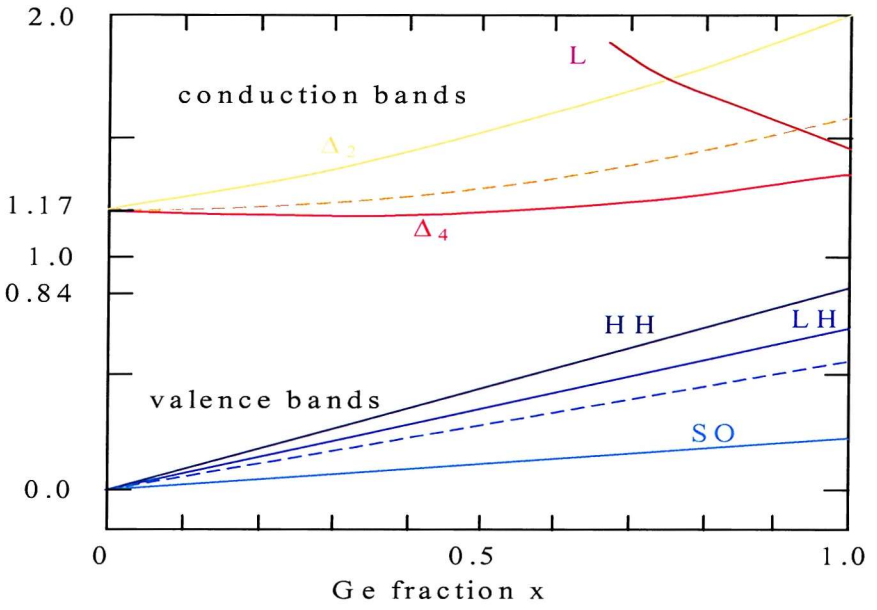


Figure 2.12: Valence and conduction bands in strained $Si_{1-x}Ge_x$ grown on Si(100) [16]

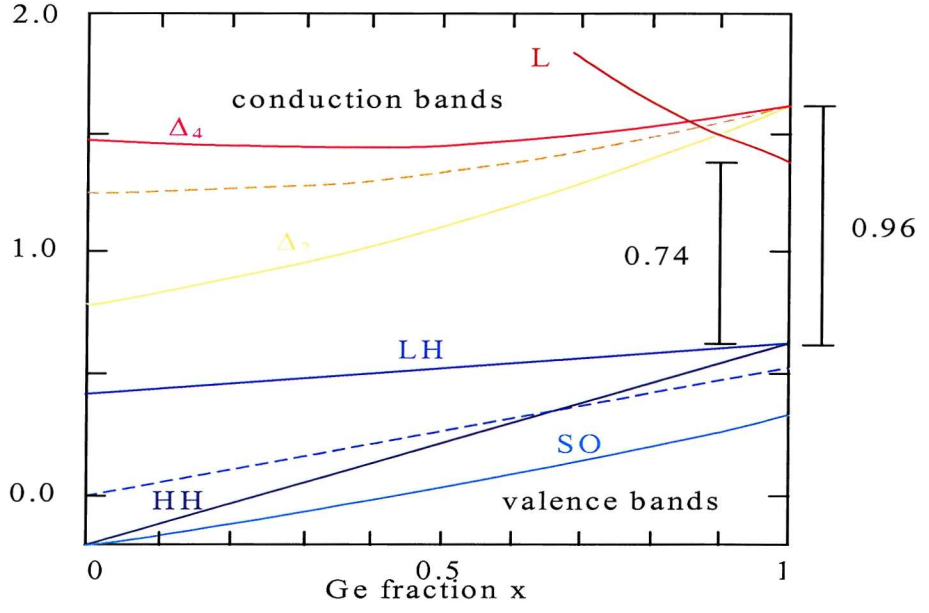


Figure 2.13: Valence and conduction bands in strained $Si_{1-x}Ge_x$ grown on Ge(100) [16]

m_{hh}/m_e	m_{lh}/m_e
bulk Si	
0.951 (0.537)	0.256 (0.153)
bulk Ge	
0.439 (0.284)	0.061 (0.044)
$Si_{1-x}Ge_x$	
$0.94 - 1.44x + 1.146x^2$	$0.251 - 0.509x + 0.396x^2$
strained Si on $Si_{1-x}Ge_x$	
$0.942x - 2.674x + 2.839x^2$	$0.259 + 0.536x - 0.28x^2$
strained $Si_{1-x}Ge_x$ on Si	
$0.927 - 2.266x + 1.827x^2$	$0.255 - 0.334x + 0.215x^2$

Table 2.4: Heavy hole and light hole masses for Si_xGe_{1-x} [1, chap 5.1]

2.5 Dielectric Constant

The dielectric constant ϵ is given by the Clausius-Mossotti relationship [31, 32]:

$$\frac{\epsilon_{Si_{1-x}Ge_x} - 1}{\epsilon_{Si_{1-x}Ge_x} + 2} = (1-x) \cdot \frac{\epsilon_{Si} - 1}{\epsilon_{Si} + 2} + x \cdot \frac{\epsilon_{Ge} - 1}{\epsilon_{Ge} + 2} \tag{2.44}$$

Where $\epsilon_{Si_{1-x}Ge_x}$ is the dielectric constant of $Si_{1-x}Ge_x$, $\epsilon_{Si} = 11.9$ the dielectric constant of Si, and $\epsilon_{Ge} = 16$ the dielectric constant of Ge. A dependence of the dielectric constant on strain is not known.

2.6 Mobility

The mobility [1, chap 5.1] is a function of the transport effective mass m^* and the transport scattering time τ_t:

$$\mu = \frac{q}{m^*}\tau_t \tag{2.45}$$

The transport scattering time can be decomposed according to Mathiessen's rule into different scattering times associated to a particular scattering mechanism:

$$\frac{1}{\tau_t} = \sum_i \frac{1}{\tau_i} \tag{2.46}$$

The mobility μ is limited by the scattering mechanism with the smallest relaxation time. In SiGe hetero structures, the main scattering mechanisms are [1, chap. 5.1]:

1. lattice scattering
2. ionized impurity scattering
3. neutral impurity scattering
4. alloy scattering
5. interface scattering

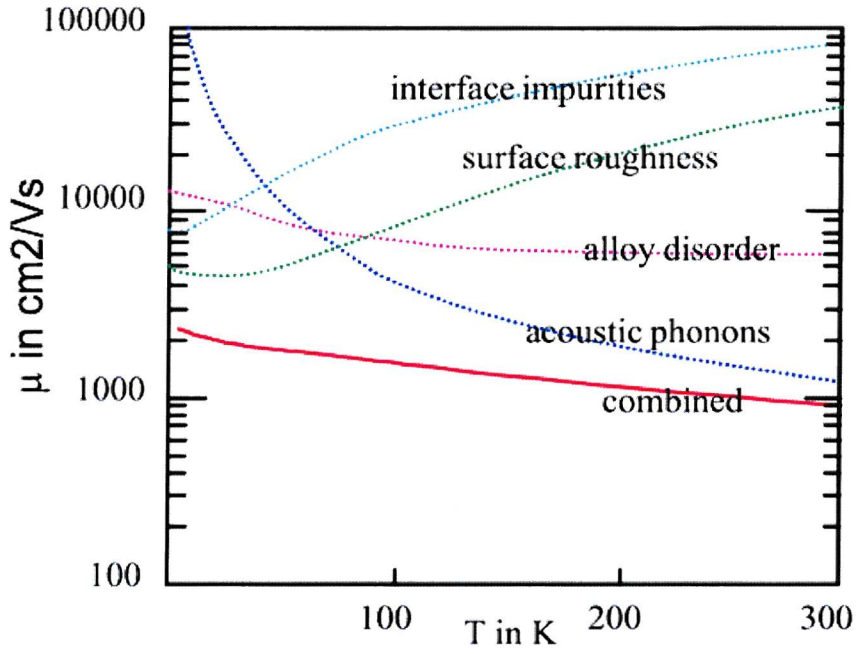


Figure 2.14: Theoretical Mobility as a function of temperature as the sum of different scattering mechanisms in a strained $Si_{0.8}Ge_{0.2}$ layer grown on a Si substrate [33]

Theoretical values for the combined hole mobility μ as a result of different scattering mechanisms are shown in fig. 2.14 for a strained $Si_{0.8}Ge_{0.2}$ layer grown on a Si substrate. Surface and interface scattering are the limiting mechanisms at low temperature, whereas the mobility is limited by acoustic phonon scattering at room temperature. The room temperature mobility is considerably higher than the mobility in Si (see tab. 2.5).

Whereas mobility data exists for pure bulk Si and Ge over a wide range of temperature and doping concentration, the mobility data available for SiGe bulk alloys and SiGe hetero structures is incomplete and often questionable Besides, experimentally determined mobilities depend on the measurement method, e.g. Hall mobility values and drift mobility values are not identical because of the impact of the magnetic field during the measurement. The theoretical treatment is in general more difficult for the hole mobilities than for the electron mobilities because of the complicated valence band structure.

In ref. [34], a monotonic increase of the hole mobility of a strained SiGe layer grown on a Si substrate is predicted the mobility in the SiGe heterostructure reaching the bulk Ge value for a Ge content of 40 %. But, theoretical predictions [35] are disputable because of the complexity of the calculations involved. Besides, the scattering mechanisms listed above are modified in hetero structures due to the reduced dimensionality of the carrier gas. Most experimental results [36-62] lack far behind the theoretical predictions. However, the discrepancy may also be attributed to

brackets found in [1]

	μ_e in cm^2/Vs	μ_h in cm^2/Vs
Si	1450	505
Ge	3900	1800

Table 2.5: Mobilities of intrinsic bulk Ge and Si at room temperature [1, chap. 5.1]

growth defects.

A 138 % hole mobility increase in a strained $Si_{0.64}Ge_{0.36}$ layer grown on Si is demonstrated in ref. [36] by our collaboration. Hole mobilities of more than 1500 cm^2/Vs have been reported for a strained Ge layer grown on a $Si_{0.3}Ge_{0.7}$ substrate [8]. The data for n-MOSFETs is as well very dispersed [63-71]. For strained Si grown on a relaxed $Si_{0.7}Ge_{0.3}$ substrate [8] an electron mobility in the order of 3000 cm^2/Vs has been found. For high electrical fields and short channel SiGe HMOSFETs [8, 72-75], velocity overshoot has been predicted and observed.

2.7 Ternary $Si_{1-x-y}Ge_xC_y$ Alloys

Only a few percent of carbon can be incorporated into $Si_{1-x}Ge_x$ until silicon carbide precipitates are formed. The lattice constant of carbon is $a_C = 0.77\mathring{A}$ and much smaller than the lattice constants of Si ($a_{Si} = 1.17\mathring{A}$) and Ge ($a_{Ge} = 1.22\mathring{A}$). Therefore, the incorporation of approximately one percent of carbon reduces the lattice mismatch between a $Si_{0.1y}Ge_{0.9}C_y$ layer and bulk Si to zero. Complete strain relieve and high thermal stability can be achieved for most $Si_{1-x-y}Ge_xC_y$ heterostructures [76-84].

Information about the band structure of $Si_{1-x-y}Ge_xC_y$ alloys is extremely sparse. In $Si_{1-x-y}Ge_xC_y$, higher electron and hole mobilities have been found than in Si.

The incorporation of carbon leads to an undersaturation of Si self-interstitials and a supersaturation of vacancies. B and P diffuse by an interstitial mechanism, whereas As and Sb diffuse by a vacancy mechanism. Therefore, the diffusion of B and P is suppressed in $Si_{1-x-y}Ge_xC_y$, whereas the the diffusion of As and Sb is enhanced [82; 83]. The outdiffusion of dopants from source and drain of MOSFETs degrades the short-channel performance and can be very severe due to interstitials created during implantation (transient enhanced diffusion, TED) [85-87]. A prospective application of $Si_{1-x-y}Ge_xC_y$ layers is the use as a barrier to B and P diffusion.

2.8 Conclusion

This second chapter provides a foundation for the whole thesis as it explains the physical properties of SiGe hetereostructures. SiGe heterostructures can be strained due to the lattice mismatch between Si and Ge. Strained SiGe layers can only be grown within a critical thickness before strain relaxation sets in. Strain alters the band structure and increases the carrier mobility.

So far, experimentally determined mobilities have not matched theoretically predicted values and the performance of HMOSFETs (hetero MOSFETs) has been rather disappointing. The reasons why it is so difficult to integrate SiGe heterolayers in HMOSFETs will be further explained in the following chapters before the electrical results for the devices fabricated in the course of this project are presented.

Bibliography

[1] Erich Kasper,"Properties of strained and relaxed Silicon-Germanium", INSPEC, 1995

[2] Suresh C. Jain, "Germanium-Silicon Strained Layers and Heterostructures", Academic Press, 1994

[3] U. Koenig, "Challenges for a Si/Ge heterodevice technology", Microelectronic Engineering, Vol. 23, pp. 3-13, 1994

[4] Friedrich Schaeffler, "High-mobility Si and Ge structures ", Semicond. Sci. Technol., Vol. 12, pp. 1515-1549, 1997

[5] Ulf Koenig, "Future Applications of Heterostructures", Physica Scripta, Vol. T68, pp. 90-101, 1996

[6] T. E. Whall, and E. H. C. Parker, "Review Silicon-germanium heterostructures- advanced materials and devices for silicon technology", Journal of Materials Science: Materials in Electronics, Vol. 6, pp. 249-264, 1995

[7] T. E. Whall and E. H. C. Parker, "Growth and characterisation of pseudomorphic Si/SiGe/Si heterostructures for p-channel field-effect transistors", Thin Solid Films, Vol. 306, pp. 338-345, 1997

[8] E. H. C. Parker and T. E. Whall, "SiGe heterostructure CMOS circuits and applications", Solid Stat. Electron., Vol. 43, pp. 1497-1506, 1999

[9] T. E. Whall and E. H. C. Parker, "Review Article SiGe heterostructures for FET applications", J. Phys. D: Appl. Phys., Vol. 31, pp. 1397-1416, pp. 1998

[10] K. L. Wang, S. G. Thomas and M. O. Tanner, "SiGe band engineering for MOS, CMOS, and quantum effect devices", J. Mat. Sci.: Mat. Electron., Vol. 6, pp. 311-324, 1995

[11] D. J. Paul, "Silicon germanium heterostructures in electronics: the present and the future", Thin Solid Films, Vol. 321, pp. 172-180, 1998

[12] C. K. Maiti, L. K. Bera and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors", Semicond. Sci. Technol., Vol. 13, pp. 1225-1246, 1998

[13] A. Sadek and K. Ismail, "Si/SiGe CMOS Possibilities", Solid-State Electronics, Vol. 38, No. 9, pp. 1731-1734, 1995

[14] T. Araki, N. Fujimura, and T. Ito, "The stability of ordered structures in SiGe films examined by strain-energy calculations", Journal of Crystal Growth, Vol. 174, pp. 675-679, 1997

[15] F. Christiano, "Extended defects in SiGe device structures formed by ion implantation", Ph.D. Thesis, University of Surrey, 1998

[16] C. G. van de Walle and R. M. Martin, "Theoretical calculations of heterojunction discontinuities in the Si/Ge system", Physical Review B, Vol. 34, No. 8, pp. 5621-5634, 1986

[17] C. G. van de Walle, "Physics of heterojunctions", Summer Course on Advanced Semiconductor Physics for Microdevices, Leuven, 1988

[18] D. Hull, and D. J. Bacon, "Introduction to dislocations", Pergamon Press, 3rd edition, 1984

[19] J. H. van der Merwe, "Strain relaxation in epitaxial overlayers", J. Electron. Mat., Vol. 200, No. 10, pp. 793-803, 1991

[20] J. W. Matthews, and A. E Blakeslee, "Defects in epitaxial multilayers", J. Crystal Growth, Vol. 27, pp. 118-125, 1974

[21] J. W. Matthews, "Defects associated with the accomodation of misfit between crystals", J. Va. Sci. Technol., Vol. 12, No. 1, pp.126-133, 1975

[22] Le Thanh Vinh, V. Aubry-Fortuna, Y. Zheng, D. Bouchier, C. Guedj and G. Hencelin, "UHV-CVD heteroepitaxial growth of $Si_{1-x}Ge_x$ alloys on Si (100) using silane and germane", Thin Solid Films, Vol. 294, pp. 59-63, 1997

[23] D. C. Paine, D. J. Howard, and N. G. Stoffel, "Strain Relief in Compositionally Graded $Si_{1-x}Ge_x$ Formed by High Dose Ion Implantation", J. Electronic Mat., Vol. 20, No. 10, pp. 735-746, 1991

[24] M. Mitchell, "SiGe Heterojunction Bipolar Transistors Fabricated by Germanium Ion Implantation", Minithesis, June 1999

[25] Henry Mathieu, "Physique des semiconducteur et des composants électroniques", Masson, 1996

[26] S. M. Sze, "Physics of Semiconductor Devices", John Wiley & Sons, 1981

[27] M. M. Rieger and P. Vogl, "Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates", Phys. Rev. B, Vol. 48, No. 19, pp. 14276-14287, 1993

[28] R. People, "Physics and Applications of Ge_xSi_{1-x}/Si Strained-Layer Heterostructures", IEEE J. Quantum Electronics, Vol. QE-22(9), pp. 1697-1710, 1986

[29] G. F. Formicone, D. Vasileska, and D. K. Ferry, "Modeling of submicron $Si1 - xGex$-based MOSFETs by self-consistent Monte Carlo", Phys. Stat. Sol. B, Vol. 204, Vol. 1, pp. 531-533, 1997

[30] J. R. Watling, A. Asenov, and J.R. Barker, Extended Abstracts of IWCE-6, No. 98EX116, p. 96, Osaka, 1998

[31] J. W Harrison and J. R. Hauser, "Theoretical calculations of electron mobility in ternary III-V compounds", J. Appl. Phys., Vol. 47, No. 1, pp. 292-300, 1976

[32] B. B. Pal, "Compositional dependence of permittivity in quarternary III-V semiconductor compounds", Sol.-St. Electron., Vol. 28, No. 12, pp. 1235-1239, 1985

[33] M. J. Kearney and A. I. Horrell, "The effect of alloy scattering on the mobility of holes in a $Si_{1-x}Ge_x$ quantum well", Semicond. Sci. Technol., Vol. 13, pp. 174-180, 1998

[34] J. M. Hinckley, and J. Singh, "Hole transport in pseudomorphic $Si_{1-x}Ge_x$ alloys grown on Si(001) substrates", Phys. Rev. B, Vol. 41, pp. 2912-2926, 1990

[35] Y. Zhang, and J. Singh, "Channel effective mass and interfacial effects in Si and SiGe metal-oxide semiconductor field effect transistor: A charge control model study", Journal of Applied Physics, Vol. 83, No. 8, pp. 4264-4271, 1998

[36] M. J. Palmer, G. Braithwaite, T. J. Grasby, P. J. Philips, M. J. Prest, E. H. C. Parker, T. E. Whall, C. P. Parry, A. M. Waite, A. G. R. Evans, S. Roy, J. R. Watling, S. Kaya, and A. Asenov, "Effective mobilities in pseudomorphic Si/SiGe/Si p-channel metal-oxide-semiconductor field-effect transistors with thin silicon capping layers", Appl. Phys. Lett., Vol. 78, No. 10, pp. 1424-1426, 2001

[37] S. Verdonckt-Vandebroek, E. F. Crabbe, B. S. Meyerson, D. L. Harame, P. S. Restle, J. M. C. Stork, A. C. Megdanis, C. L. Stanis, A. A. Bright, G. M. W. Kroesen, and A. C. Warren , "High-Mobility Modulation-Doped Graded SiGe-Channel p-MOSFET'S", IEEE Electron Device Letters, Vol. 12, No. 8, pp. 447-449, August 1991

[38] T. Manku, and A. Nathan, "Effective mobility in p-channel Si-SiGe metal oxide semiconductor field effect transistors (MOSFET's), Can. J. Phys, Vol. 70, No. 10-11, pp. 959-962, 1992

[39] T. Manku, J. M. McGregor, A. Nathan, D. J. Roulston, J.-P. Noel, and D. C. Houghton, "Drift Hole Mobility in Strained and Unstrained Doped $Si_{1-x}Ge_x$ Alloys", IEEE Transactions on Electron Devices, Vol. 40, No. 11, pp. 1990-1996, 1993

[40] T. Manku, and A. Nathan, "Lattice Mobility of Holes in Strained and Unstrained $Si_{1-x}Ge_x$ Alloys", IEEE Electron Device Letters, Vol. 12, No. 12, pp. 704-706, 1991

[41] T. E. Whall, "Fully pseudomorphic Si/SiGe/Si heterostructures for p-channel field effect devices", Thin Solid Films, Vol. 294, pp. 160-165, 1997

[42] U. Koenig and F. Schaeffler, "p-Type Ge channel MODFET's with High Transconductance Grown on Si Substrate", IEEE Electron Device Lett., Vol. 14, No. 4, pp. 205-207, 1993

[43] U. Koenig, M. Glueck, and G. Hoeck , "Si/SiGe Field-Effect Transistors", J. Vac. Sci. Technol. B, Vol. 16, No. 5, pp. 2609-2614, 1998

[44] C. K. Maiti, L. K. Bera, S. S. Dey, D. K. Nayak, and N. B. Chakrabarti, "Hole Mobility Enhancement in Strained-Si p-MOSFETs under high vertical field", Solid-St. Electron., Vol. 41, No. 12, pp. 1863-1869, 1997

[45] R. J. P. Lander, C. J. Emeleus, B. M. McGregor, E. H. C Parker, T. E. Whall, A. G. R. Evans, and G. P. Kennedy, "Study of Hall and effective mobilities in pseudomorphic $Si_{1-x}Ge_x$ p-channel metal-oxide-semiconductor field-effect-transistors at room temperature and 4.2 K", J. Appl. Phys., Vol. 82, No. 9, pp. 1-7, 1997

[46] W.-C. Tsai, C.-Y. Chang, G.-W. Huang, F.-F. Fang, Y.-H. Chang, and C.-F. Huang, "Very High Hole Mobility in P-Type Si/SiGe Modulation-Doped Heterostructures", Jpn. J. Appl. Phys., Vol. 36, Part 2, No. 3B, pp. L323-L326, 1997

[47] R. J. P. Lander, C. J. Emeleus, B. M. McGregor, E. H. C Parker, T. E. Whall, A. G. R. Evans, and G. P. Kennedy, "Study of Hall and effective mobilities in pseudomorphic $Si_{1-x}Ge_x$ p-channel metal-oxide-semiconductor field-effect-transistors at room temperature and 4.2 K", J. Appl. Phys., Vol. 82, No. 9, pp. 1-7, 1997

[48] A. Sadek, K. Ismail, M. A. Armstrong, D. A. Antoniadis, and F. Stern, "Design of Si/Ge Heterojunction Complementary Metal-Oxide-Semiconductor Transistors", IEEE Transactions on Electron Devices, Vol. 43, No. 8, pp. 1224-1232, August 1996

[49] M. Glueck, U. Koenig, W. Winter, K. Brunner, and K. Eberl, "Modulation-doped $Si_{1-x-y}Ge_xC_y$ p-type Hetero-FETs", Conference on Modulated Semiconductor Structures, 1997

[50] K. Ismail, J. O. Chu, and B. S. Meyerson, "High hole mobility in SiGe alloys for device applications", Appl. Phys. Lett., Vol. 64, No. 23, pp. 3124-3126, 1994

[51] G. Hoeck, M. Glueck, T. Hackbarth, H.-J. Herzog, and E. Kohn, "Carrier mobilities in modulation doped $Si_{1-x}Ge_x$ heterostructures with respect ot FET applications", Thin Solid Films, Vol. 336, pp. 141-144, 1998

[52] B. R. Cyca, K. G. Robins, N. G. Tarr, D. X. Xu, J.-P. Noel, D. Landheed, and M. Simard-Normadin, "Hole confinement and mobility in heterostructure Si/Ge/Si p-channel metal-oxide-semiconductor field effect transistors", J. Appl. Phys., Vol. 81, No. 12, pp. 8079-8083, 1997

[53] K. Ismail, "Si/SiGe High-Speed Field-Effect Transistors", IEDM, pp. 509-512, 1995

[54] D. K. Nayak, K. Goto, A. Yutani, J. Murota and Y. Shiraki, "High-Mobility Strained-Si PMOSFET's", IEEE Trans. Electron. Dev., Vol. 43, No. 10, pp. 1709-1716,1996

[55] Y. H. Wu and A. Chin, "High Temperature Formed SiGe P-MOSFET's with Good Device Characteristics", IEEE Electron Device Lett., Vol. 21, No. 7, pp. 350-352, 2000

[56] R. M. Sidek, U. N. Straube, A. M. Waite, A. G. R. Evans, C. Parry, P. Phillips, T. E. Whall, and E. H. C. Parker, "SiGe CMOS fabrication using SiGe MBE and anodic/LTO gate oxide", Semicond. Sci. Technol., Vol. 15, pp. 135-138, 2000

[57] G. Hoeck, T. Hackbarth, U. Erben, E. Kohn, and U. Koenig, " High performance 0.25 μm p-type Ge/SiGe MODFETs", Electronics Lett., Vol. 34, No. 19, pp. 1888-1889, 1998

[58] S. P. Voinigescu, C. A. T. Salama, J.-P. Noel, and T. I Kamins, "Optimized Ge channel Profiles for VLSI Compatible Si/SiGe p-MOSFET's", IEDM Tech. Digest, pp. 369-372, 1994

[59] C. Pawlowicz, N. G. Tarr, L. P. Berndt, R. L. Williams, D. Landheer, D.-X. Xu, R. Abid, and S. P. McAlister, " Heterostructure $Si_{1-x}Ge_x$ channel p-MOSFETs with high Ge concentration", J. Vac. Sci. Technol. A, Vol. 16, No. 2, pp. 864-867, 1998

[60] R. Oberhuber, G. Zandler, and P. Vogl, "Subband structure and mobility of two-dimensional holes in strained Si/SiGe MOSFET's", Physical Review, Vol. B 58, No. 15, pp. 9941-9948, 1998

[61] H. Jiang and R. G. Elliman, "Electrical Properties of GeSi Surface- and Buried-Channel p-MOSFET's Fabricated by Ge Implantation", IEEE Transactions on Electron Devices, Vol. 43, No. 1, pp. 97-103, 1996

[62] M. J. Kearney and A. I. Horrell, "An analysis of the mobility of holes in p-type SiGe quantum wells using multiple scattering theory", Semicond. Sci. Technol., Vol. 14, pp. 211-214, 1999

[63] T. Ostermann, M. Glueck, R. Hagelauer, U. Koenig, M. Dobersberger, M. Birk, and F. Oehler, "Inverter Circuits with $Si/SiGe$ n-type MODFETs", ISDR, 1997

[64] M. Glueck, T. Hackbarth, U. Koenig, A. Hass, G. Hoeck, and E. Kohn, " High f_{max} n-type Si/SiGe MODFETs", Electronics Letters, Vol. 33, No. 4, pp. 335-336, 1997

[65] V. I. Kuznetsov, K. Werner, S. Radelaar, and J. W. Metselaar, "Characterisation of an n-type Si/SiGe modulation doped field-effect transistor", Thin Solid Films, Vol. 294, pp. 263-266, 1997

[66] M. Glueck, T. Hackbarth, U. Koenig, A. Hass, G. Hoeck, and E. Kohn, " High f_{max} n-type Si/SiGe MODFETs", Electronics Letters, Vol. 33, No. 4, pp. 335-336, 1997

[67] A. C. Churchill, D. J. Robbins, D. J. Wallis, N. Griffin, D. J. Paul, and A. J. Pidduck, "High-mobility two-dimensional electron gases in Si/SiGe heterostructures on relaxed SiGe layers grown at high temperature", Semicond. Sci. Technol., Vol. 12, pp. 943-946, 1997

[68] C.-L. Huang and G. S. Gildenblat, "Measurements and Modeling of the n-Channel MOSFET Inversion Layer Mobility and Device Characteristics in the Temperature Range 60-300 K", IEEE Transactions on Electron Devices, Vol. 37, No. 5, pp. 1289-1300, 1990

[69] S. F. Nelson, K. Ismail, T. N. Jackson, J. J. Nocera, and J. O. Chu, "Systematics of electron mobililty in Si/SiGe heterostructures", Appl. Phys. Lett., Vol. 63, No. 6, pp. 794-796, 1993

[70] R. S. Prasad, T. J. Thornton, S. Kanjanachuchai, J. Fernandez, and A. Matsumura, "Mobility degradation in gated Si:SiGe quantum wells with thermally grown oxides", Electronics Lett., Vol. 31, No. 21, pp. 1876-1878, 1995

[71] N. Griffin, D. J. Paul, M. Pepper, S. Taylor, J. P. Smith, W. Eccleston, J. M. Fernandez, and B. A. Joyce, "Gating High Mobility Silicon-Germanium Heterostructures", Microelectron. Eng., Vol. 35, pp. 309-312, 1997

[72] F. Assaderaghi, D. Sinitsky, J. Bokor, P. K. Ko, H. Gaw, and C. Hu, "High-Field Transport of Inversion-Layer Electrons and Holes Including Velocity Overshoot", IEEE Trans. Electron Devices, Vol. 44, Vol. 4, pp. 664-671, 1997

[73] K Bhaumik, Y. Shacham-Diamand, J.-P. Noel, J. Bevk, and L. C. Feldman, "Theory and Observation of Enhanced, High Field Hole Transport in $Si_{1-x}Ge_x$ Quantum Well p-MOSFET's", IEEE Trans. Electron Devices, Vol. 43, Vol. 11, pp. 1965-1971, 1996

[74] S. Kaya, Y. P Zhao, J. R. Watling, A. Asenov. J. R. Barker, G. Ansaripour, G. Braithwaite, T. E. Whall, and E. H. C. Parker, "Indication of velocity overshoot in strained $Si_{0.8}Ge_{0.2}$ p-channel MOSFETs", Semicond. Sci. Technol., Vol. 15, No. 6, pp. 573-578, 2000

[75] Y. P. Zhao, J. R. Watling, S. Kaya, A. Asenov, and J. R. Barker, "Drift diffusion and hydrodynamic simulations of Si/SiGe p-MOSFETs", Mat. Sci. Eng., Vol. B72, pp. 180-183, 2000

[76] H. J. Osten, M. Kim, G. Lippert, and P. Zaumseil, "Ternary SiGeC alloys: growth and properties of a new semiconducting material", Thin Solid Films, Vol. 294, pp. 93-97, 1997

[77] K. Eberl, K. Brunner, and W. Winter,"Pseudomorphic $Si_{1-y}C_y$ and $Si_{1-x-y}Ge_xC_y$ alloy layers on Si", Thin Solid Films, Vol. 294, pp. 98-104, 1997

[78] S. K. Ray, D. W. McNeill, D. L. Gay, C. K. Maiti, G. A. Armstrong, B. M. Armstrong and H. S. Gamble, "Comparison of $Si_{1-y}C_y$ films produced by solid-phase epitaxy and rapid thermal chemical vapour deposition", Thin Solid Films, pp. 149-152, 1997

[79] A. St. Amour, L. D. Lanzerotti, C. L. Chang, and J. C. Sturm, "Optical and electrical properties of $Si_{1-x-y}Ge_xC_y$ thin films and devices", Thin Solid Films, Vol. 294, pp. 112-117, 1997

[80] R. Duschl, H. Seeberger, and K. Eberl, "Hole mobilities in pseudomorphic $Si_{1-x-y}Ge_xC_y$", Thin Solid Films, Vol. 336, pp. 336-339, 1998

[81] A. C. Mocuta and D. W. Greve, "$Si_{1-x-y}Ge_xC_y$-Channel p-MOSFET's with Improved Thermal Stability", IEEE Electron Dev. Lett., Vol. 21, No. 6, pp. 292-294, 2000

[82] R. Butz and H. Lueth, "The influence of carbon on the surface morphology fo Si (100) and on subsequent Ge island formation", Thin Solid Films, Vol. 336, pp. 69-72, 1998

[83] H. Ruecker, B. Heinemann, D. Bolze, D. Knoll, D. Krueger, R. Kurps, H. J. Osten, P. Schley, B. Tillack, and P. Zaumseil, "Dopant Diffusion in C-doped Si and SiGe: Physical Model and Experimental Verification", IEDM Conference Proceedings, 1999

[84] H. Ruecker and B. Heinemann, "Tailoring dopant diffusion for advanced SiGe:C heterojunction bipolar transistors", Sol.-St. Electron., Vol. 44, pp. 783-789, 2000

[85] A. Nejim, N. P. Barradas, C. Jeynes, F. Cristiano, E. Wendler, K. Gaertner, and B. J. Sealy, "Residual post anneal damage of Ge and C co-implantation of Si determined by quantitative RBS-channeling", Nucl. Instr. and Meth. in Phys. Res. B 139, pp. 244-248, 1998

[86] F. Christiano, C. Bonafos, A. Nejim, S. Lombardo, M. Omri, D. Alquier, A. Martinez, S. U. Campisano, P. L. F. Hemment, and A. Claverie,"Interstitial trapping efficiency of C^+ implanted into preamorphised silicon - control of EOR defects", Nucl. Instr. and Meth. in Phys. Res. B 127 / 128, pp. 22-26, 1997

[87] M. Yang, C.-L. Chang, M. Carroll, and J. C. Sturm, "25-nm p-Channel Vertical MOSFET's with SiGeC Source-Drains", IEEE Electron Dev. Lett., Vol. 20, No. 6, pp. 301-303, 1999

Chapter 3

MOSFET Theory

The third chapter reviews the theory of MOSFETs. HMOSFETs and conventional MOSFETs display qualitatively an identical behavior. However, quantitative drain current equations can in general only be given for conventional MOSFETs.

Furthermore, the behavior of modern MOSFETs with a short channel and ultrathin gate oxide is studied. The classical MOSFET equations describe the behavior of modern MOSFETs with a short channel and ultrathin gate oxide only approximately.

Finally, the use of strained SiGe layers for HMOSFETs and the impact of strained SiGe layers on the MOSFET characteristics is discussed. SiGe heterojunctions can be used in the channel, the gate, and the source and drain. A strained SiGe channel can enhance the drain current drive with respect to conventional MOSFETs due to superior transport properties. In source and drain, SiGe can improve the short channel characteristics and reduce the source and drain resistance. Poly-SiGe as a gate material allows the adjustment of the threshold voltage and the reduction of the gate resistance.

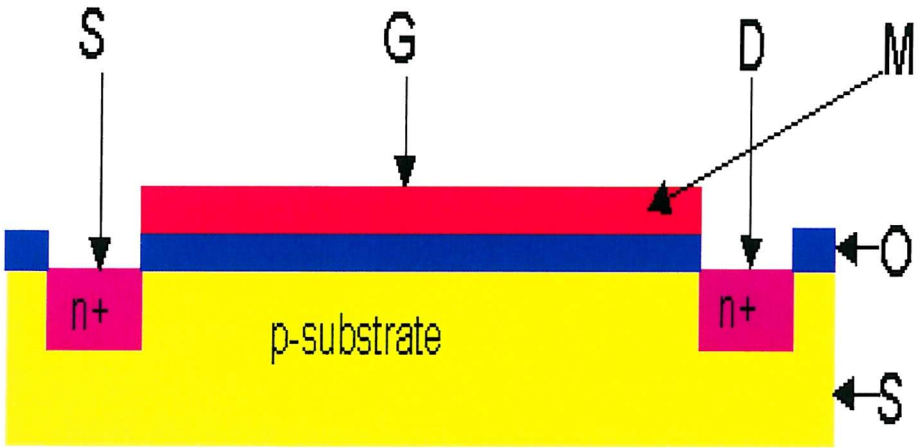


Figure 3.1: Classical n-MOSFET

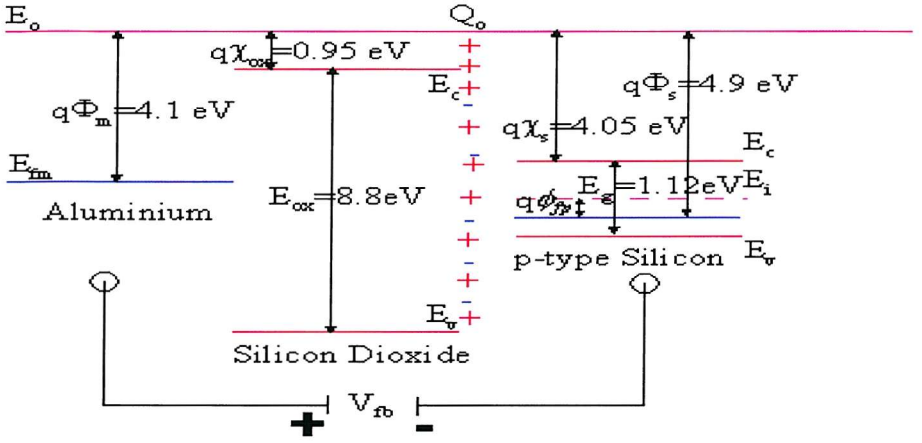


Figure 3.2: MOSFET structure without band bending for Al, thermally grown SiO_2 and $N_a = 10^{15} cm^{-3}$

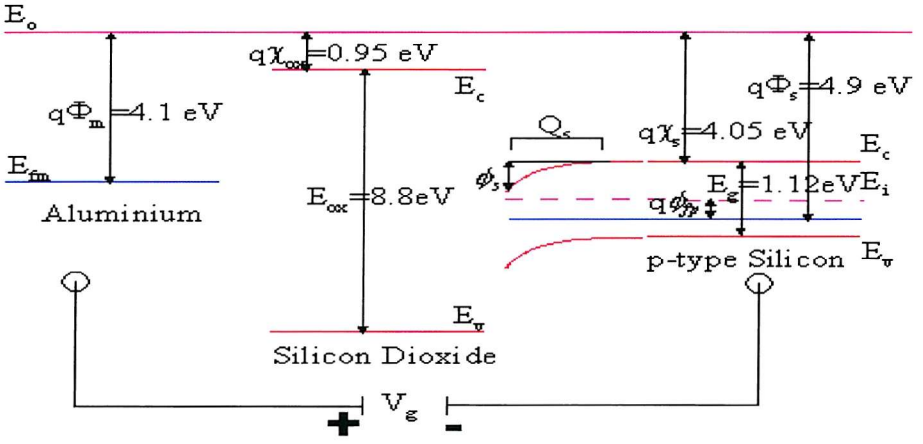


Figure 3.3: The same MOSFET structure for an applied gate voltage V_g without interface charge

3.1 Gate Voltage

The charge neutrality condition requires that the total charge at the gate surface Q_g, in the oxide and in the semiconductor space charge region Q_s is zero. The oxide charge and interface charge are represented by the effective interface charge Q_o [1] [1, p. 133]:

$$Q_g + Q_o + Q_s = 0 \tag{3.1}$$

Whereas the effective interface charge is always fixed, the charges at the gate and semiconductor surface are determined by the applied voltage.

Electrons and holes possess a different potential energy in metal and semiconductor. In a metal, the work function Φ_m is defined as the energy that must be supplied to an electron to take it across the surface energy barrier (i.e. energy difference between the vacuum energy E_o

[1] The charge and capacitance refer in the MOSFET theory always to a normalized area.

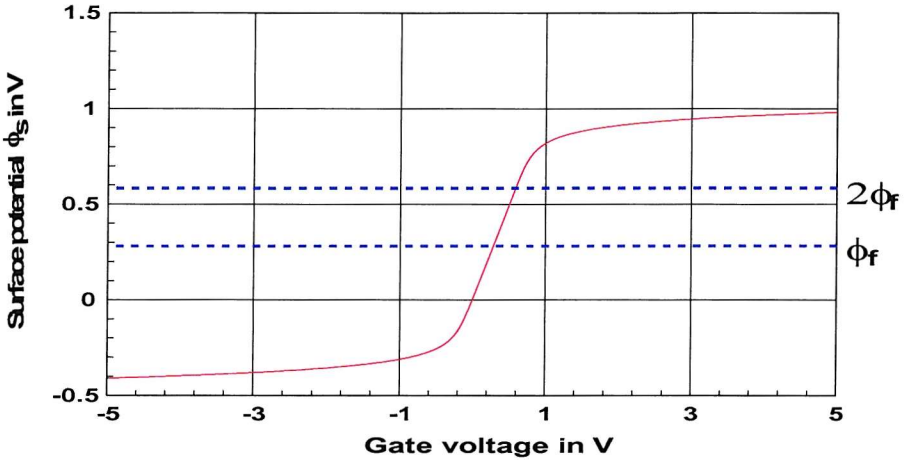


Figure 3.4: Surface potential ϕ_s as a function of the gate voltage V_g for $V_{fb} = 0\ V$ $N_b = 1 \cdot 10^{15}/cm^3$, $T = 300\ K$, and a 6 nm thick gate oxide

and the Fermi energy of the metal E_{fm}). [2]

In a semiconductor, the height of the surface barrier is described by the electron affinity χ_s. The electron affinity is defined as the difference in the vacuum energy and the conduction band edge at the surface and is a property of the semiconductor whereas the work function of the semiconductor depends on the doping concentration. Neglecting the temperature dependence of the Fermi level, the work function of a p-type and n-type semiconductor respectively is given by ([1, p. 124], see fig. 3.2):

$$\Phi_s \simeq \chi_s + E_g/2 + q\phi_{fp} \tag{3.2}$$

$$\Phi_s \simeq \chi_s + E_g/2 - q\phi_{fn} \tag{3.3}$$

ϕ_{fn} and ϕ_{fp} are the the Fermi potentials ϕ_f of electrons and holes when the intrinsic Fermi level is taken as the zero reference level as it is often done in literature [1, 2] and q is the elementary charge. ϕ_f can be calculated according to:

$$\phi_f = V_t \ln \frac{N_b}{n_i} \tag{3.4}$$

where V_t is the thermal voltage ($V_t = kT/q$), N_b the net bulk doping (ie the magnitude of the difference between the concentration of acceptors N_a and donors N_d; $N_b = |N_a - N_d|$) and n_i the intrinsic carrier concentration.

In thermal equilibrium, the bias between gate and substrate is determined by the potential difference between gate and substrate and the effective interface charge. This bias is called the flat band voltage because it corresponds to the bias opposite in sign that is necessary in order to flatten the surface barrier. Thus, the following relation must be valid (see fig. 3.2):

$$\Phi_m = \Phi_s + \frac{Q_o}{C_{ox}} + V_{fb} \tag{3.5}$$

[2] Here we still use the theory for a metal gate. This is the usual approach in books, although polysilicon gates are more common today.

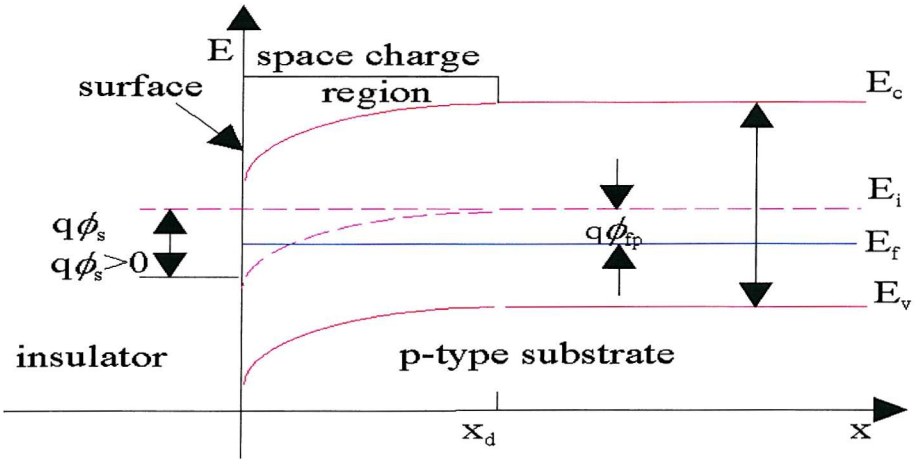


Figure 3.5: Space charge region

Using $\Phi_{ms} = \Phi_m - \Phi_s$, we get the flat band voltage V_{fb} [1, p. 131]:

$$V_{fb} = \Phi_{ms} - \frac{Q_o}{C_{ox}} \tag{3.6}$$

In the MOSFET theory, the interface charge is neglected. Then, the magnitude of the charge at the gate surface $Q_g = \epsilon_o \epsilon_{ox} E_{ox}$ and the magnitude of the charge in the semiconductor space charge region $Q_s = -\epsilon_o \epsilon_{si} E_{si}$ must be identical (Where ϵ_o is the vacuum permittivity, ϵ_{ox} the dielectric constant of SiO_2, ϵ_{si} the dielectric constant of silicon, E_{ox} the electric field at the oxide interface in SiO_2, and E_{si} the electric field at the oxide interface in silicon.).
An applied external bias V_g (see fig. 3.3) is shared between the voltage across the oxide V_{ox}, the surface potential ϕ_s, and the work function difference Φ_{ms}:

$$V_g = V_{ox} + \phi_s + \Phi_{ms} \tag{3.7}$$

Using $E_{ox} = V_{ox}/t_{ox}$ (Where t_{ox} is the oxide thickness.), the gate voltage can be expressed as [1, p. 134]:

$$V_g = V_{fb} + \phi_s - \frac{Q_s}{C_{ox}} \tag{3.8}$$

3.1.1 Space Charge in the MOS Structure

The space charge Q_s is calculated solving the Poisson equation in the space charge region [1, p. 38; 2, p. 278] ($0 \leq x \leq x_d$; see fig. 3.5). Using the charge neutrality in the bulk, the Poisson equation for the electrical potential ϕ becomes:

$$\frac{d^2\phi}{dx^2} = -\frac{q\rho(x)}{\epsilon_o \epsilon_{si}} = -\frac{q(N_d(x) - N_a(x) + p(x) - n(x))}{\epsilon_o \epsilon_{si}} \tag{3.9}$$

The electron density $n(x)$ and hole density $p(x)$ are given by:

$$n(x) = n_i \exp\left[(\phi(x) - \phi_f(x)/V_t)\right] = n_o(x) \exp\left[(\phi(x)/V_t)\right] \tag{3.10}$$

$$p(x) = n_i \exp\left[-(\phi(x) - \phi_f(x))/V_t\right] = p_o(x) \exp\left[-(\phi(x)/V_t\right] \tag{3.11}$$

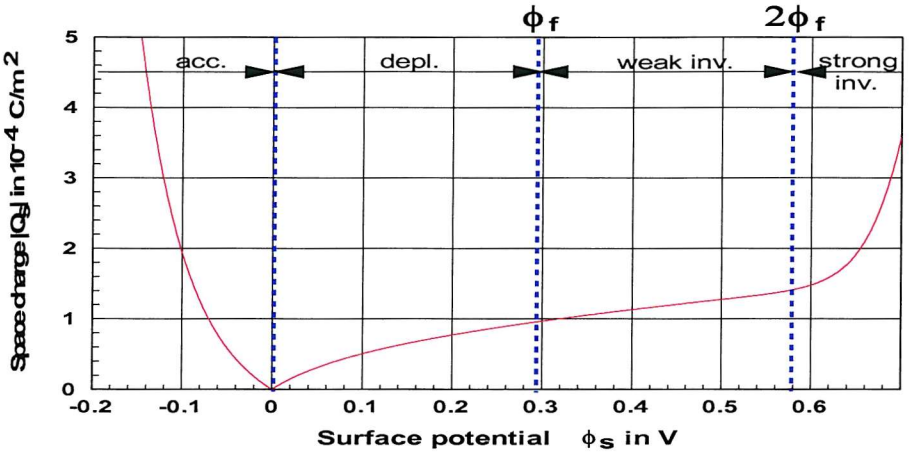


Figure 3.6: Space charge Q_s for $N_b = 1 \cdot 10^{15}/cm^3$ at $T = 300\ K$

The dopant profile is often uniform ($n_o(x) = n_o$, $p_o(x) = p_o$).
Using the boundary conditions at $x = 0$:

$$\phi(0) = \phi_s, \frac{d\phi(0)}{dx} = -E(0) \tag{3.12}$$

And at the end of the space charge region ($x = x_d$):

$$\phi(x_d) = \frac{d\phi(x_d)}{dx} = 0 \tag{3.13}$$

The electrical field becomes [2, p. 279]:

$$E(x) = Sign(\phi_s)\frac{V_t}{L_D}F(\phi(x)) \tag{3.14}$$

Where L_D is the extrinsic Debye length for holes:

$$L_D = (\frac{\epsilon_o \epsilon_{si}}{2qV_t p_o})^{1/2} \tag{3.15}$$

The *Sign*-function gives us the sign of the surface potential at the semiconductor surface. $F(\phi(x))$ is given by:

$$F(\phi(x)) = [(\exp(-\phi(x)/V_t) + \frac{\phi(x)}{V_t} - 1) + \frac{n_o}{p_o}(\exp(\phi(x)/V_t) - \frac{\phi(x)}{V_t} - 1)]^{1/2} \tag{3.16}$$

At the semiconductor surface ($\phi(0) = \phi_s$), equation (3.16) becomes:

$$F(\phi_s) = (\exp(-\phi_s/V_t) + \frac{\phi_s}{V_t} + \exp(\phi_s - 2\phi_f/V_t) - 1 - \frac{n_o}{p_o}(\frac{\phi_s}{V_t} + 1))^{1/2} \tag{3.17}$$

Using the Gaussian law at the semiconductor surface:

$$\int \vec{E}d\vec{s} = Q_s/\epsilon_o\epsilon_{si} \tag{3.18}$$

We get the total charge (see fig. 3.6):

$$Q_s = -Sign(\phi_s)\frac{\epsilon_o\epsilon_{si}V_t}{L_D}F(\phi_s) \tag{3.19}$$

Applying equation (3.4), the electron and hole density can be written as:

$$n(x) = N_b \exp\left(\phi(x) - 2\phi_f/V_t\right) \tag{3.20}$$

$$p(x) = N_b \exp\left(-\phi(x)/V_t\right) \tag{3.21}$$

Thus, the space charge can be expressed as [1, p. 145]:

$$Q_s = -\sqrt{2\epsilon_o\epsilon_{si}qN_b}f(\phi_s) \tag{3.22}$$

Where $f(\phi_s)$ is given by:

$$f(\phi_s) = [\phi_s + \exp\left(-2\phi_f/V_t\right)(V_t \exp\left(\phi_s/V_t\right) - V_t - \phi_s) + V_t \exp\left(-\phi_s/V_t\right) - V_t]^{1/2} \tag{3.23}$$

3.1.2 Gate Capacitance

The gate capacitance is defined as:

$$C_g = \frac{dQ_g}{dV_g} \tag{3.24}$$

With $dQ_s = -dQ_g$ and a fixed or zero effective interface charge (see eq. (3.8)), the variation of the gate voltage is given by:

$$dV_g = d\phi_s + \frac{dQ_g}{C_{ox}} \tag{3.25}$$

Hence, the gate capacitance can be expressed as [1, p. 148; 2, p. 292]:

$$\frac{1}{C_g} = \frac{1}{C_{ox}} + \frac{1}{C_s} = \frac{1}{C_{ox}} + \frac{1}{dQ_g/d\phi_s} \tag{3.26}$$

Where C_{ox} is given by:

$$C_{ox} = \epsilon_o\epsilon_{si}/t_{ox} \tag{3.27}$$

The variation of the gate charge corresponds to the variation of the semiconductor charge that is given by the depletion charge Q_b and the inversion charge Q_i:

$$C_s = -dQ_s/d\phi_s = -(dQ_b/d\phi_s + dQ_i/d\phi_s) \tag{3.28}$$

Posing $C_b = -dQ_b/d\phi_s$ and $C_i = -dQ_i/d\phi_s$, we get:

$$C_s = C_b + C_i \tag{3.29}$$

The total capacitance is written:

$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{1}{C_b + C_i} \tag{3.30}$$

Accumulation Region In accumulation, the surface charge is approximately given by [1, p. 293]:

$$Q_s \simeq -Sign(\phi_s)\frac{\epsilon_o\epsilon_{si}V_t}{L_d} \exp\left(-\phi_s/2V_t\right) \tag{3.31}$$

For a p-type semiconductor, this charge consists of holes which are attracted to the surface.

Thus, the dynamic capacitance that is associated with the space charge is given by:

$$C_s \simeq C_i = \frac{\epsilon_o\epsilon_{si}}{2L_D} \exp\left(-\phi_s/2V_t\right) \tag{3.32}$$

Depletion and Weak Inversion Region In depletion, the space charge of a p-type semiconductor (n-MOSFET) is formed by ionized acceptors. The charge is then given by the bulk charge:

$$Q_s \simeq Q_b = -\sqrt{-2qN_b\epsilon_o\epsilon_{si}\phi_s} = C_{ox}\gamma\sqrt{\phi_s} \tag{3.33}$$

Where γ is given by:

$$\gamma = \frac{\sqrt{2\epsilon_o\epsilon_{si}N_a}}{C_{ox}} \tag{3.34}$$

Thus, the depletion capacitance for a p-type semiconductor is given by:

$$C_s \simeq C_b = \sqrt{-qN_a\epsilon_o\epsilon_{si}/2\phi_s} \tag{3.35}$$

Strong Inversion In strong inversion, electrons are attracted to the surface and the space charge Q_s can be approximated by [2, p. 294]:

$$\begin{aligned} Q_s &\simeq -\frac{\epsilon_o\epsilon_{si}V_t}{L_D}(\phi_s/V_t + \exp(\phi_s - 2\phi_f/V_t))^{1/2} \\ &= -\sqrt{2\epsilon_o\epsilon_{si}qN_a}[\phi_s + V_t\exp((\phi_s - 2\phi_f)/V_t)]^{1/2} \end{aligned} \tag{3.36}$$

The total charge can be divided into depletion charge Q_b and inversion charge Q_i. As the depletion charge is given by eq. (3.33), the inversion charge becomes:

$$Q_i = -\sqrt{2\epsilon_o\epsilon_{si}qN_a}[\sqrt{\phi_s + V_t\exp(\phi_s - 2\phi_f/V_t)} - \sqrt{\phi_s}] \tag{3.37}$$

The two terms do not respond to the variation of the gate voltage in the same way. The depletion charge results from the evacuation of the majority carriers and follows instantaneously the variation of the gate voltage, whereas the inversion charge results from the thermal creation of minority carriers and is therefore much slower. Thus, the capacitance depends on the frequency.

Low Frequency Capacitance In inversion, the inversion charge is much higher than the depletion charge and able to follow the variation of the bias for low frequencies. As the inversion charge is given by:

$$Q_s \simeq Q_i = -\frac{\epsilon_o\epsilon_{si}kT}{qL_D}\exp((\phi_s - 2\phi_f)/2V_t) \tag{3.38}$$

we get the following expression for the capacitance:

$$C_s \simeq C_i = \frac{\epsilon_o\epsilon_{si}}{2L_D}\exp((\phi_s - 2\phi_f)/2V_t) \tag{3.39}$$

High Frequency Capacitance At high frequency, the mobile charge is not able to follow the variation of the gate voltage. The modulation of the space charge results exclusively from the depletion charge:

$$C_s = C_b = \epsilon_o\epsilon_{si}/x_d \tag{3.40}$$

3.2 MOSFET DC Model

The standard DC model [1, p. 230] for the n-MOSFET is derived under the following assumptions [3]:

1. The variation of the electric field in the y-direction (along the channel) is much less than the variation of the electrical field in the x-direction (see fig. 3.5 and fig. 3.7, gradual channel approximation, GCA). Using the GCA, the Poisson equation becomes one-dimensional:

$$\frac{d^2\phi}{dx^2} = -\frac{\rho(x)}{\epsilon_o \epsilon_{si}} \tag{3.41}$$

   As it is verified, the GCA is valid for the whole channel except close to the drain and source terminals. However, charge and potential depend on the y-direction.

2. In a n-channel device, the hole current can be neglected for the normal regions of operation. In fact, the hole current gets important only in the avalanche region.

3. Recombination and generation processes are neglected. The drain current is constant at any point along the channel.

4. Current flows in the y-direction only ($\partial\phi/\partial x = 0$).

5. The surface potential is approximated by a value which is independent of the gate voltage.

6. The bulk charge is approximated by the value calculated for a constant value of the surface potential.

7. The electrons are only due to the inversion of the surface.

8. The mobility is taken to be constant ($\mu(x,y) = \mu_{eff}$; effective mobility).

9. The diffusion current is taken into account in the subthreshold region only.

3.2.1 Linear Region of MOSFET Operation

The drift current of a n-channel device which flows in the y-direction is given by:

$$J_n(x,y) = -qn(x,y)\mu_{eff}\frac{\partial\phi_s}{\partial y} \tag{3.42}$$

Integrating equation (3.42) across the channel width W (z-direction) and depth (x-direction) gives the drain current I_{ds}

$$I_{ds}(y) = -W\int_0^\infty [qn(x,y)\mu_{eff}\frac{\partial\phi_s}{\partial y}]dx = const \tag{3.43}$$

[3] The p-MOSFET model is analogous. However, the electron mobility has to be replaced by the hole mobility. Different results for n- and p-MOSFETs are noted.

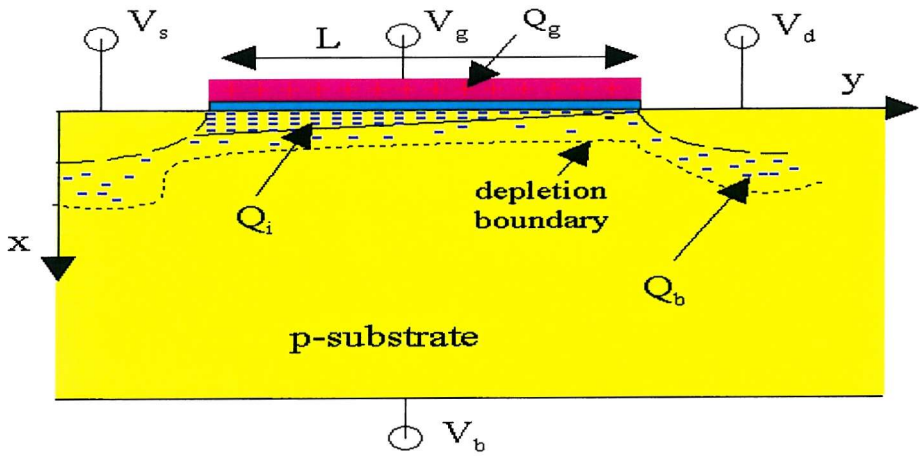


Figure 3.7: n-MOSFET

Using $V_{cb}(y) = \phi_s(y) - \phi_s(0)$, we can write equation (3.43) as:

$$I_{ds} = -W\frac{dV_{cb}}{dy}\int_0^\infty qn(x,y)\mu_{eff}dx \tag{3.44}$$

Equation (3.44) can be written as:

$$I_{ds}(y)dy = -\mu_{eff}WQ_i(y)dV_{cb} \tag{3.45}$$

The variation of the surface potential as a function of the gate voltage and its contribution to the gate capacitance are neglected. The surface potential is approximated by (see fig. 3.4 and fig. 3.6, [1, p. 243]:

$$\phi_s(y) = 2\phi_f + V_b + V(y) \tag{3.46}$$

Where V_b is the voltage between source and bulk that is supposed to be zero.

After integrating equation (3.45), we get [1, p. 243]:

$$I_{ds} = -\mu_{eff}\frac{W}{L}\int_0^{V_{ds}} Q_i(y)dV \tag{3.47}$$

The induced charge is related to the gate voltage by equation (3.8):

$$V_g = V_{fb} + \phi_s(y) - \frac{Q_s}{C_{ox}} \tag{3.48}$$

Substituting $\phi_s(y)$ from equation (3.46) in (3.48) leads to the following expression for the inversion charge density Q_i:

$$Q_i(y) = -C_{ox}[V_g - V_{fb} - 2\phi_f - V(y)] - Q_b(y) \tag{3.49}$$

The bulk charge density is assumed to be constant along the channel length and independent of the applied drain voltage V_{ds} so that the surface potential $\phi_s(y) = 2\phi_f + V_b$ is constant along the length of the channel. Then, equation (3.33) becomes:

$$Q_b(y) = -C_{ox}\gamma\sqrt{2\phi_f + V_b} \tag{3.50}$$

Combining equations (3.49) and (3.50) results in:

$$Q_i(y) = -C_{ox}[V_g - V_{th} - V(y)] \tag{3.51}$$

Where V_{th} is the threshold voltage for the n-MOSFET that is defined as:

$$V_{th} = V_{fb} + 2\phi_f + \gamma\sqrt{2\phi_f + V_b} \tag{3.52}$$

The threshold voltage for the p-MOSFET is given by [3]:

$$V_{th} = V_{fb} - 2\phi_f - \gamma\sqrt{2\phi_f - V_b} \tag{3.53}$$

Inserting $Q_i(y)$ from equation (3.51) into equation (3.47) and integrating, we get the following equation for the n-MOSFET drain current I_{ds} [1, p. 244]:

$$I_{ds} = \frac{\mu_{eff}C_{ox}W}{L}[V_g - V_{th} - 0.5V_{ds}]V_{ds} = \beta[V_g - V_{th} - 0.5V_{ds}]V_{ds} \tag{3.54}$$

Analogously, we get for the p-MOSFET:

$$I_{ds} = -\frac{\mu_{eff}C_{ox}W}{L}[V_g - V_{th} - 0.5V_{ds}]V_{ds} = -\beta[V_g - V_{th} - 0.5V_{ds}]V_{ds} \tag{3.55}$$

β is called the gain factor. For the p-MOSFET, we have to use the hole mobility instead of the electron mobility.

When the drain voltage is small, equation (3.54) becomes approximately:

$$I_{ds} \simeq \beta(V_{gs} - V_{th})V_{ds} \tag{3.56}$$

Hence, we get the effective channel resistance:

$$R_{ch} = \frac{V_{ds}}{I_{ds}} \simeq [\beta(V_g - V_{th})]^{-1} \tag{3.57}$$

The transconductance g_m and the conductance g_d are respectively given by:

$$g_m = \frac{\partial I_{ds}}{\partial V_g} = \frac{W\mu_{eff}C_{ox}}{L}V_{ds} \tag{3.58}$$

$$g_d = \partial I_{ds}/\partial V_{ds} = \frac{W\mu_{eff}C_{ox}}{L}(V_g - V_{th}) \tag{3.59}$$

The mobility that is extracted using the transconductance is called the field-effect mobility μ_{FE}:

$$\mu_{FE} = \frac{g_m L}{WC_{ox}V_{ds}} \tag{3.60}$$

The field-effect mobility is usually considerably lower than the physical effective mobility because of the negligence of the mobility dependence on the gate voltage.

We obtain the maximum of I_{ds} as a function of V_{ds} by differentiating equation (3.54) and equating the result to zero:

$$\frac{dI_{ds}}{dV_{ds}} = \frac{\mu_{eff}C_{ox}W}{L}[V_g - V_{th} - V_{ds}] = 0 \tag{3.61}$$

The drain voltage that gives us the maximal value of the drain source is called the saturation voltage V_{dsat}:

$$V_{dsat} = V_g - V_{th} \tag{3.62}$$

For bias that are higher than the saturation voltage, equation (3.54) is no longer valid and we have to use an other model.

3.2.2 Saturation Region

Replacing V_{ds} by V_{dsat} in equation (3.54) leads to the following equation for the drain current I_{dsat} at the pinch-off point:

$$I_{dsat} = \frac{\beta}{2}(V_g - V_{th})^2 \tag{3.63}$$

Eq. (3.63) is based on the assumption that the drain current is independent of V_{ds} beyond the pinch-off. In reality the channel length varies as a function of V_{ds}. We modify expression (3.63) by introducing an effective channel length:

$$I_{dsat} = \frac{\mu_{eff} C_{ox} W}{2(L - l_d)}(V_g - V_{th})^2 \tag{3.64}$$

Where l_d increases with V_d. Using eq. (3.63), we get:

$$I_{ds} = \frac{I_{dsat}}{1 - l_d/L} \tag{3.65}$$

In general l_d is small compared to L and to a first approximation:

$$(1 - \frac{l_d}{L})^{-1} \simeq 1 + \frac{l_d}{L} = 1 + \lambda V_{ds} \tag{3.66}$$

With this approximation, equation (3.65) for a n-MOSFET becomes:

$$I_{ds} = I_{dsat}(1 + \lambda V_{ds}) \tag{3.67}$$

Similarly, we have the following equation for a p-MOSFET:

$$I_{ds} = -I_{dsat}(1 - \lambda V_{ds}) \tag{3.68}$$

(3.54) and (3.68) are the drain current equations of the SPICE MOS level 1 model. The subthreshold current is considered first in higher spice level models.

3.2.3 Subthreshold Region Model

So far, the diffusion current has always been neglected while deriving the drain current. However, the diffusion current dominates the subthreshold behavior. Subthreshold current is a leakage current and affects therefore dynamic circuits and standby power. In the subthreshold region, the charge density is nearly constant along the y-direction and we can replace the surface potential $\phi_s(y)$ by a constant value ϕ_{ss}. The depletion charge Q_b can be expressed as (see eq. (3.33)):

$$Q_b = -C_{ox}\gamma\sqrt{\phi_s(y)} = -C_{ox}\gamma\sqrt{\phi_{ss}} \tag{3.69}$$

Since $Q_i << Q_b$, we have $Q_s \simeq Q_b$ and equation (3.8) becomes:

$$V_g = V_{fb} + \phi_{ss} - \frac{Q_b}{C_{ox}} \tag{3.70}$$

Solving equations (3.69) and (3.70), we get

$$\begin{aligned}\phi_{ss} &= V_g - V_{fb} + \frac{\gamma^2}{2}[1 - \sqrt{1 + \frac{4}{\gamma^2}(V_g - V_{fb})}] \\ &= [-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + V_g - V_{fb}}]^2\end{aligned} \tag{3.71}$$

The diffusion current is given by:

$$J_n = qD_n \frac{dn}{dy} \tag{3.72}$$

Using the Einstein equation $D_n = \mu_{eff} V_t$, integrating eq. (3.72) across the channel of thickness x_d gives us I_{ds} in the subthreshold region as:

$$I_{ds} = \mu_{eff} W V_t \frac{dQ_i}{dy} \tag{3.73}$$

We further integrate from y=0 to y=L:

$$I_{ds} = \frac{W}{L} \mu_{eff} V_t \int_{Q_{is}}^{Q_{id}} dQ_i = \frac{W}{L} \mu_{eff} V_t (Q_{id} - Q_{is}) \tag{3.74}$$

Where Q_{is} and Q_{id} are the inversion charge densities at the source and the drain end respectively when the device is in the subthreshold region.

In weak inversion, the expression for the inversion charge density $Q_i(y)$ (3.37) can be simplified [1, p. 143]:

$$Q_i(y) = \frac{\gamma C_{ox}}{2\sqrt{\phi_{ss}}} V_t \exp\left((\phi_{ss} - 2\phi_f - V_{cb}(y))/V_t\right) \tag{3.75}$$

Thus, we get:

$$I_{ds} = \frac{\mu_{eff} W C_{ox} \gamma}{2L\sqrt{\phi_{ss}}} V_t^2 \exp\left((\phi_{ss} - 2\phi_f - V_{sb})/V_t\right)(1 - \exp\left(-V_{ds}/V_t\right)) \tag{3.76}$$

The above equation is often rewritten in terms of the surface concentration n_s:

$$I_{ds} = \frac{W D_n x_d q n_s}{L} (1 - \exp\left(-V_{ds}/V_t\right)) \tag{3.77}$$

In general, a simplified form of equation (3.77) is used. The bulk charge Q_b is a weak function of ϕ_s and can be expanded into a Taylor series around ϕ_{ss}:

$$Q_b(\phi_s) \simeq Q_b(\phi_{ss}) + (\phi_s - \phi_{ss}) \frac{\partial Q_b}{\partial \phi_s} \tag{3.78}$$

Using equation (3.69), it follows:

$$\frac{\partial Q_b}{\partial \phi_s} = C_d = \frac{\gamma C_{ox}}{2\sqrt{\phi_s}} \tag{3.79}$$

C_d is called the depletion region capacitance. Combining equations (3.78), (3.79), and (3.70) yields:

$$V_g = V_{fb} + \phi_{ss} + \frac{Q_b(\phi_{ss})}{C_{ox}} + (\phi_s - \phi_{ss}) \frac{C_d}{C_{ox}} \tag{3.80}$$

An appropriate value for ϕ_{ss} would lie in the middle of the subthreshold region. However, using $\phi_{ss} = 2\phi_f + V_b$ as in many circuit models, equation (3.80) becomes:

$$\phi_s - 2\phi_f - V_b = \frac{V_g - V_{th}}{\eta} \tag{3.81}$$

Where:

$$\eta = 1 + \frac{C_d}{C_{ox}} = 1 + \frac{\gamma}{2\sqrt{2\phi_f + V_{sb}}} \tag{3.82}$$

Combining equations (3.76), (3.79), and (3.81) leads to:

$$I_{ds} = \beta \frac{C_d}{C_{ox}} V_t^2 \exp\left[\frac{V_g - V_{th}}{\eta V_t}\right](1 - \exp\left(-V_{ds}/V_t\right)) \tag{3.83}$$

Neglecting the dependence on the drain voltage V_{ds}, a simplified version of equation (3.83) can be obtained [1, p. 262]:

$$I_{ds} = I_{on} \exp\left(\frac{V_g - V_{th}}{\eta V_t}\right) \tag{3.84}$$

3.2.4 Device Speed

The transconductance $g_m(\omega)$ depends on the frequency ω [4]:

$$g_m(\omega) = \frac{g_m}{1 + \frac{j\omega}{\omega_c}} \tag{3.85}$$

Where g_m is the DC value of the transconductance defined in equation (3.58).

In saturation, the cutoff frequency ω_c is given by [4]:

$$\omega_c = \frac{\mu_{eff}}{2L^2} \cdot (V_g - V_{th}) \tag{3.86}$$

The device speed is therefore proportional to μ_{eff} and $1/L^2$. In practice, the extrinsic capacitances lead to a much lower switching speed than predicted.

3.3 Modern MOSFETs

The modern MOSFET is a short geometry MOSFET with a thin gate oxide. Then, the approximations above for the inversion and bulk capacitance is not valid.

The depletion layer widths of source and drain junctions are comparable to the channel length and the potential becomes two-dimensional. An effective channel length L_{eff} is introduced which can differ considerably from the drawn channel length L.

The off-state current is in general high and hot carrier effects occur because the scaling of the device dimensions is not followed by a proportional scaling of the supply voltage. The source and drain series resistance can not be neglected.

3.3.1 Change in Device Behavior

When the channel length is reduced, the depletion layer widths of source and drain junctions are comparable to the channel length and the device characteristics change:

1. The punch-through is characterized by a super-linear increase of the drain current as a function of the drain voltage. The punch-through can take place in all regions of device operation. The punch-through effect occurs when the applied drain voltage surpasses the punch-through voltage V_{pt} and the depletion width of the source touches the depletion

width of the drain. An important current can flow even when the gate voltage is lower than the threshold voltage.

For an n-MOSFET, the punch-through voltage V_{pt} is given by [1, p. 43, p. 74, p. 81]:

$$V_{pt} \simeq \frac{eN_a}{2\epsilon_o\epsilon_{si}}(L - \sqrt{\frac{2\epsilon_o\epsilon_{si}\phi_{bi}}{qN_a}})^2 - \phi_{bi} \tag{3.87}$$

N_a refers to the substrate, whereas N_d refers to the source and drain. ϕ_{bi} is the built-in potential of a pn-junction:

$$\phi_{bi} = V_t \ln(\frac{N_a N_d}{n_i^2}) \tag{3.88}$$

For low substrate doping and a long channel, the punch-through voltage is very high and, in general, the avalanche breakdown will take place first.

2. When the channel length is reduced and the drain voltage is increased, the source and drain depletion regions move more closely together. Their electrical fields penetrate the channel region. The threshold voltage for a short channel MOSFET V_{thd} is dependent on the drain voltage due to DIBL [1, p. 212]:

$$V_{thd} = V_{th}(V_{ds}) = V_{th} - \sigma V_{ds} = V_{th} - \delta V_{th} \tag{3.89}$$

σ is called the DIBL parameter.

The subthreshold slope increases as the channel length is shrunken. The subthreshold behavior is critical in CMOS in order to reduce off-state current and power consumption. A similar effect leads to the decrease of the threshold voltage in very narrow devices.

3. Source and drain resistance are no longer negligible in comparison to the channel resistance. Now, they act as additional series resistance R_{sd} and change the device characteristics.

4. The variation of the gate capacitance as a function of the gate voltage cannot be neglected for devices with a thin gate oxide. The drain current for small drain voltages is given by [5]:

$$I_{sd} = \frac{Q_i \cdot \mu_{eff} \cdot W \cdot (V_{sd} - R_{sd} \cdot I_{sd})}{L} \tag{3.90}$$

5. Carriers experience in short channel devices a high electrical field. The velocity of electrons and holes saturates respectively at $1 \cdot 10^6 V/m$ and $2 \cdot 10^6 V/m$ leading to mobility degradation.

3.3.2 Hot Carrier Effects

Hot carrier effects [6-10] are indirectly linked to short channel effects as the continuous scaling of the channel length was not accompanied by the proportional scaling of the supply voltage. Therefore, the maximum electrical field experienced by the carriers in the channel region has steadily increased and hot carrier effects may occur when the carriers move from source to drain and acquire a kinetic energy that is higher than the thermal energy of the lattice (hot carriers).

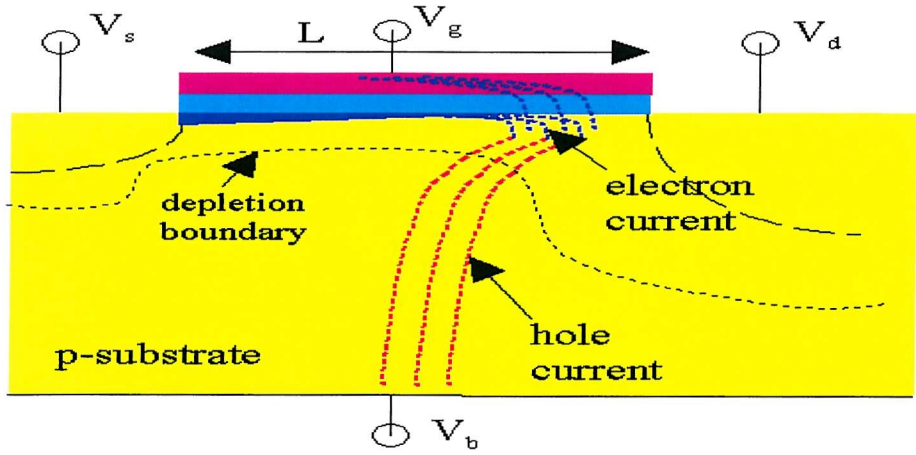


Figure 3.8: Electron and hole current due to impact ionization

Oxide charging Hot carriers can surmount the $Si - SiO_2$ band offset (3.1 eV for electrons and 4.8 eV for holes) and enter into the oxide (see fig. 3.8). Hole injection is much less efficient than electron injection due to the higher energy barrier. The carriers can either be trapped in the gate oxide or result is a gate current (n-MOSFETs: channel hot electrons, CHE). Oxide charging seriously affects long-term operating devices as the charge continues to increase during operation. It deteriorates the transconductance and limits the maximum voltage that can be applied. The gate current results in the degradation of the oxide quality and makes the device eventually unreliable.

Avalanche multiplication Besides, hot carriers can create electron-hole pairs by impact ionization. Both can be injected into the oxide (secondarily generated hot electrons, SGHE). The majority carriers that are not injected into the gate oxide increase the substrate current (see fig. 3.8), whereas the minority carriers contribute to the drain current. The carriers created by impact ionization can create further electron-hole pairs if they acquire a sufficient energy (drain avalanche hot carriers, DAHC). The avalanche multiplication is characterized by a strong increase of the drain current as a function of the drain voltage. Due to the high resistivity of the substrate, the bulk voltage close to the devices can easily be changed by the substrate current leading to the malfunction of circuits.

Parasitic bipolar transistor action A CMOS process leads to the creation of two parasitic bipolar transistors (see fig. 3.9). In a single MOSFET, source, substrate and drain form an inherent parasitic bipolar transistor (PBT) in parallel with the MOSFET. In a n-well process, p^+source, n-well and p-substrate form a vertical pnp transistor and n-well, p-substrate and n^+ source a lateral npn transistor. The base of each PBT is driven by the collector of the other (feed back loop). These two parasitic transistors constitute a silicon controlled rectifier (SCR). A substrate current due to hot carriers can forward bias emitter-base junctions resulting in large

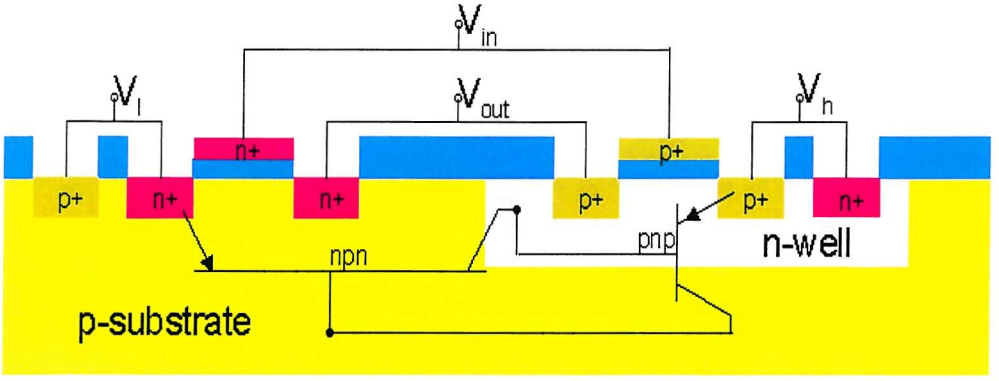


Figure 3.9: n-MOSFET

conduction (latchup) which can easily destroy a chip.

3.4 SiGe for MOSFET applications

SiGe heterojunctions can be used in the channel, the gate, the source and the drain of MOSFETs. The interest in SiGe for MOSFET applications arises from the band offset and the enhanced transport properties as discussed in chap. 2. Additionally, a better dopant activation can be achieved in SiGe than in Si.

As the work function depends on the Ge content [11], the use of poly-SiGe as a gate material allows the adjustment of the threshold voltage. Besides, the better dopant activation in poly-SiGe reduces the gate resistance [12-14].

In source and drain [15-19], SiGe can reduce the offstate current, the parasitic bipolar gain and as a consequence punch-through and latch-up. The better dopant activation in SiGe can lower the resistivity of source and drain and the contact resistance [20].

This project focuses on the use of SiGe in the channel. The interest in SiGe heterojunctions in the channel originates from the enhanced transport properties leading to an increased cut-off frequency (see eq. (3.86)). Besides, HMOSFETs display a better noise performance than conventional MOSFETs [21, 22].

Although HMOSFETs display qualitatively the same behavior as conventional MOSFETs, the quantitative drain current equations derived in this chapter are not valid for HMOSFETs as it will be demonstrated in the following chapter. In contrast to conventional MOSFETs, the space charge region of HMOSFETs has to be divided into three regions - the cap layer, the heterochannel, and the buffer layer. A general analytical expression for the space charge region of the HMOSFET is not known. Because of the band offset, the magnitude of the threshold voltage is lowered for HMOSFETs [23].

3.5 Conclusion

In this chapter, the electrical characteristics of MOSFETs have been reviewed. Strained SiGe heterojunctions can enhance the speed performance, reduce noise and short-channel effects of MOSFETs.

The drain current equations derived for devices with a long channel are not strictly valid for modern MOSFETs with a short channel and a thin gate oxide. This is as well the case for HMOSFETs which display qualitatively the same electrical characteristics as conventional MOSFETs. The analytical formulas provide us with insight about the dependence of the device performance on single parameters. Besides, HMOSFETs are always compared to conventional MOSFETs. In the following chapter, it will be demonstrated that mobility extraction methods developed for MOSFETs with a thin gate oxide can be employed for HMOSFETs.

Bibliography

[1] N. Arora, "MOSFET Models for VLSI Circuit Simulation/Theory and Practice", Springer-Verlag, 1992 Instr. and Meth., Vol. B 96, pp. 281-285, 1995

[2] H. Mathieu, "Physique des semiconducteur et des composants électroniques", Masson, 1996

[3] R. L. Geiger, "VLSI Design Techniques for Analog and Digital Circuits", McGraw Hill, 1990

[4] C. T. Sah, "Characteristics of the Metal-Oxide-Semiconductor Transistors", IEEE Transactions on Electron Devices, ED 11, pp. 324-345, 1964

[5] G. Niu, J. D. Cressler, S. J. Mathew, and S. Subbanna, "A Total Resistance Slope-Based Effective Channel Mobility Extraction Method for Deep Submicrometer CMOS Technology", IEEE Trans. Electron Dev., Vol. 46, No. 9, pp 1912-1914, 1999

[6] S. M. Sze, "Physics of Semiconductor Devices", John Wiley & Sons, 1981

[7] B. Janossy, Y. Haddab, J.-M. Villiot, and R. S. Popovic, "Hot carrier Hall devices in CMOS technology", Sensors and Actuators, Vol. A 71, pp. 172-178, 1998

[8] N. Afshar-Hanaii, "Some Aspects of Submicron CMOS Technology", Ph.D. thesis, University of Southampton, 1993

[9] A. Waite, "Some Aspects of Submicron CMOS Devices", Minithesis, University of Southampton, June 1995

[10] C. T. Wang, "Hot Carrier Design Considerations for MOS Devices and Circuits", Van Nostrand Reinhold, 1992

[11] T.-J. King, J. R. Pfiester, and K. C. Saraswat, "A Variable-Work-Function Polycrystalline $Si_{1-x}Ge_x$ Gate Material for Submicrometer CMOS Technologies", IEEE Electron Device Letters, Vol. 12, No. 10, pp. 533-535, 1991

[12] W.-C. Lee, T.-J. King, and C. Hu, "Observation of Reduced Boron Penetration and Gate Depletion for $Poly - Si_{0.8}Ge_{0.2}$ Gated PMOS Devices", IEEE Electron Dev. Lett., Vol. 20, No. 1, pp. 9-11, 1999

[13] T. Noguchi, J. A. Tsai, A. J. Tang, and R. Reif, "Resistivity Study of $P-$, $B-$, and BF_2-Implanted Poycrystalline $Si_{1-x}Ge_x$ Films with Subsequent Annealing", Jpn. J. Appl. Phys., Vol. 33, Part 2, No. 12 B, pp. L 1748-L 1750, 1994

[14] V. E. Houtsma, J. Holleman, C. Salm, F. P. Widdershoven, and P. H. Woerlee, "Stress-Induced Leakage Current in P^+ Poly MOS Capacitors with Poly-Si and Poly-$Si_{0.7}Ge_{0.3}$ Gate Material", IEEE Electron Dev. Lett., Vol. 20, No. 7, pp. 314-316, 1999

[15] R. M. Sidek, A. G. R. Evans, and R. A. Kubiak, "Reduction of parasitic bipolar transistor action and punchthrough susceptibility in MOSFETs using $Si/Si_{1-x}Ge_x$ sources and drains", Electron Lett., Vol. 32, No. 3, pp. 269-270, 1996

[16] J.-H. Sim, C.-H. Choi, and K. Kim, "Elimination of Parasitic Bipolar-Induced Breakdown Effects in Ultra-Thin SOI MOSFET's using Narrow-Bandgap-Source (NBS) Structure", IEEE Trans. Electron Dev., Vol. 42, No. 8, pp. 1495-1502, 1995

[17] X. Chen, Q. Ouyang, D. M. Onsongo, S. K. Jayanarayanan, A. Tasch, and S. Banerjee, "SiGe heterojunction vertical p-type metal-oxide-semiconductor field-effect transistors with Si cap", Appl. Phys. Lett., Vol. 77, No 11, pp. 1656-1658, 2000

[18] N. Collaert and K. DeMeyer, "Modeling the Short-Channel Threshold Voltage of a Novel Vertical Heterojunction pMOSFET", IEEE Trans. Electron Dev., Vol. 46, No. 5, pp. 933-939, 1999

[19] N. Collaert, P. Verheyen, and K. De Meyer, "On the Reverse Short Channel Effect in Deep Submicron Heterojunction MOSFET's and Its Impact on the Current-Voltage Behavior", IEEE Trans. Electron Dev., Vol. 47, No. 6, pp. 1214-1220, 2000

[20] I. Mizushima, A. Murakoshi, K. Suguro, N. Aoki, and J. Yamauchi, "Ultra high dose boron ion implantation: super-saturation of boron and its application", Mat. Chem. Phys., Vol. 54-59, 1998

[21] A. F. M. Anwar, K.-W. Liu, and V. P. Kesan, "Noise Performance of $Si/Si_{1-x}Ge_x$ FET's", IEEE Trans. Electron. Dev., Vol. 42, No. 10, pp. 1841-1846, 1995

[22] A. D. Lambert, B. Alderman, R. J. P. Lander, E. H. C. Parker, and T. E. Whall, "Low Frequency Noise Measurements of P-Channel $Si_{1-x}Ge_x$ MOSFET's", IEEE Trans. Electron Dev. Vol. 46, No. 7, pp. 1484-1486, 1999

[23] A. Sadek, S. E.-D. Habib, and K. Ismail, "Analytical Model of Si/ SiGe Field-Effect Transistors", Solid-State Electronics, Vol. 38, No. 11, pp. 1969-1971, 1995

Chapter 4

Electrical Characterization of MOSFETs

In this chapter, a few available characterization routines for MOSFETs out of the vast number of existing routines are presented. Charge pumping measurements allow the characterization of the oxide interface which contributes to carrier scattering, s-parameter measurements the assessment of MOSFETs at high frequencies and possibly even mobility extraction. Simple and robust extraction routines of MOSFET parameters are based on the approximation of the gate capacitance by a constant. But, the constant capacitance approximation may not be justified for HMOSFETs. Schroedinger-Poisson simulations are used in order to demonstrate that split-CV measurements are ideal for the mobility extraction on HMOSFETs.

4.1 Split-CV Measurements

Split-CV methods were originally developed for MOSFET with an ultrathin gate oxide [1, 2]. The constant capacitance approximation is not suitable for these devices and leads to a systematic error in the mobility extraction because the varying capacitances associated with the inversion charge and the bulk charge cannot be neglected.

In split-CV methods [3-7], the currents I_1 and I_2 are measured while a voltage V_g is applied to the gate (see fig. 4.1). The split-CV method is based on the assumption that the current I_1 consists of minority carriers which flow into the channel underneath the gate from the source and the drain when the substrate is inverted, whereas the current I_2 is due to the variation of the bulk charge. In quasistatic split-CV [3], the inversion charge Q_i is obtained by:

$$Q_i(V_g) = \int_{-\infty}^{V_g} I_1(V) \frac{dt}{dV} dV \tag{4.1}$$

Using eq. (3.8) in chap. 3, a theoretical expression for of the variation of the inversion charge as

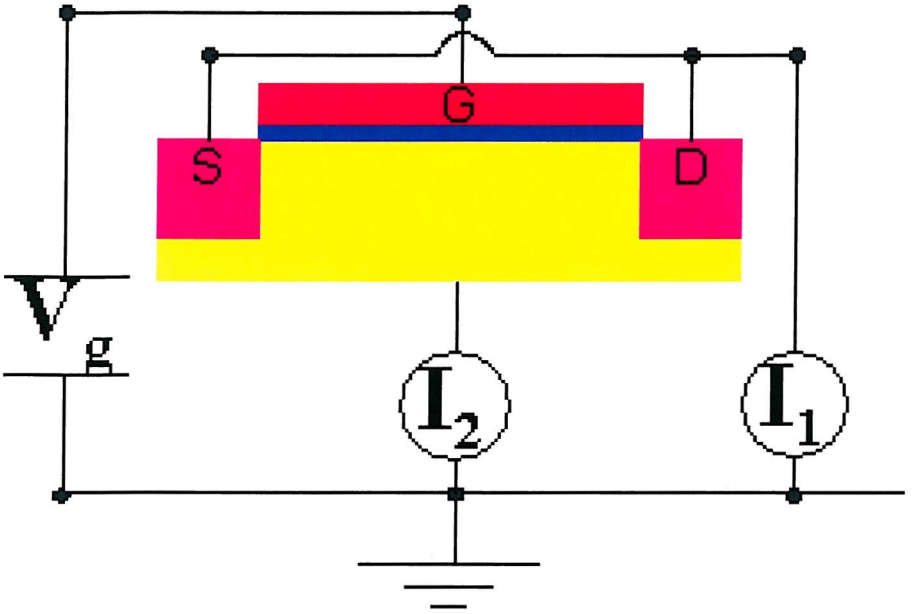


Figure 4.1: Setup for split-CV measurements

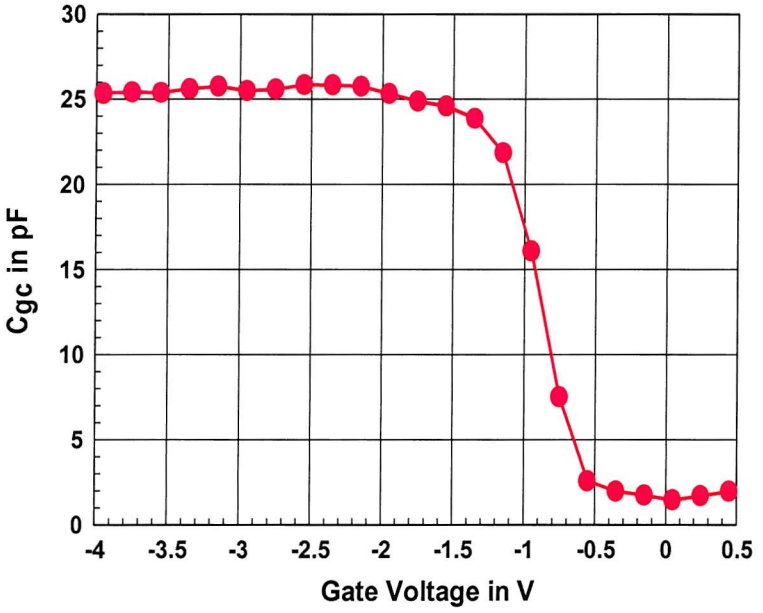


Figure 4.2: C_{gc} extracted for a p-MOSFET

a function of the gate voltage can be determined:

$$V_g = V_{fb} + \phi_s - \frac{Q_s}{C_{ox}} \tag{4.2}$$

Differentiating eq. (4.2) with respect to V_g, the following equation is obtained:

$$1 = \frac{d\phi_s}{dV_g} - \frac{1}{C_{ox}} \frac{dQ_s}{d\phi_s} \frac{d\phi_s}{dV_g} \tag{4.3}$$

Using $Q_s = Q_i + Q_b$, eq. (4.3) can be rearranged:

$$\frac{d\phi_s}{dV_g} = \frac{C_{ox}}{C_{ox} - \frac{dQ_i}{d\phi_s} - \frac{dQ_b}{d\phi_s}} \tag{4.4}$$

Using eq. (4.4), the variation of the inversion charge as a function of the gate voltage, ie the gate-channel capacitance $dQ_i/dV_g = C_{gc}$, can be expressed as [5]:

$$C_{gc} = \frac{dQ_i}{d\phi_s} \frac{d\phi_s}{dV_g} = \frac{\frac{dQ_i}{d\phi_s} C_{ox}}{C_{ox} - \frac{dQ_i}{d\phi_s} - \frac{dQ_b}{d\phi_s}} = \frac{C_{ox} C_l}{C_{ox} + C_l + C_D} \tag{4.5}$$

Where $C_l = -dQ_i/d\phi_s$ and $C_D = -dQ_b/d\phi_s$.
Thus, the inversion charge for a p-MOSFET is given by:

$$Q_i = \int_{V_g}^{\infty} C_{gc}(V) dV \tag{4.6}$$

In practice, the integration should start at the flatband voltage V_{fb}.

A measured C_{gc}-curve is shown in fig. 4.2 for a p-MOSFET with a 6 nm thick gate oxide ($W = 1000\ \mu m$ and $L = 6\ \mu m$). The fact that C_{gc} does even for accumulation not become zero can be attributed to the variation of the gate length as a function of the gate voltage and the small width of the gate.

The gate-channel capacitance can also be obtained from LF [4] and HF [5] measurements. Quasistatic split-CV measurements require large MOSFETs with a gate to channel capacitance of at least a few tens of pF. Recent split-CV methods [5] are even able to account for the drain voltage and interface traps [6].

4.2 S-Parameter Measurements

S-parameter measurements are used to characterize the performance of devices at high frequencies. They are based on the description of electrical signals as power waves and require the extensive use of calibration routines for the extraction of the intrinsic device parameters.

4.2.1 Measurement of a Two-Port System at High Frequencies

At low frequencies, a two-port system like a transistor is usually described by y-parameters [7-9]:

$$\begin{aligned} I_1 &= y_{11} V_1 + y_{12} V_2 \\ I_2 &= y_{21} V_1 + y_{22} V_2 \end{aligned} \tag{4.7}$$

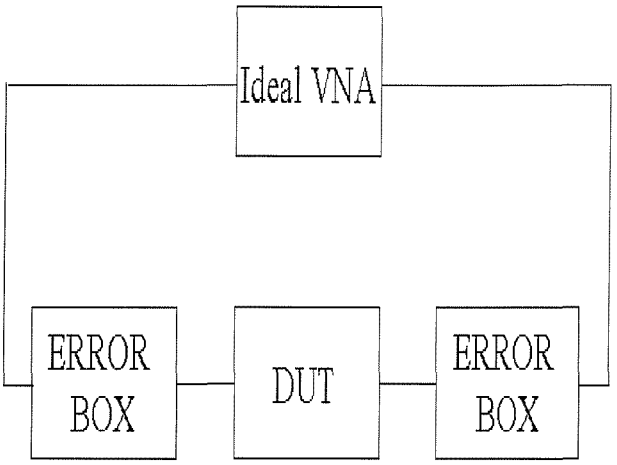


Figure 4.3: Two-error-two-port model

Where I_1 and V_1 denote the current and voltage at port 1, i. e. the gate and I_2 and V_2 denote the current and voltage at port 2, i. e. the drain.

At higher frequencies, voltages and currents are difficult to measure and their definitions may even be questionable. Then, circuits are described by traveling 'power' waves related to each other by s- or scattering parameters in the form of linear simultaneous equations. S-parameters relate the square roots of the reflected power at port 1 b_1 and at port 2 b_2 to the square roots of the incident power at port 1 a_1 and port 2 a_2 [7-9]:

$$\begin{aligned} b_1 &= s_{11}a_1 + s_{12}a_2 \\ b_2 &= s_{21}a_1 + s_{22}a_2 \end{aligned} \tag{4.8}$$

The y-parameters can be represented by s-parameters [8-10]:

$$\begin{aligned} y_{11} &= \frac{(1+s_{22})(1-s_{11})+s_{12}s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} \cdot \frac{1}{Z_0} \\ y_{12} &= \frac{-2s_{12}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} \cdot \frac{1}{Z_0} \\ y_{21} &= \frac{-2s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} \cdot \frac{1}{Z_0} \\ y_{22} &= \frac{(1+s_{11})(1-s_{22})+s_{12}s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} \cdot \frac{1}{Z_0} \end{aligned} \tag{4.9}$$

Where Z_0 is the internal impedance of the signal generator.

The current gain H_{21} which is 0 dB for the unity-gain frequency f_T is given by [9]:

$$H_{21} = \frac{y_{21}}{y_{11}} = \frac{-2s_{21}}{(1-s_{11})(1+s_{22})+s_{12}s_{21}} \tag{4.10}$$

4.2.2 Calibration and Deembedding

In s-parameter measurements, the DUT (device under test) is described by the two-error-two-port model (see fig. 4.3). In order to get rid of the two error boxes which are parasitics of the

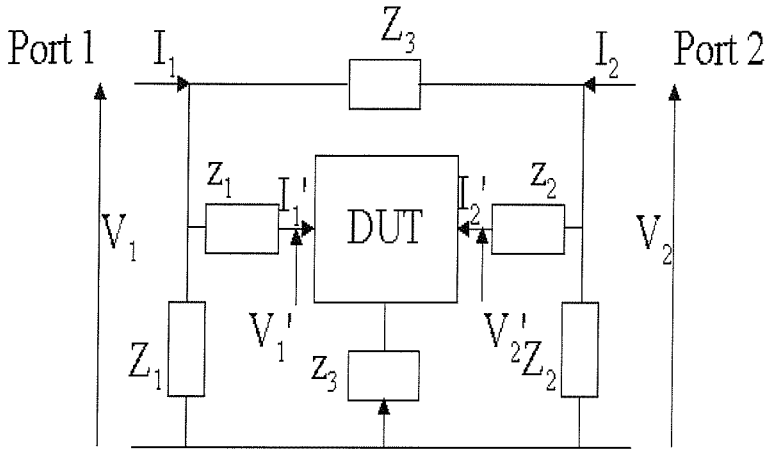


Figure 4.4: DUT with parasitics

measurement system and the vector network analyzer (VNA), several known standards are measured (open, load, short, and through). Calibration is an active research area [10-21]. Common calibration routines are LRM, LRRM, SOLT, SORL, and TRL.

In order to find the intrinsic device performance, the parasitic impedances z_1, z_2, z_3, Z_1, Z_2, and Z_3 introduced by the wiring and contact pads on the chip and shown schematically in fig. 4.4 have to be extracted. The open, shortened, and connected device are measured. The matrices Y_{con}, Y_{short}, and Y_{open} are obtained for the connected, shortened, and open device.

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix}_{con} = Y_{con} \cdot \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{4.11}$$

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix}_{short} = Y_{short} \cdot \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{4.12}$$

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix}_{open} = Y_{open} \cdot \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{4.13}$$

Using the decomposition $\vec{I}_{short} = \vec{I}_{open} + \vec{I}'_{short}$, the current $\vec{I}'_{short}$ can be found and the matrix $Z_{short} = [Y_{short} - Y_{open}]^{-1}$ can be determined:

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = Z_{short} \cdot \begin{pmatrix} I'_1 \\ I'_2 \end{pmatrix} = \begin{pmatrix} z_1 + z_3 & z_3 \\ z_3 & z_2 + z_3 \end{pmatrix} \begin{pmatrix} I'_1 \\ I'_2 \end{pmatrix} \tag{4.14}$$

Using the decomposition $\vec{I}_{con} = \vec{I}_{open} + \vec{I}'_{con}$, the current $\vec{I}'_{con}$ can be found and the matrix $Z_{con} = [Y_{con} - Y_{open}]^{-1}$ can be determined:

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = Z_{con} \cdot \begin{pmatrix} I'_1 \\ I'_2 \end{pmatrix} \tag{4.15}$$

The currents passing through the device have to pass through z_1, z_2, and z_3. The matrix Z_{con} can therefore be decomposed $Z_{con} = Z_{DUT} + Z_{short}$. Then, the following relation for the device

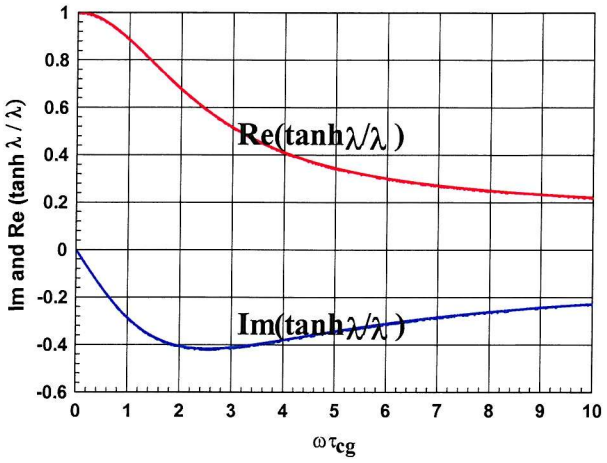


Figure 4.5: Real and imaginary part of the function $\tanh\lambda/\lambda$

without parasitics can be found:

$$\begin{pmatrix} V_1' \\ V_2' \end{pmatrix} = Z_{DUT} \cdot \begin{pmatrix} I_1' \\ I_2' \end{pmatrix} \tag{4.16}$$

The matrix Z_{DUT} can be transformed into a Y-matrix for the intrinsic device $Y_{DUT} = Z_{DUT}^{-1}$:

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = Y_{DUT} \cdot \begin{pmatrix} V_1' \\ V_2' \end{pmatrix} \tag{4.17}$$

4.2.3 Mobility Extraction

The split-CV method in ref. [5] which is not influenced by interface traps and the applied drain voltage might be adapted for VNAs. A realization of this mobility extraction method using VNAs would require six devices - three devices for the determination of I_1 and three devices for the determination of I_2 (short, open connected). The setup is the same as the one used for the other split-CV methods given in fig. 4.1. The gate voltage V_g consists of a base voltage V_{base} and a small superimposed signal v_o at high frequency:

$$V_g = V_{base} + v_o \exp(j\omega t) \tag{4.18}$$

In a transmission line model for high frequencies, the currents I_1 and I_2 in fig. 4.1 generated by v_o are given by [5]:

$$I_1 = -j\omega WL \frac{C_{ox}C_l}{C_{ox}+C_D+C_l} \frac{\tanh\lambda}{\lambda} v_o \tag{4.19}$$

$$I_2 = -j\omega WL[\frac{C_{ox}C_D}{C_{ox}+C_D} - \frac{C_{ox}C_DC_l}{(C_{ox}+C_D)(C_{ox}+C_D+C_l)} \frac{\tanh\lambda}{\lambda}]v_o \tag{4.20}$$

Where $C_l = -dQ_i/d\phi_s$ and $C_D = -dQ_b/d\phi_s$.

The propagation constant λ is given by:

$$\lambda = \sqrt{j\omega\tau_{gc}} = \sqrt{\frac{j\omega L^2 C_{gc}(C_{ox}+C_D)}{4Q_i C_{ox}\mu_{eff}}} \tag{4.21}$$

The function $Im[\tanh\lambda/\lambda]$ has a minimum of -0.4172 for $\omega\tau_{gc} = 2.54$. C_{gc} (see eq. (4.5)) and τ_{gc} can be obtained from the real part of I_1 for a fixed base voltage.

Then, Q_i can be obtained from eq. (4.6):

$$Q_i(V) = \int_{-\infty}^{V_{base}} C_{gc}(V)dV \tag{4.22}$$

The ratio $(C_{ox} + C_D)/C_{ox}$ is given by:

$$\frac{C_{ox} + C_D}{C_{ox}} = \frac{Re[I_1 + I_2]}{Re[I_1]} \tag{4.23}$$

Using, eq. (4.21), the mobility can be expressed as:

$$\mu_{eff} = C_{gc}\frac{Re[I_1 + I_2]}{Re[I_1]}\frac{L^2}{4Q_i\tau_{gc}} \tag{4.24}$$

This mobility extraction method does not take into account the gate resistance and might therefore not be suitable for extremely high frequencies. It is important to take care in the design of the investigated devices that this assumption is justified. A finger-structure can help to reduce the gate resistance.

Besides, the source and drain series resistance is not included in the model. The investigated devices should therefore be relatively long. A large device may also be favorable because the effect of the parasitics is less important. In ref. [5], a 200 μm long and 200 μm wide device is studied. For such a long device, the gate resistance might be reduced considerably by using a metal line on top of the poly-Si gate. However, the low frequency range which would be of interest for devices with a long gate might not be suitable for VNAs.

4.3 Charge Pumping Measurements

Charge pumping measurements [22-32] can provide us with information about the average interface trap density [23], the capture cross section of interface traps [24], the energy and spatial distribution of the interface traps [22]. In comparison with CV measurements, charge pumping measurements are far more sensitive and less prone to oxide leakage. Devices with a gate capacitance in the order of 1 pF can easily be measured.

A pulse train V_g is applied at the gate of a MOSFET and a reverse bias V_{rev} is applied at the source and drain of the MOSFET (see fig. 4.6). The substrate current I_{cp} is measured. The MOSFET goes through three different modes when the transistor is switched from accumulation towards inversion and vice-versa.

The cycle T_p for an n-MOSFET is considered. The form of the applied gate pulse is given in fig. 4.7. The pulse consists of a base voltage V_{base} and a superimposed signal of amplitude ΔV_A. The MOSFET is initially in accumulation.

During t_r, the gate voltage is increased. The first holes which are emitted from traps towards the valence band flow back to the substrate. When the gate voltage approaches the threshold

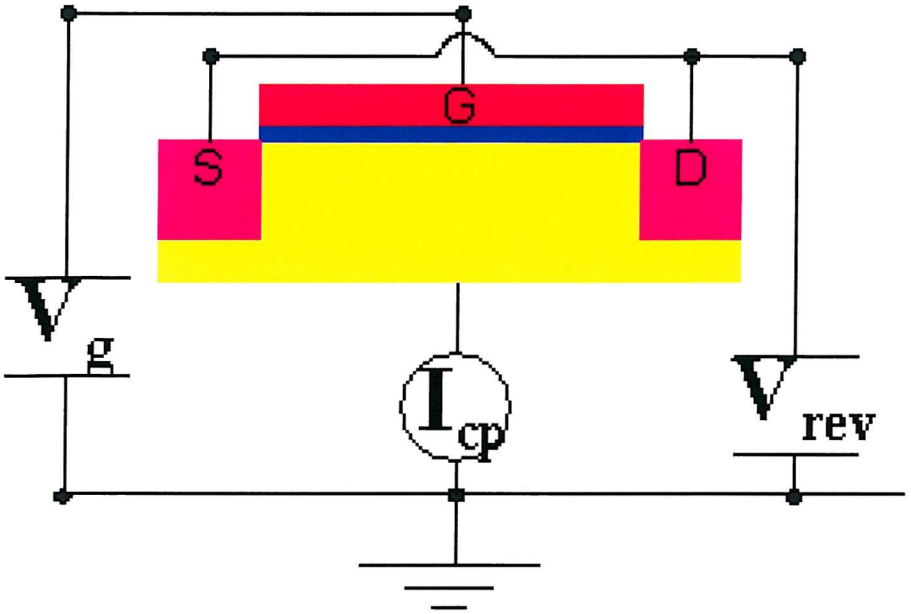


Figure 4.6: Setup for charge pumping measurements

voltage, the remaining holes recombine with electrons coming from source and drain. Further electrons from source and drain fill electron traps at the interface.

The gate voltage reaches the maximum value ($V_g = \Delta V_A + V_{base}$). As the capture processes are nearly instantaneous, all the electron traps with an energy lower than the Fermi level are filled during t_h.

The gate voltage decreases again during t_f. Electrons captured in interface states above the Fermi level are emitted and flow back to source and drain. However, the emission process is not instantaneous, but exponentially dependent on time and temperature. When the gate voltage surpasses the flatband voltage, holes from the substrate recombine with the electrons left in the channel and fill the hole traps.

During t_l, all hole traps above the Fermi energy are filled with holes coming from the substrate. The gate voltage reaches the minimum and the cycle starts from the beginning.

The resulting DC substrate current is due to holes which recombine with electrons. It is a direct measure for the number of hole and electron traps swept through by the Fermi level. However, the electrons and holes which are emitted from traps while the channel is in inversion or accumulation, respectively, do not contribute to I_{cp}. As the emission process depends on temperature, the charge pumping current decreases with increasing temperature.

In case that the emission is neglected for small t_r and t_f, the charge pumping current is given by:

$$I_{cp} = q \cdot f \cdot A_G \cdot \overline{D_{it}} \tag{4.25}$$

Where f is the frequency of the gate pulse, A_G the gate area, and $\overline{D_{it}}$ the mean surface state density averaged over the energy levels swept through by the Fermi level.

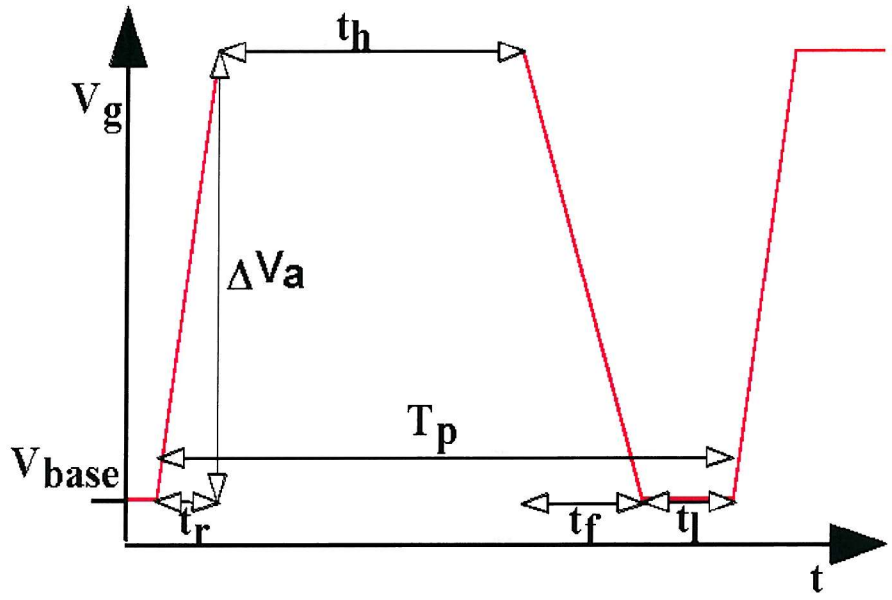


Figure 4.7: Wave form of the applied pulse

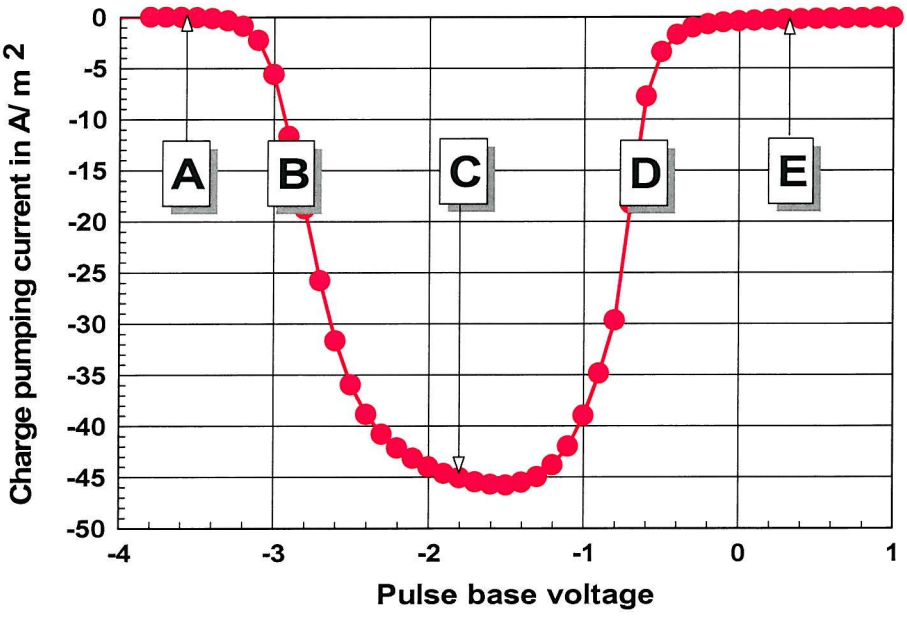


Figure 4.8: Elliot charge pumping curve

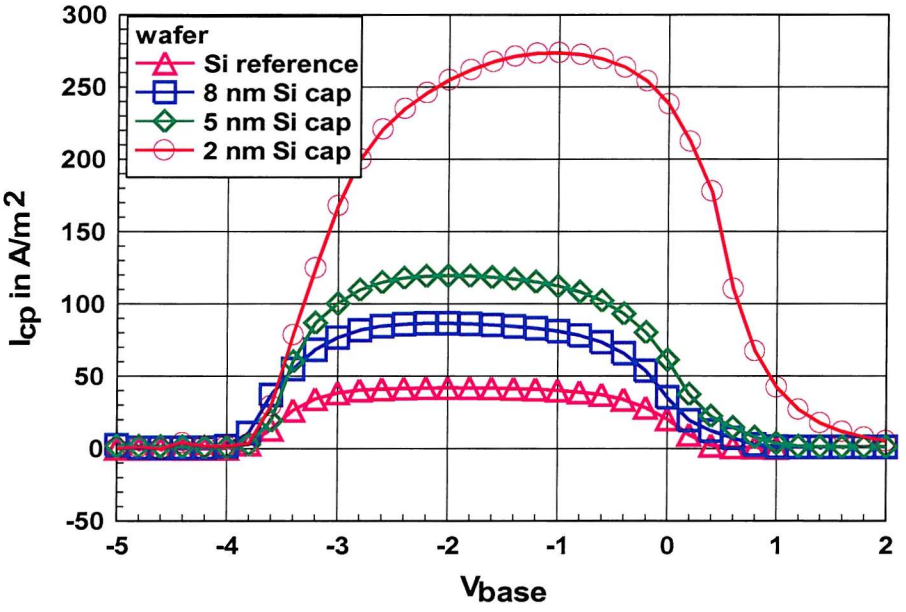


Figure 4.9: I_{cp} for HMOSFETs with different Si cap layers on top of the 36 % Ge channel, $t_{ox} = 6\ nm$, for $\Delta V_A = 4\ V$, and $f = 100\ kHz$

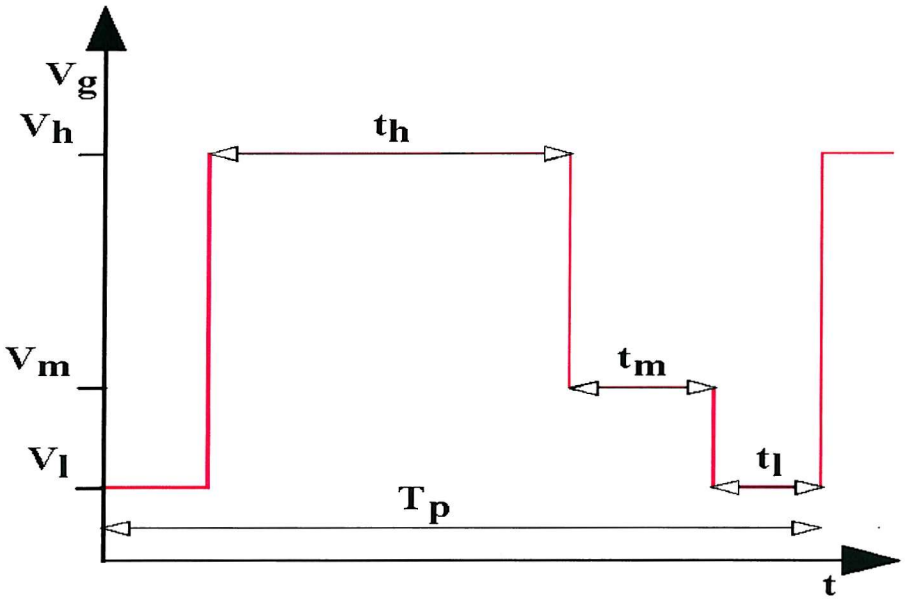


Figure 4.10: Wave form for the three level charge pumping method

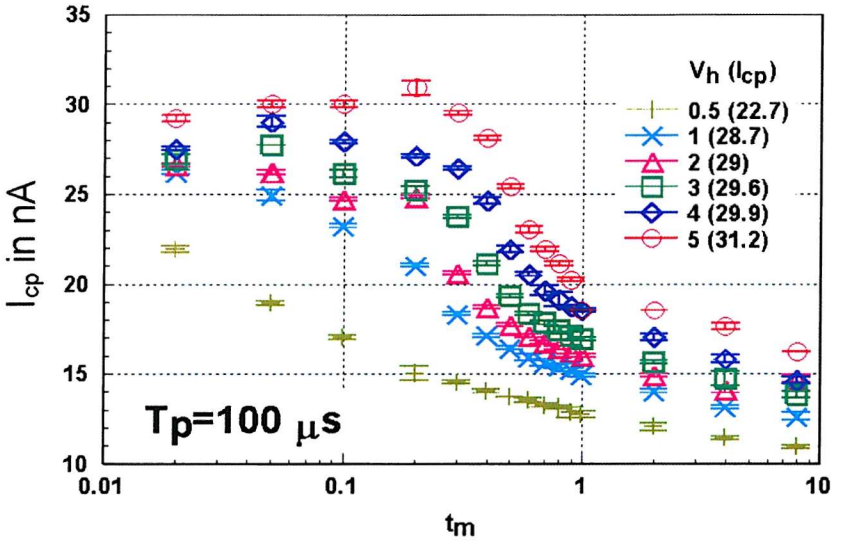


Figure 4.11: Charge pumping current as a function t_m for an n-MOSFET with a 12 nm thick gate oxide, $L = 6\ \mu m$, and $W = 1000\ \mu m$ for $V_l = -2\ V$, $V_m = 0\ V$, and various V_h

Using a simple emission model, the resulting substrate or charge pumping current I_{cp} for a square pulse is given by [22, 23]:

$$I_{cp} = 2 \cdot q \cdot f \cdot \overline{D_{it}} \cdot A_G \cdot k \cdot T[\ln(v_{th} \cdot n_i \sqrt{\sigma_n \sigma_p}) + \ln(\frac{|V_{fb} - V_{th}|}{|\Delta V_A|}\sqrt{t_f \cdot t_r})] \tag{4.26}$$

Where k is the Boltzmann constant, T the absolute temperature, v_{th} the thermal velocity of the carriers, n_i the intrinsic concentration, σ_n the capture cross section of electrons, σ_p the capture cross section of holes, t_f the pulse fall time, and t_r the pulse rise time.

In the Elliot charge pumping technique [25], a constant reverse bias is applied at the source and drain, while the base level of the gate pulse train is swept from a low accumulation to a high inversion level. An increase of the reverse bias voltage reduces the part of the channel that can be measured. Usually, no reverse bias is applied (ie $V_{rev} = 0\ V$). The amplitude of the gate pulse is kept constant, the fall and rise times can be neglected.

The Elliot curve comprises five distinct regions (see fig. 4.8 for an n-MOSFET). In strong inversion (A), the surface potential is pinned close to $2\phi_f$. There is little change in the occupancy of the surface states and the charge pumping current is negligible. When the base voltage is varied between accumulation and strong inversion, many surface states are swept through by the Fermi level resulting in a high charge pumping current (C). When the gate voltage is permanently below the threshold voltage with or without the applied pulse, the surface states can no longer be filled by carriers from the source and the drain and I_{cp} is small (E). (B) and (D) are transition regions. If ΔV_A is larger than the interval between the regions (A) and (E), I_{cp} saturates in the region (C). The maximum of $|I_{cp}|$ can be used in order to calculate $\overline{D_{it}}$ according to eq. (4.25).

Fig. 4.9 compares the Elliot charge pumping curves of HMOSFETs with different Si cap

layers. The apparent increase of the interface state density of HMOSFETs with the decrease of the cap layer thickness can be attributed to Ge segregation during gate oxide growth (see section 6.3). A screening effect of the Si cap layer is not observed.

Fig.'s 4.11-4.13 show a few factors which may cause systematic errors in the extraction of the interface state density. Using the pulse in fig. 4.10 with negligible fall and rise times, detailed information should be obtainable. The MOSFET is in accumulation at $V_g = V_l$ and in inversion at $V_g = V_h$. During t_m, some carriers captured at voltages between V_h and V_m are emitted. The number of emitted carriers depends on the emission rate and t_m. Varying t_m, the emission rate and the exact number of interface traps might be determined.

In fig. 4.11, the charge pumping current is given as a function of t_m for an n-MOSFET with a 12 nm thick gate oxide, $L = 6\ \mu m$, and $W = 1000\ \mu m$, $V_l = -2\ V$, $V_m = 0\ V$, and various V_h. Instead of decreasing exponentially, the charge pumping current forms as plateau for small t_m and relatively high V_h. As confirmed by "HP Advanced Design System" simulations, the plateau is due to the time it takes the wide gate to decharge. The charge pumping current saturates as a function of V_h. If t_m is smaller than the time it takes the gate voltage to reach the value of the saturation voltage, it has no effect on I_{cp}. The plateau gets wider with increasing V_h. As the emission process is not solely determined by the applied pulse, equation (4.26) is not suitable for these devices with an extremely wide gate. However, such devices will be used in chapter 8 about the $SiGe$ CMOS/Ge^+ implantation batch as the charge pumping current is not stable for devices with a small gate and influenced by the contact between the probes and the metal pads. Instead, equation (4.25) will be used although it might underestimate the interface state density slightly as it does not take into account the emission processes.

Fig. 4.12 shows the saturation of the charge pumping current as a function of T_p. I_{cp} does not increase further for T_p lower than a few hundreds of nanoseconds because the gate voltage is not able to follow the applied voltage. The value T_p for which I_{cp} saturates is similar to the width of the plateau in fig. 4.11. Fig. 4.13 illustrates how the charge pumping current increases overproportional for long gate lengths because the minority carriers in the inversion layer are not completely evacuate while the gate voltages passes from inversion to accumulation.

4.4 Parameter Extraction for Conventional MOSFETs

In the linear region of device operation for small drain voltages [33], the drain current equation for MOSFETs is given by:

$$I_{ds} = \frac{Q_i \cdot \mu_{eff} \cdot W \cdot (V_{ds} - R_{sd} \cdot I_{ds})}{L} \tag{4.27}$$

Where I_{ds}, Q_i, V_{ds}, R_{sd}, μ_{eff}, L, and W are respectively the drain current, the inversion charge per unit area, the drain voltage, the source and drain series resistance, the effective mobility, the effective channel length, and the effective channel width.

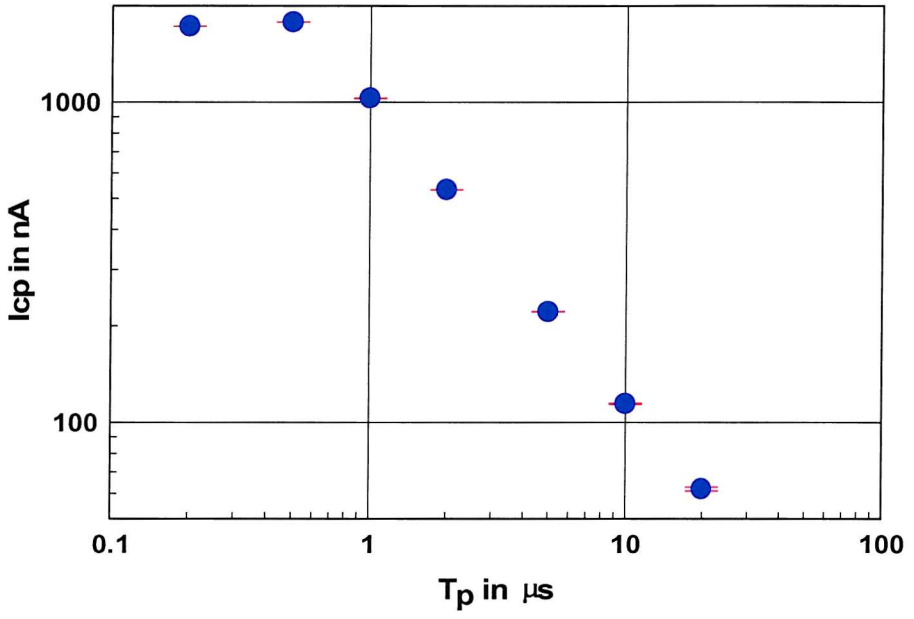


Figure 4.12: Charge pumping current as a function of T_p for an n-MOSFET with a 12 nm thick gate oxide, $L = 6\ \mu m$, $W = 1000\ \mu m$, $V_l = -2\ V$, $V_h = 5\ V$, and $t_m = 0\ \mu s$

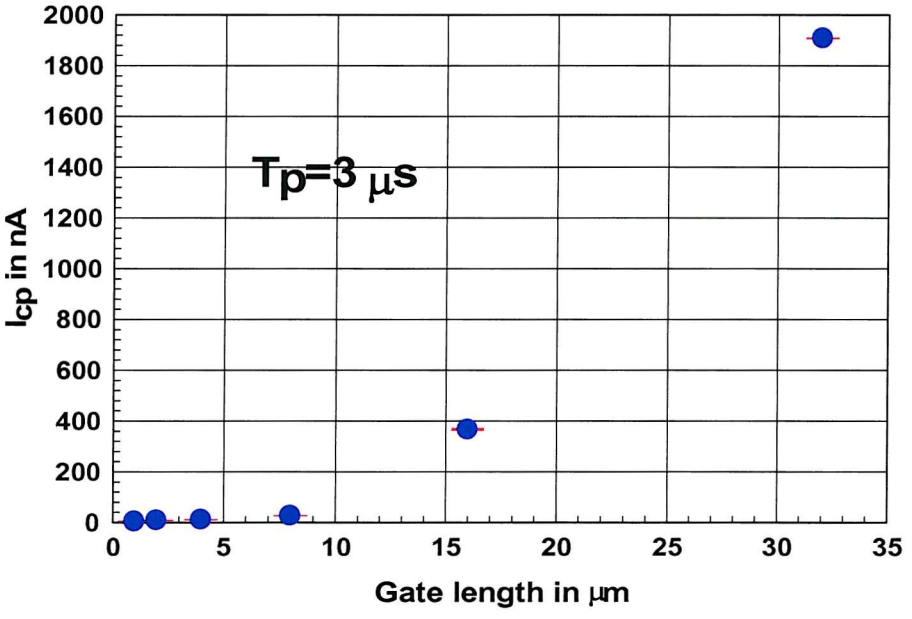


Figure 4.13: Charge pumping current as a function the gate length L for an n-MOSFETs with a 12 nm thick gate oxide, $W = 20\ \mu m$, $V_l = -2\ V$, $V_h = 5\ V$, $t_m = 0\ \mu s$, and $T_p = 3\ \mu s$

The effective mobility μ_{eff} can be described by [7]:

$$\mu_{eff} = \frac{\mu_o}{1 + \Theta_o(V_g - V_{th})} \tag{4.28}$$

Where Θ_o is the mobility degradation factor, and μ_o is the low field mobility.

The inversion charge can be described by (see chap. 3):

$$Q_i = C_{ox} \cdot (V_g - V_{th}) \tag{4.29}$$

Where V_g, V_{th}, and C_{ox} are respectively the gate voltage, the threshold voltage, and the gate oxide capacitance.

Using eq. (4.28) and eq. (4.29), equation (4.27) can be expressed as [24, 25]:

- for n-MOSFETs

$$I_{ds} = \frac{\beta(V_g - V_{th})V_{ds}}{1 + \Theta(V_g - V_{th})} \tag{4.30}$$

- for p-MOSFETs

$$I_{ds} = \frac{\beta(-V_g + V_{th})V_{ds}}{1 + \Theta(-V_g + V_{th})} \tag{4.31}$$

β is given by:

$$\beta = \frac{\mu_o \cdot C \cdot W}{L} \tag{4.32}$$

As typically wide devices are used, we do not have to distinguish between the effective and drawn channel width.

Θ depends on the source and drain series resistance $R_{sd} = R_s + R_d$ and the mobility degradation factor Θ_o:

$$\Theta = \beta R_{sd} + \Theta_o \tag{4.33}$$

The parameters given in eq.'s (4.30) and (4.31) are determined by curve fitting for a particular device, whereas the determination of the low field mobility μ_o, the mobility degradation factor Θ_o, the source and drain series resistance R_{sd}, and the difference between the drawn and effective channel length $\Delta L = L_o - L$ require the measurement of several devices which are identical except for the gate length.

Determination of the Source and Drain Series Resistance and the Mobility Degradation Factor Using N sets of devices with different channel lengths which are identical in every other aspect, R_{sd} and Θ_o can be extracted by plotting Θ as a function of $\beta = \beta_1, \beta_2 \ldots \beta_N$:

$$\Theta = \beta R_{sd} + \Theta_o \tag{4.34}$$

Θ_o can be expressed according to equation (see appendix A):

$$\Theta_o = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i^2 \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\Theta_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i\Theta_i}{\Delta} \tag{4.35}$$

$$\sigma_{\Theta_o} = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i^2}{\Delta}} \tag{4.36}$$

R_{sd} is given according to equation (see appendix A):

$$R_{sd} = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i\Theta_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\Theta_i}{\Delta} \tag{4.37}$$

$$\sigma_{R_{sd}} = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}{\Delta}} \tag{4.38}$$

σ_i and Δ are defined by:

$$\Delta = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i^2 - (\sum_{i=1}^{N} \frac{1}{\sigma_i^2}\beta_i)^2 \tag{4.39}$$

$$\sigma_i = \sqrt{(\sigma_\Theta)_i^2 + (R_{sd}^*(\sigma_\beta)_i^2)} \tag{4.40}$$

Where R_{sd}^* is obtained from equation (4.37) when $(\sigma_\beta)_i$ is neglected. A more precise value can be obtained by iteration.

Determination of the Low Field Mobility and ΔL Using the same results, ΔL and μ_o can be extracted by plotting a graph of $1/\beta$ against the drawn channel length:

$$\frac{1}{\beta} = \frac{1}{\mu_o CW} L_o - \frac{1}{\mu_o CW}\Delta L = B \cdot L_o + A \tag{4.41}$$

A and B are given by (see appendix A):

$$A = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i^2 \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(\frac{1}{\beta})_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i(\frac{1}{\beta})_i}{\Delta} \tag{4.42}$$

$$B = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i(\frac{1}{\beta})_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(\frac{1}{\beta})_i}{\Delta} \tag{4.43}$$

The standard deviations σ_A and σ_B can be expressed as:

$$\sigma_A = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i^2}{\Delta}} \tag{4.44}$$

$$\sigma_B = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}{\Delta}} \tag{4.45}$$

Where Δ and σ_i are given by:

$$\Delta = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i^2 - (\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(L_o)_i)^2 \tag{4.46}$$

$$\sigma_i = (\sigma_{\frac{1}{\beta}})_i \tag{4.47}$$

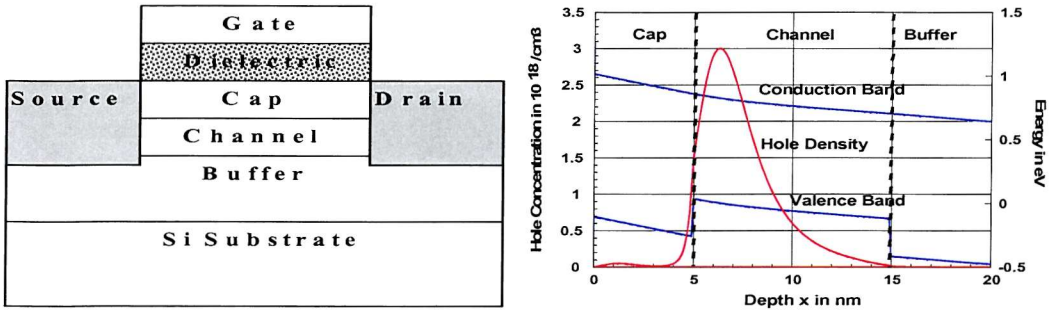
Thus, μ_o and ΔL can be calculated:

$$\mu_o = \frac{1}{BCW} \tag{4.48}$$

$$\sigma_{\mu_o} = \frac{\sigma_B}{B^2 CW} \tag{4.49}$$

$$\Delta L = A\mu_o CW = \frac{A}{B} \tag{4.50}$$

$$\sigma_{\Delta L} = \sqrt{(\frac{\sigma_A}{B})^2 + (\frac{A\sigma_B}{B^2})^2} \tag{4.51}$$



(a) (b)

Figure 4.14: (a) Cross section of a HMOSFET (b) Space charge region of a p-MOSFET with a 6 nm gate oxide, a 5 nm Si cap, a 10 nm 40 % Ge channel, and a Si buffer at a gate voltage $V_g = -1\ V$

4.5 Parameter Extraction Routine for HMOSFETs

HMOSFETs have an epitaxial $Si_{1-x}Ge_x$ layer underneath the gate. The $Si_{1-x}Ge_x$ layer can be divided into a buffer layer, a channel and a cap layer (see fig. 4.14 (a)). Typically, the Ge content in the strained heterochannel is higher for the p-MOSFETs and lower for the n-MOSFETs than the Ge content in the unstrained cap and buffer layers. The hole mobility for p-MOSFETs and the electron mobility for n-MOSFETs are therefore enhanced while the carriers are confined in the hetero channel. However, the confinement might not be perfect and the gate capacitance might vary as a function of the gate voltage (see fig. 4.14 (b)).

In the linear region of operation, for small drain voltages, the MOSFET can be described by a resistance R_{to} that is the sum of the source and drain series resistances R_{sd} and the channel resistance R_{ch} [36] according to eq. (4.27):

$$R_{to} = R_{sd} + R_{ch} \tag{4.52}$$

Whereas the channel resistance R_{ch} is dependent on the gate voltage, inversely proportional to the channel width, and proportional to the channel length, R_{sd} is independent of the gate voltage and the channel length, but inversely proportional to the channel width.

The channel resistance R_{ch} is given by:

$$\frac{1}{R_{ch}} = \frac{W}{L}\int_0^d \mu(x)\rho(x)dx = \frac{W \cdot Q_i \cdot \mu_{eff}}{L} \tag{4.53}$$

Where μ, ρ, x, and d are respectively the mobility, the charge density, the depth in the substrate direction (the oxide-silicon interface is situated at $x = 0$, see fig. 4.14 (b)), and the end of the space charge region.

Taking into account ΔL, the difference between drawn channel length L_o and effective channel length L, R_{ch} is given by:

$$R_{ch} = \frac{L_o - \Delta L}{W} r_{ch}(V_g - V_{th}) \tag{4.54}$$

Where r_{ch} is the channel sheet resistance and given as a function of the gate drive as the dispersion in the values of the threshold voltage renders the determination of device parameters otherwise difficult.

The total resistance can be written as:

$$\begin{aligned} R_{to} &= R_{sd} + R_{ch} \\ &= R_{sd} - \frac{\Delta L}{W} r_{ch}(V_g - V_{th}) + \frac{L_o}{W} r_{ch}(V_g - V_{th}) \\ &= R_{indL} + L_o \cdot R_{depL} \end{aligned} \tag{4.55}$$

Where R_{indL} is in contrast with R_{depL} independent of L.

Determination of the Threshold Voltage The extrapolation method for the determination of the threshold voltage is not necessarily suitable for HMOSFETs because the dependence of the mobility on the gate voltage is in general not known. Instead, the constant current method seems to be most appropriate for the determination of the threshold voltage. In the constant current method, the drain current is measured for a low drain voltage. At the threshold voltage, the drain current I_{ds} is given by [7, p. 438]:

$$I_{ds} = I_{th} \frac{W}{L} \tag{4.56}$$

For drain voltages typically lower than 0.1 V, $I_{th} = 1 \cdot 10^{-7}$ A results often in a good fit.

Determination of R_{sd} and ΔL Using N devices with different channel lengths which are identical in every other aspect, $R_{depL}(V_g - V_{th}) = R_{depL}$ and R_{indL} can be extracted by plotting R_{to} as a function of $L = L_1, L_2 \ldots L_N$ for a particular gate drive. R_{indL} can be expressed as (see appendix A):

$$R_{indL} = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i^2 \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (R_{to})_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i (R_{to})_i}{\Delta} \tag{4.57}$$

$$\sigma_{R_{indL}} = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i^2}{\Delta}} \tag{4.58}$$

R_{depL} is given by (see appendix A):

$$R_{depL} = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i (R_{to})_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (R_{to})_i}{\Delta} \tag{4.59}$$

$$\sigma_{R_{depL}} = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}{\Delta}} \tag{4.60}$$

σ_i and Δ are defined by:

$$\Delta = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i^2 - (\sum_{i=1}^{N} \frac{1}{\sigma_i^2} (L_o)_i)^2 \tag{4.61}$$

$$\sigma_i = \sigma_{(R_{to})_i} \tag{4.62}$$

Varying the gate drive $V_g - V_{th}$, R_{indL} can be expressed as a function of R_{depL}:

$$R_{indL} = R_{sd} + \Delta L \cdot R_{depL} \tag{4.63}$$

R_{sd} can be expressed as (see appendix A):

$$R_{sd} = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i^2 \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{indL})_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i (R_{indL})_i}{\Delta} \tag{4.64}$$

$$\sigma_{R_{sd}} = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i^2}{\Delta}} \tag{4.65}$$

ΔL is given by (see appendix A):

$$\Delta L = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i (R_{indL})_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{indL})_i}{\Delta} \tag{4.66}$$

$$\sigma_{\Delta L} = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}{\Delta}} \tag{4.67}$$

σ_i and Δ are defined by:

$$\Delta = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i^2 - (\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(R_{depL})_i)^2 \tag{4.68}$$

$$\sigma_i = \sqrt{(\sigma_{R_{indL}})_i^2 + (\Delta L^* (\sigma_{R_{depL}})_i^2} \tag{4.69}$$

Where ΔL^* is obtained from equation (4.66) when $(\sigma_f)_i$ is neglected. A more precise value can be obtained by iteration.

Determination of the Channel Sheet Resistance and the Effective Mobility Alternatively, parameter extraction routines described in [33, 36-40] can be applied to HMOSFETs in order to extract the effective channel length and the source and drain series resistance as long as they are not based on unjustified assumptions about the dependence of the inversion charge on the gate voltage. Once the effective channel length and the source and drain series resistance are known, the channel sheet resistance, which is inversely proportional to the effective mobility, can be calculated.

The channel resistance is calculated according to:

$$R_{ch} = R_{to} - R_{sd} \tag{4.70}$$

$$\sigma_{R_{ch}} = \sqrt{\sigma_{R_{to}}^2 + \sigma_{R_{sd}}^2} \tag{4.71}$$

The channel sheet resistance can also be extracted for devices with an extremely long gate for which R_{sd} can be neglected. This approximation is very common for split-CV measurements which require large MOSFETs.

The channel sheet resistance for a device of a particular channel length is given by:

$$r_{ch} = \frac{W}{L} R_{ch} \tag{4.72}$$

$$\sigma_{r_{ch}} = \sqrt{(\frac{W}{L}\sigma_{R_{ch}})^2 + (\frac{W}{L^2}R_{ch}\sigma_{\Delta L})^2} \tag{4.73}$$

The measurements for devices with different channel lengths can be combined using the weighted averages:

$$r_{ch} = \frac{\sum_{i=1}^{N} \frac{1}{\sigma^2_{(r_{ch})_i}}(r_{ch})_i}{\sum_{i=1}^{N} \frac{1}{\sigma^2_{(r_{ch})_i}}} \tag{4.74}$$

$$\sigma_{r_{ch}} = \frac{1}{\sum_{i=1}^{N} \frac{1}{\sigma^2_{(r_{ch})_i}}} \tag{4.75}$$

The channel sheet resistance r_{ch} can also be written as:

$$\frac{1}{r_{ch}} = \int_0^d \mu(x) \cdot \rho_i(x)dx = \mu_{eff} \cdot Q_i \tag{4.76}$$

Where Q_i, μ, ρ_i, x, and d are respectively the inversion charge, the mobility, the inversion charge density, the depth in the substrate direction (the dielectric-silicon interface is situated at $x = 0$, see fig. 4.14), and the depth of the space charge region.

Knowing the channel sheet resistance, the accurate extraction of the effective mobility requires the determination of the inversion charge which might be approximated by:

$$Q_i(V_g - V_{th}) = C \cdot (V_g - V_{th}) \tag{4.77}$$

Where V_g is the gate voltage, V_{th} the threshold voltage and C the series capacitance of the gate oxide C_{ox} and the Si cap C_{cap} per unit area as all carriers would ideally be confined in the SiGe channel:

$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{1}{C_{cap}} \tag{4.78}$$

Substituting equation (4.77) into equation (4.76), the effective mobility can be expressed as:

$$\mu_{eff} = \frac{1}{C \cdot (V_g - V_{th}) \cdot r_{ch}} \tag{4.79}$$

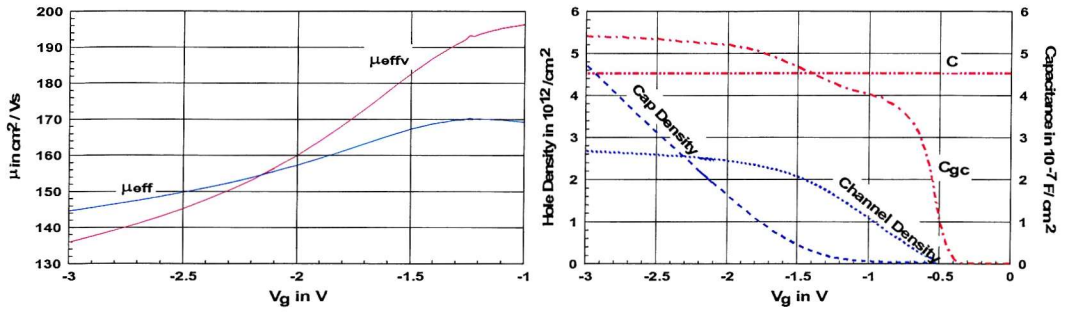
Alternatively, the capacitance associated with the inversion charge can be obtained from split-CV measurements [1-6]. Substituting equation (4.6) into equation (4.76) yields the following equation for the effective mobility μ_{effv} which takes account of the variable capacitance:

$$\mu_{effv} = \frac{1}{r_{ch} \cdot \int_{V_{fb}}^{V_g} C_{gc}(V)dV} \tag{4.80}$$

If a split-CV method is used which takes into account the interface traps, the drain voltage, expression (4.80) is exact except for the neglect of the diffusion current.

4.6 Validity of the Constant Capacitance Approximation

The error in the mobility extraction for a typical p-channel HMOSFET with a thin gate oxide is examined [41] when the constant capacitance approximation is applied. A 1D Schroedinger-Poisson solver [42] is used to calculate the hole density. It takes account of QM effects and includes the strain dependence of the light and heavy hole masses.



(a) (b)

Figure 4.15: (a) Comparison between μ_{effv} and μ_{eff} for a p-MOSFET with a 6 nm gate oxide, 5 nm cap, and a 10 nm 40 % Ge channel (b) C, C_{gc} and hole density in the Si cap and SiGe channel for a p-MOSFET with a 6 nm gate oxide, 5 nm cap, and a 10 nm 40 % Ge channel

The band offset [43] in the valence band of the SiGe channel is described by (see fig. 4.14 (b)):

$$\Delta E_v = -0.74x \ (eV) \tag{4.81}$$

where x is the Ge fraction.

The effective light hole m_{lh} and heavy hole m_{hh} masses fitted to the density of states effective mass at 300 K obtained from k.p theory [44] are given by:

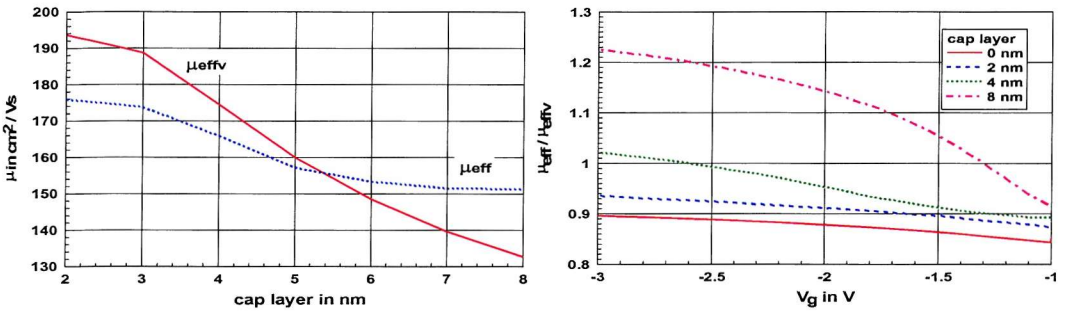
$$m_{lh} = 0.255 - 0.334x + 0.215x^2 \tag{4.82}$$

$$m_{hh} = 0.927 - 2.266x + 1.827x^2 \tag{4.83}$$

Typical hole mobilities of 100 cm^2/Vs and 200 cm^2/Vs are respectively used for μ_{Si} and μ_{SiGe}. The 10 nm thick heterochannel has a Ge content of 40 %. The oxide is 6 nm thick whereas the thickness of the Si cap is varied between 0 and 8 nm. The substrate is uniformly doped ($1 \cdot 10^{17}\ cm^{-3}$). The threshold voltage is defined as the voltage at which the hole density reaches $1 \cdot 10^{11} cm^{-2}$.

Fig. 4.15 (b) compares μ_{effv} and μ_{eff} as a function of the gate voltage for a HMOSFET with a 5 nm cap layer. For a small gate drive, the carriers are confined in the the SiGe channel and the constant capacitance approximation underestimates the effective mobility as for conventional MOSFETs due to the contribution of the inversion layer capacitance associated with the inversion charge. At higher gate drive, the Si cap is populated. Then, C_{gc} (see fig. 4.15 (a)) and μ_{effv} are incorrectly estimated.

Fig. 4.16 (a) illustrates the variation of μ_{eff} with respect to μ_{effv} as a function of the cap layer thickness at $V_g = -2V$. For thicker cap layers, the effective mobility can be severely overestimated by the constant capacitance approximation (see fig. 4.16 (b)). The quotient μ_{eff}/μ_{effv} varies 30 % for the devices with an 8 nm cap layer over a gate voltage range of 2 V.



(a) (b)

Figure 4.16: (a) Comparison between μ_{effv} and μ_{eff} for a p-MOSFET with a 6 nm gate oxide, and a 10 nm 40 % Ge channel at a gate voltage $V_g = -2\ V$ (b) μ_{eff}/μ_{effv} as a function of the gate voltage for p-MOSFETs with a 6 nm gate oxide, and a 10 nm 40 % Ge channel

4.7 Conclusion

In this chapter, a few techniques which are used in this work for the characterization of HMOSFETs have been presented. Charge pumping measurements are a useful technique for the characterization of the oxide interface of HMOSFETs which may have interface traps due to the presence of Ge. The speed performance can be assessed by s-parameter measurements. The presented mobility extraction routine which may make use of s-parameter extraction in case that the VNA possesses the required sensitivity at low frequencies is ideal for HMOSFETs.

HMOSFETs require in general a mobility extraction routine which takes account of the variable gate capacitance. Schroedinger-Poisson simulations show that the constant capacitance approximation produces a considerable error in the mobility extraction for typical HMOSFETs. The effective mobility extracted with the constant capacitance approximation can vary 30.0 % with respect to the correct value for a 40 % SiGe MOSFET with a 6 nm thick oxide and 8 nm thick cap layer over a voltage range of 2 V.

Bibliography

[1] M. Liang, J. Y. Choi, P. Ko, and C. Hu, "Inversion-Layer Capacitance and Mobility of Very Thin Gate-Oxide MOSFETs", IEEE Trans. Electron Devices, Vol. ED-33, No. 3, pp. 409-413, 1986

[2] S.-I. Takagi, M. Takayanagi, and A. Toriumi, "Impact of Electron and Hole Inversion-Layer Capacitance on Low Voltage Operation of Scaled n- and p-MOSFETs", IEEE Trans. Electron Dev., Vol. 47, No. 5, pp. 999-1005, 2000

[3] C. G. Sodini, T. W. Ekstedt, and J. L. Moll, "Charge Accumulation and Mobility in Thin Dielectric MOS Transistors", Solid-St. Electron., Vol. 25, No. 9, pp. 833-841, 1982

[4] J. Koomen, "Investigation of the MOST Channel Conductance in Weak Inversion", Solid-St. Electron., Vol. 16, pp. 801-810, 1973

[5] P. M. Chow and K. Wang, "A New AC Technique for Accurate Determination of Channel Charge and Mobility in Very Thin Gate MOSFET's", IEEE Trans. Electron Devices, Vol. ED-33, No. 9, pp. 1299-1304, 1986

[6] C. L. Huang, J. V. Faricelli, N. Arora, "A New Technique for Measuring MOSFET Inversion Layer Mobility", IEEE Trans. Electron Dev., Vol. 40, No. 6, pp. 1134-1139, 1993

[7] N. Arora, "MOSFET Models for VLSI Circuit Simulation", Springer-Verlag, New York, 1993

[8] C. L. Yip, "High-Frequency Circuit Design and Measurements", Chapman and Hall, London, 1990

[9] R. K. Hoffmann, "Handbook of Microwave Integrated Circuits", Artech House, Norwood, 1987

[10] N. d'Halleweyn, "Caracterisation et Modelisation des Transistor MOS/SOI Destines a Fonctionner en Hyperfrequences", Memoire de fin d'etudes, INPG, 1997

[11] H. Heuermann and B. Schiek, "Procedures for the Determination of the Scattering Parameters for Network Analyzer Calibration", IEEE Trans. Instr. Meas., Vol. 42, No. 2, pp. 528-531, 1993

[12] G. J. Scalzi, A. J. Slobodnik, and G. A. Roberts, "Network Analyzer Calibration Using Offset Shorts", IEEE Trans. Microwave Theory Tech., Vol. 36, No. 6, pp. 1097-1100, 1988

[13] K. J. Silvonen, "A General Approach to Network Analyzer Calibration", IEEE Trans. Microwave Theory Tech., Vol. 40, No. 4, pp. 754-759, 1992

[14] C. Wan, B. Nauwelears, W. de Raedt, and M. van Rossum, ""1 Thru+2.5 Reflects:" A new technique for the de-embedding MIC/MMIC device measurements in fixed-length fixtures ", EuMC, pp. 174-177, Prague, 1996

[15] R. B. Marks, "A Multiline Method of Network Analyzer Calibration", IEEE Trans. Microwave Theory Tech, Vol. 39, No. 7, pp. 1205-1214, 1991,

[16] H. Heuermann and B. Schiek, "Robust Algorithms for Txx Network Analyzer Self-Calibration Procedures", IEEE Trans. Instr. Meas., Vol. 43, No. 1, pp. 18-23, 1994

[17] S. P. Yeo and M. Cheng, "Network analyzer calibration for measurements near waveguide cutoff frequency", Microwave J., Vol. 42, No. 3, pp. 134-142, 1999

[18] J. Pence, "Technique verifies LRRM calibrations on GaAs substrates", Microwaves & RF, Vol. 33, No. 1, pp. 69-76, 1994

[19] W. Sansen and P. Vandeloo, "Modelling the MOS Transistor at High Frequencies", Electronics Lett., Vol. 22, No. 15, pp. 810-812, 1986

[20] R. Singh, A. Juge, R. Joly, and G. Morin, "An investigation into the Nonquasi-Static Effects in MOS Devices with On-Wafer S-Parameter Techniques", Proc. IEEE Int. Conference on Microelectronic Test Structures, Vol. 6, pp. 21-25,1993

[21] C. Raynaud, J. Gautier, G. Guegan, M. Lerme, E. Playez, and G. Dambrine, "High Frequency Performance of Submicrometer Channel Length Silicon MOSFET's", IEEE Electron Device Lett., Vol. 12, No. 12, pp. 667-669, 1991

[22] C. T. Wang, "Hot Carrier Design Considerations for MOS Devices and Circuits", Van Nostrand Reinhold, New York, 1992

[23] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. de Keersmaecker, "A Reliable Approach to Charge-Pumping Measurements in MOS Transistors", IEEE Trans. Electron Dev., Vol. ED-31, No. 1, pp. 42-53, 1984

[24] N. S. Saks and M. G. Ancona, "Determination of Interface Trap Capture Cross Sections Using Three-Level Charge Pumping", IEEE Electron Dev. Lett., Vol. 11, No. 8, pp. 339-342, 1990

[25] A. B. M. Elliot, "The Use of Charge Pumping Currents to Measure Surface State Densities in MOS Transistors", Sol.-St. Electron., Vol. 19, pp. 241-247, 1976

[26] C. R. Viswanathan and V. Ramgopal Rao, "Application of Charge Pumping Technique for sub-micron MOSFET Characterization", Microelectronic Eng., Vol. 40, pp. 131-146, 1998

[27] J. G. Simmons and L. S. Wei, "Theory of dynamic charge current and capacitance characteristics in MIS systems containing distributed surface traps", Solid.-St. Electron., Vol. 16, pp. 53-66, 1973

[28] G. Groeseneken and H. E. Maes, "Basics and applications of charge pumping in submicron MOSFETs", Microelectronics Reliability, Vol. 38, pp. 1379-1389, 1998

[29] Y. Chu, D.-W. Lin, and C.-Y. Wu, "A New Charge-Pumping Technique for Profiling the Interface-States and Oxide-Trapped Charges in MOSFET's, IEEE Trans. Electron Dev., Vol. 47, No. 2, pp. 348-353, 2000

[30] J. S. Brugler and P. G. A. Jespers, "Charge Pumping in MOS Devices", IEEE Trans. Electron Dev., Vol. ED-16, No. 3, pp. 297-302, 1969

[31] W. L. Tseng, "A new charge pumping method of measuring $Si - SiO_2$ interface states", J. Appl. Phys., Vol. 62, No. 2, pp. 591-599, 1987

[32] P. Masson, J.-L. Autran, and G. Ghibaudo, "An improved time domain analysis of the charge pumping current", J. Non-Crystalline Solids, Vol. 280, pp. 255-260, 2000

[33] G. Niu, J. D. Cressler, S. J. Mathew, and S. Subbanna, "A Total Resistance Slope-Based Effective Channel Mobility Extraction Method for Deep Submicrometer CMOS Technology", IEEE Trans. Electron Devices, Vol. 46, No. 9, pp. 1912-1914, 1999

[34] R. Oakley, and R. Hocking, "CASMOS- an Accurate MOS Model with Geometry-Dependant Parameters:1", IEE Proc., Vol. 128, Part 1, No. 6, 1981

[35] A. Waite, "Elevated Source/ Drain MOSFETs for Deep Submicron VLSI", Ph.D. thesis, University of Southampton, May 1999

[36] K. Terada and H. Muta, "A New Method to Determine Effective MOSFET Channel Length", Jap. J. Appl. Phys., 18 (5), pp. 953-959, 1979

[37] J. G. J. Chern, P. Chang, R. F. Motta, and N. Godhino, "A New Method to Determine MOSFET Channel Length", IEEE Elect. Dev. Lett., EDL-1(9), pp. 170-173, 1980

[38] K. K. Ng and J. Brews, "Measuring the Effective Channel Length of MOSFETs", IEEE Circuits and Devices Magazine, 6, pp. 33-38, 1990

[39] J. Whitfield, "A Modification on "An improved Method to Determine MOSFET Channel Length"", IEEE Electron Dev. Lett., Vol. EDL-6, No. 3, pp. 109-110, 1985

[40] Y. Taur, "MOSFET Channel Length: Extraction and Interpretation", IEEE Trans. Electron Dev., Vol. 47, No. 1, pp. 160-170, 2000

[41] U. N. Straube, A. G. R. Evans, G. Braithwaite, S. Kaya, J. Watling, and A. Asenov, "On the mobility extraction for HMOSFETs", Sol.-St. Electron., Vol. 45, pp. 527-529, 2000

[42] I. H. Tan, G. L. Snider, and E. L. Hu, "A sef-consistent solution of the Schroedinger-Poisson equations using a nonuniform mesh", J. Appl. Phys., Vol. 68, No. 8, pp. 4071-4076, 1990

[43] G. Abstreiter , "Physics and Perspectives of Si/Ge Heterostructures and Superlattices", Phys. Scr., pp. 42-45, Vol. T49A, 1993

[44] J. R. Watling, A. Asenov, and J.R. Barker, "Efficient Hole Transport Model in Warped Bands for Use in the Simulation of Si/SiGe MOSFETs", Extended Abstracts of IWCE-6, No. 98EX116, pp.96, Osaka, 1998

Chapter 5

Vapor-Phase Deposition

In IC fabrication, vapor-phase methods are predominant in the deposition of thin films. Vapor-phase epitaxy methods can be classified into chemical (CVD) and physical (PVD) vapor deposition methods. Physical vapor deposition methods do not involve chemical reactions. Nevertheless, the distinction begins to blur because CVD growth techniques are extended to ultrahigh vacuum and MBE, a physical vapor deposition method, can use gaseous sources.

This chapter focuses on the application of vapor-phase methods relevant to conventional processing and this work. Chemical vapor phase deposition methods are used in conventional circuit manufacturing for the deposition of dielectric, polysilicon, and epitaxial silicon layers whereas sputtering is the only common physical vapor phase deposition method. However, molecular beam epitaxy, a physical vapor phase deposition method, is ideal for the growth of SiGe heterostructures.

5.1 Chemical Vapor Deposition

Chemical vapor deposition [1, pp. 312; 2] is a material synthesis process whereby the constituents of the vapor phases react chemically near or on the substrate surface to form a nonvolatile solid film. All CVD methods are based upon the same growth process and possess a similar growth reactor. However, they use different temperature and pressure regimes that are the most suitable for certain layers. Some CVD techniques enhance their capability making use of additional physical effects.

5.1.1 CVD Growth and Reactor

A detailed understanding of CVD is a formidable task because it involves chemical reactions and flow dynamics. Chemical vapor deposition is described in the laminar flow pattern where the input reactant gases are flowing in one direction and passing over the substrate to the exhaust. Then, a growth process consists of the following steps [2, pp. 108]:

1. Introduction of the reactant species to the substrate region
2. Transfer to the substrate surface
3. Absorption or chemisorption on the substrate surface
4. Surface diffusion, site accommodation, chemical reaction, and layer deposition
5. Desorption of residual reactants and by-products
6. Transfer of residual reactants and by products from the substrate surface
7. Removal of residual reactants and by-products from the substrate region

The steps can occur consecutively or in series. The slowest step is called the rate-limiting step and determines the overall rate of the growth process. A growth process that is dominated by 3, 4 or 5 is called a surface-controlled process, whereas a growth process that is dominated by step 1 and 2 is called a mass-transport-controlled process.

Mass-transport-limited processes occur at high temperatures when temperature has little influence on the growth rate. The deposition rate is controlled by the concentration of the deposition gases. Mass-transport-limited processes require a uniform concentration of reactants in the bulk gas.

In contrast, surface-controlled processes occur at low temperatures when the reaction rate determines growth. Then, the deposition rate R is given by the Arrhenius equation [2, p. 228]:

$$R = A \exp\left(-E_a/kT\right) \tag{5.1}$$

Where the frequency factor, the activation energy, temperature and the Boltzmann constant are respectively given by A , E_a, T and k. E. g. the activation energy E_a for polysilicon is 1.7 eV. The uniformity of reactants in the bulk gas is not critical, but temperature control has to be excellent.

The deposition of a thin layer on the substrate can be decomposed into nucleation and growth. In the nucleation process, a newly formed phase grows to a critical size and becomes stable (nucleus). The nucleation process can be homogenous or heterogenous. The heterogenous nucleation process occurs in contrast to the homogenous nucleation process in the presence of other objects like the substrate surface and is thermodynamically always preferable to homogenous nucleation.

The chemical reactions lead to the formation of a solid material as well on the substrate surface (heterogenous reaction) as in the gas phase (homogeneous reaction). Heterogenous reactions occur selectively on the the heated surface and produce therefore good quality films. Homogenous reactions are not desirable because they form gas-phase clusters of the depositing material and result in poorly adhering, low density films with defects.

The design and operation of CVD reactors are determined by various factors. The CVD reactors can be classified according to the method that is used to heat up the wafers and the

pressure regime of operation. The reactor geometry is determined by the pressure regime, energy source and throughput. They can be further split into subgroups defined by the method of heating [2, pp. 210]:

- Resistance heating
- RF induction heating
- Heating with a glow discharge (plasma)
- Heating by photon energy

Energy is transferred either to the reactant gases or directly to the substrate. Reaction chambers where the whole reaction tube is surrounded by resistance-heated coils and not only the wafers become hot are known as hot-wall reactors. In these systems, chemical reactions occur as well on the wafer surface as on the reaction chamber walls. Hence, these systems have to be cleaned regularly. In contrast, energy input via RF induction or infrared lamps mounted within the reactor does not heats up the reaction chamber walls appreciably (cold-wall systems). In some cold wall systems, even means of cooling the walls are implemented. Most gases used in film deposition are toxic or at least hazardous. They can be categorized into four different groups:

- pyrophoric
- poisonous
- corrosive
- dangerous combinations of gases

Combinations as silane with halogens, silane with hydrogen and oxygen with hydrogen cause safety problems. Silane can react with air to form solid products causing particle contamination or plugging of the pipes. More over, an inert carrier gas is often used in order to improve the uniformity of the reactant species.

5.1.2 Different CVD Methods

In this section, a short overview over the most common CVD techniques is given. The basics of each method and its applications are discussed.

Atmospheric Pressure CVD Atmospheric pressure CVD (APCVD) is an early and simple CVD technique. APCVD takes place in the mass-transport-limited regime and has therefore large reaction rates, but it does not allow a high throughput as the wafers cannot be vertically stacked. APCVD is suitable for thick dielectrics. It operates at atmospheric pressure. Deposition rates can be in excess of 1000 Å/ min. Particle formation due to homogenous reactions is a major drawback in APCVD. In order to achieve a good morphology, APCVD requires a diluent, e. g. N_2 for SiO_2-deposition.

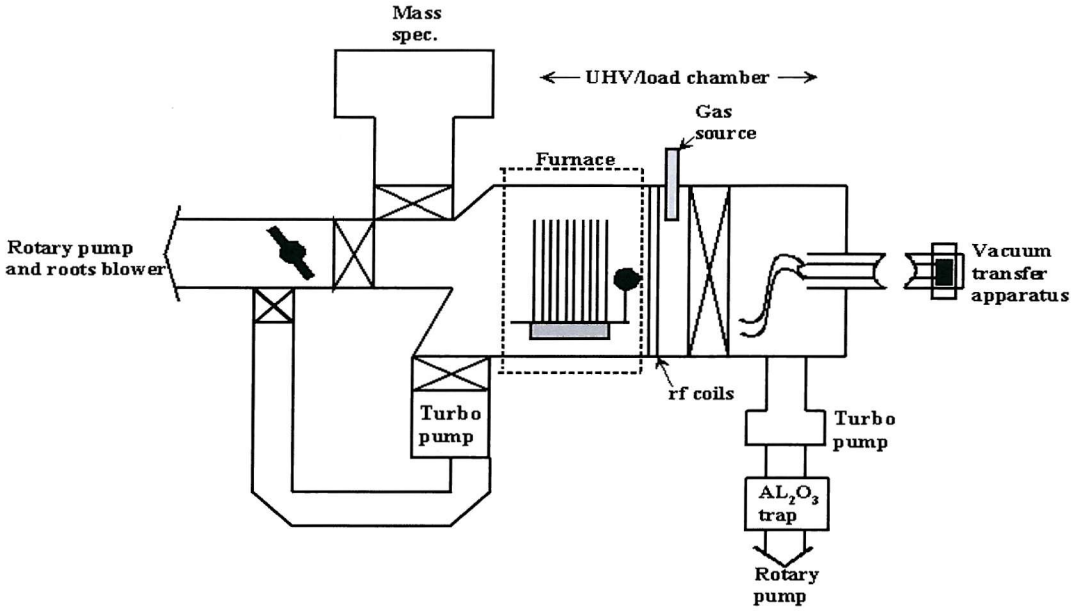


Figure 5.1: Schematic of an UHVCVD system

Low Pressure CVD Low pressure CVD (LPCVD) enables uniform step coverage, precise control of composition and structure, low-temperature processing, high deposition rate and high throughput. LPCVD is more popular than APCVD due to its superior properties. It requires no carrier gas as gas phase nucleation is reduced by low pressures. LPCVD operates in general in the surface-reaction-rate-limited mode. Therefore, the wafers can be stacked vertically at very close spacing. The chamber holds up to 200 wafers.

Plasma Enhanced CVD Plasma enhanced CVD (PECVD) does not rely solely on thermal energy and, therefore, allows film deposition at very low substrate temperatures. The energy provided to the reactant species by an RF plasma allows them to react and to diffuse further along the surface at low temperature. The RF frequency is normally less than 1 MHz. The process fills small feature sizes easily. Although particle contamination and limited capacity pose still major concerns, it has been used for the growth of SiGe layers [3].

Ultra High Vacuum CVD The UHVCVD [4-7] growth reactor is very similar to a conventional LPCVD reactor (see fig. 5.1). It is not commonly used in industrial fabrication. In comparison to conventional CVD, ultrahigh vacuum CVD (UHVCVD) is based on a ultrapure growth environment throughout the deposition process (low limits of oxygen and carbon). The base pressure is typically as low as 10^{-9} torr.

UHVCVD is suitable for the growth of high quality epitaxial Si and SiGe layers. Growth occurs in the range of 550^oC in about a mtorr of silane instead of 850^oC or higher for conventional CVD. Thus, little silane is decomposed in the gas phase. The low growth rate of UHVCVD (a few Å/ min.) is acceptable because up to 35 wafers can be loaded in the current generation

research reactors.

Further CVD Methods Metal organic CVD [9], laser ablation CVD, hybrid excitation CVD, photon-induced CVD [9, 10], electron cyclotron resonance CVD, rapid thermal CVD [4] are examples for further CVD methods. Photon-induced CVD (PHCVD) may allow film deposition at very low temperatures without the disadvantages of PECVD. PHCVD employs either a laser, UV or IR lamp. In UV lamp reactors, mercury vapor absorbs the UV light in the CVD chamber [2, p. 249]. The energy that is transferred to the silane lowers the activation energy for the Si film deposition.

Electron cyclotron resonance CVD (ECRCVD) is very similar to PECVD. However, the RF-field in order to generate a plasma is replaced by a microwave power. ECRCVD requires less energy for the generation of a plasma than PECVD resulting in a low defect density of deposited films.

Metal organic CVD (MOCVD) is a competitor to MBE as it is also able to grow thin, atomically abrupt, extremely perfect epilayers. MOCVD uses organometallics. It is used for the fabrication of GaAs devices, but not for the fabrication of SiGe devices.

Rapid thermal CVD combines rapid thermal processing technology and CVD. The wafer is heated by lamps above and below the flow tube. It can be used for SiGe heterostructures.

5.1.3 Deposition of Thin Films by CVD

Whereas sputtering is the most suitable methods for metal and alloy deposition, CVD is mainly used for the deposition of dielectric , poly- and silicon films. The deposition of the most common layers and SiGe layers by CVD is described in this section.

Deposition of Dielectric Layers Dielectric films, such as silicon dioxide and silicon nitride, are used as isolation, mask, and passivation layers. Silicon dioxide is formed either thermally at high temperature in oxygen or by chemical reaction. Thermal oxide is stoichiometric, displays excellent properties and is used as gate oxide, whereas CVD oxide is used for most other applications (poly-metal interlevel dielectrics, intermetal dielectrics, dopant mask, passivation layers, diffusion sources). Silicon nitride films possess a higher dielectric constant than silicon dioxide and are not suitable as stand-alone intermetal dielectrics or poly-metal interlevel dielectrics (PMD). Requirements for dielectric CVD films are [2, pp. 215]:

- no residual constituents outgassing during later processing
- good step coverage
- ease of etching
- thermal stability

- good adhesion
- low dielectric constant (isolation)
- high breakdown field strength
- no moisture absorption

PMDs can be deposited and flowed at temperatures in excess of 800^oC, whereas dielectric films employed between metal layers (intermetal dielectrics) have to be processed below 450^oC due to the presence of Al.

Deposition of Silicon Dioxide Films Furnace oxide possess a unbeatable quality, but requires a high thermal budget. The quality of deposited silicon dioxide films is affected by pressure, deposition temperature, reactant concentration and dopant concentration. High temperature leads as well to an increase of the density of the deposited oxide and of the deposition rate for all CVD methods as to good step coverage.

Feeding dopant hydrides such as phosphine, arsine and diborane into the CVD reactor results in doped silicon dioxides. Trimethylphosphite (TMOP) is used for phosphorus-doped oxides. The viscosity decreases with the concentration of phosphorus and boron (PSG: Phosphosilicate glass, BPSG: borophosphosilicate glass). Doped dioxide used for the reflow process contain 6 to 9 wt % phosphorus. Unfortunately, silicon dioxide with high phosphorus concentration results in aluminium corrosion. Boron dopants can be added in order to reduce the reflow temperature further. The boron concentration should not exceed 5 % because the silicon dioxide films become unstable.

In the following chemical reactions, s, l, g means respectively solid, liquid, and gaseous. Silicon dioxide films can be formed by different chemical reactions:

- Below 500^oC, the following reaction is suitable for the growth of intermetal dielectrics by APCVD and LPCVD [2, p. 216]:

$$SiH_4(g) + O_2(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{5.2}$$

$$4PH_3(g) + 5O_2(g) \rightarrow 2P_2O_5(s) + 6H_2(g) \tag{5.3}$$

  The dielectrics exhibit high particle concentration and poor step coverage.

- An LPCVD process taking place at about 900^oC is [2, p. 217]:

$$SiCl_2H_2(g) + 2N_2O(g) \rightarrow SiO_2(s) + 2N_2(g) + 2HCl(g) \tag{5.4}$$

  The disadvantages of the particular process are that the chlorine damages the underlying polysilicon and that the deposition temperature is too high for the use as intermetal dielectric.

- The following reaction uses tetraethyl orthosilicate (TEOS, $Si(OC_2H_5)_4$) for the formation of silicon dioxide between 650 and 750^oC [2, p. 216]:

$$Si(OC_2H_5)_4(l) \rightarrow SiO_2(s) + by - products(g) \tag{5.5}$$

  TEOS silicon dioxide displays very good uniformity and step coverage.

- In PECVD, silicon dioxide can be formed according to [2, p. 244]:

$$SiH_4(g) + O_2(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{5.6}$$

$$SiH_4(g) + 4N_2O(g) \rightarrow SiO_2(s) + 4N_2(g) + 2H_2O(g) \tag{5.7}$$

  For PECVD, radicals are formed and further chemical reactions occur in the plasma discharge. The silicon dioxide is not perfectly stoichiometric and contains some by-products.

Deposition of Silicon Nitride Films Silicon nitride (Si_3N_4) possesses the ability to protect against the diffusion of water, dopants and impurities. In LOCOS, it is used as a mask and passivation layer. The most common silicon nitride deposition methods are:

- At temperatures between 700 and 800^oC, silicon nitride is deposited by APCVD according to [2, p. 225]:

$$3SiH_4(g) + 4NH_3(g) \rightarrow Si_3N_4(s) + 12H_2(g) \tag{5.8}$$

- It can also be deposited by LPCVD in the same temperature range [2, p. 225]:

$$3SiCl_2H_2(g) + 4NH_3(g) \rightarrow Si_3N_4(s) + 6H_2(g) \tag{5.9}$$

  The properties of both LPCVD and APCVD silicon nitride depends on total pressure, reactant concentrations, deposition temperature and temperature gradient. The deposition rate increases with temperature, total pressure and partial pressure of dichlorosilane, whereas ammonia reduces the deposition rate. Deposited silicon nitride is amorphous and can contain a large amount of hydrogen. High ammonia to dichlorosilane ratio and low deposition temperature increase the hydrogen content. The hydrogen atoms passivate the dangling bonds of silicon. The refractive index of deposited silicon nitride (2.01) is higher than the refractive index of silicon dioxide (1.46). The refractive index of silicon oxinitride can vary between both of these indices. Relatively thick silicon nitride layers ($> 200nm$) may crack because of high tensile stress.

- In PECVD, the reaction of silane with ammonia or nitrogen results in the formation of a silicon nitride film [2, p. 244]:

$$SiH_4(g) + NH_3(g) \rightarrow SiN:H(s) + 3H_2(g) \tag{5.10}$$

$$SiH_4(g) + N_2(g) \rightarrow 2SiN:H(s) + 3H_2(g) \tag{5.11}$$

Recently, thin CVD nitrides have been used extensively as dielectrics for storage capacitors of dynamic random-access memories (DRAMs) and as interpoly dielectrics for nonvolatile memory devices. This application of nitride oxides requires smooth surfaces that can be grown solely at low temperature.

Deposition of Polysilicon and Amorphous Silicon Polycrystalline films are widely deposited by LPCVD and PECVD according to [2, p. 227]:

$$SiH_4(g) \rightarrow Si(s) + 2H_2(g) \tag{5.12}$$

Deposition pressure, deposition temperature, silane content and dopant concentration affect the film properties. At temperatures between 500 and 650^oC, pure silane or 20 to 30 % silane, diluted with nitrogen, is fed into the LPCVD or PECVD reactor. The deposition is either surface-reaction-limited or mass-transport-limited. High temperature results in a rough surface and loosely adhering deposit with poor uniformity. The in situ doping of polysilicon is possible by adding phosphine, arsine, diborane to the reaction gases. The introduction of an inert gas into the growth chamber can improve the film uniformity. Diborane enhances the deposition rate of polysilicon films, whereas phosphine and arsine seriously reduce the deposition rate.

Deposited silicon films exhibit an amorphous structure for temperatures below 600^oC, whereas polysilicon is formed above 600^oC. A subsequent anneal is necessary for the activation of dopants or the crystallization of the amorphous film. A typical grain size for polysilicon films is between 0.03 and 0.1 μm increasing to the top surface. Annealing does not affect the size of polysilicon grains considerably. For amorphous silicon films, polysilicon grains of a size up to 0.2 μm grow during annealing.

The resistivity of dopant-implanted polysilicon films is affected by implant energy, implant species, implant dose, annealing temperature and annealing time. Some dopants diffuse to the boundaries of grains where they are segregated resulting in low resistivity.

Silicon Epitaxy Epitaxy [12, 13] is a process to grow a single crystal layer with well-controlled doping profile on a crystal substrate. Epitaxy enables the fabrication of bipolar transistors with high collector-substrate breakdown voltage and low collector resistance and the fabrication of MOSFETs with high channel mobility and high speed performance. In CMOS, latchup can be reduced considerably.

The term epitaxy means "upon-ordered" in Greek. Homoepitaxy occurs when the single crystal substrate and the single crystal layer are of exactly the same material. Heteroepitaxy occurs when the substrate and the single crystal layer are different in any aspect. In reality, almost all epitaxial processes are heteroepitaxial because there are always some differences between the substrate and the epilayer. However, it is broadly accepted that homoepitaxy is the process that occurs when substrate and epilayer are chemically identical.

In industry, conventional Si and selective epitaxy are exclusively employed. For chlorosilane chemistry, the overall reaction can be expressed as [2, p. 117]:

$$SiCl_4 + 2H_2 \rightarrow 4HCl \tag{5.13}$$

An alternative reaction is based on silane chemistry [2, p. 118]:

$$SiH_4 \rightarrow Si + 2H_2 \tag{5.14}$$

In selective epitaxial growth (SEG), a selective epitaxial layer is grown on a bare substrate while no silicon (type 1) or polysilicon (type 2) is grown on a dielectric mask. High surface mobility and $SiCl_4$ based chemistry is characteristic for type 1. Adding HCl or Cl_2 increases the selectivity. Type 2 is typical of SiH_4 based chemistry.

Conventional epitaxy is carried out at temperatures between 1100 and 1300^oC. Silane enables the growth of silicon epitaxial layers at lower temperatures than chlorosilane. However, the Si layers display poor surface morphology and might even be polycrystalline due to homogenous reactions. More over, SiH_4 is explosive and susceptible to oxidation. Metals are often incorporated into the epitaxial layer because no etching process occurs in silane chemistry.

Epitaxy might result in pattern shift, distortion and washout. All these effects are undesirable because they cause the displacement of alignment marks. Pattern shift is less severe for high growth temperature, low growth pressure and rate. SiH_4 suppresses pattern shift that is very important for $SiCl_4$ [2, p. 123], but enhances pattern distortion. In contrast to pattern shift, pattern distortion augments with increasing growth temperature and diminishing growth rate.

Pattern shift is usually not a problem for the SiGe CMOS and PMOS devices with an MBE layer. The active area mask is the first mask used after MBE growth and all critical alignments are done with respect to the active area mask. Pattern shift might be a problem for the virtual substrate/limited area batch presented in chapter 8 because the devices are fabricated on tiny pillars surrounded by trenches etched before MBE deposition. The alignment with respect to the trenches is critical. As fig. 5.2 shows, the thick MBE layer leads apart from washout to a narrowing of the trenches. As the size of the pillars gets effectively larger, the alignment is now affected by washout.

In UHVCVD [6], germane GeH_4 is used as a source gas for the growth of strained $Si_{1-x}Ge_x$ epilayers exceeding the critical thickness [7]. High temperature processing would lead to strain relaxation and islanding. However, the growth temperature must be high enough in order to ensure a sufficient surface mobility resulting in good quality epitaxial growth. An additional benefit of UHVCVD is that outdiffusion and outgassing (autodoping) can be suppressed due to growth temperatures of 550^oC or even less.

In UHVCVD, epitaxial growth is based on the high quality substrate, the purity of the source gas and the low growth pressure. The control of the oxygen and water partial pressures is essential to reduce the SiO_2 film. The surface preparation is an even more crucial step in

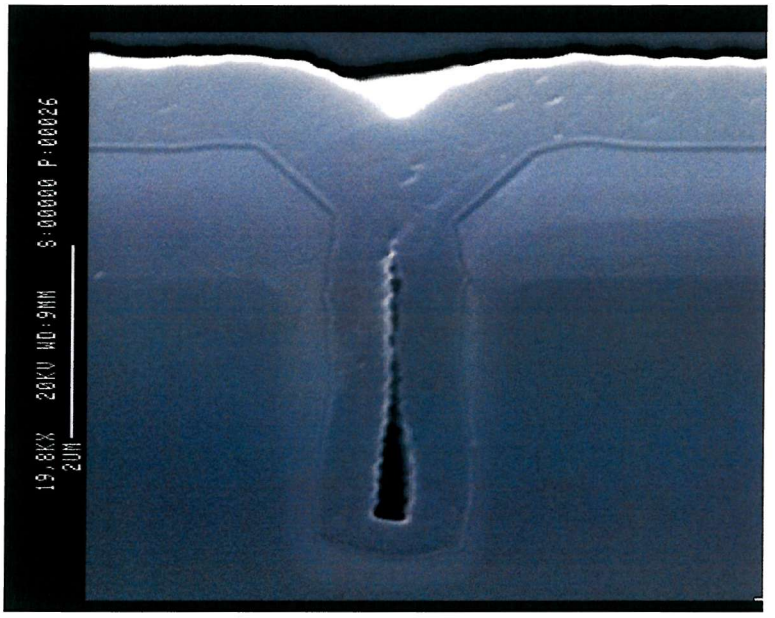


Figure 5.2: Washout of patterns by MBE deposition for the virtual substrate/limited area batch

UHVCVD than in conventional Si epitaxy. The surface cleaning techniques can be classified into ex-situ (surface cleaning outside the growth reactor) and in-situ (surface cleaning inside the growth reactor) techniques.

5.2 Physical Vapor Phase Deposition

Typical physical vapor phase deposition methods are evaporation, sputtering, and molecular beam epitaxy (MBE). Evaporation was used for the early semiconductor technologies. However, multi-component films, step coverage and film uniformity are primary problems of this technique. Sputtering is primarily used for the deposition of metals and alloys. MBE is the ideal method for the fabrication of well-defined SiGe heterostructures.

5.2.1 Molecular Beam Epitaxy

BCF Theory A 2-D growth theory (BCF theory) for the growth of MBE layers was developed by Burton, Cabrera, and Frank (see fig 5.4). In comparison to CVD, the MBE growth theory is quite accurate. In the BCF theory, the surface of the wafer consists of two levels of terraces where the upper terrace is one atomic layer higher than the lower terrace. The terraces can either be due to dislocations, previous growth or imperfect crystal orientation.

The BCF theory is based on the evaporation model. In order to liberate itself from the upper level, an atom at a kink site has just to break a few bonds. Such an atom diffuses across the surface and eventually desorbs. Assuming that the activation energy for liberating an atom from

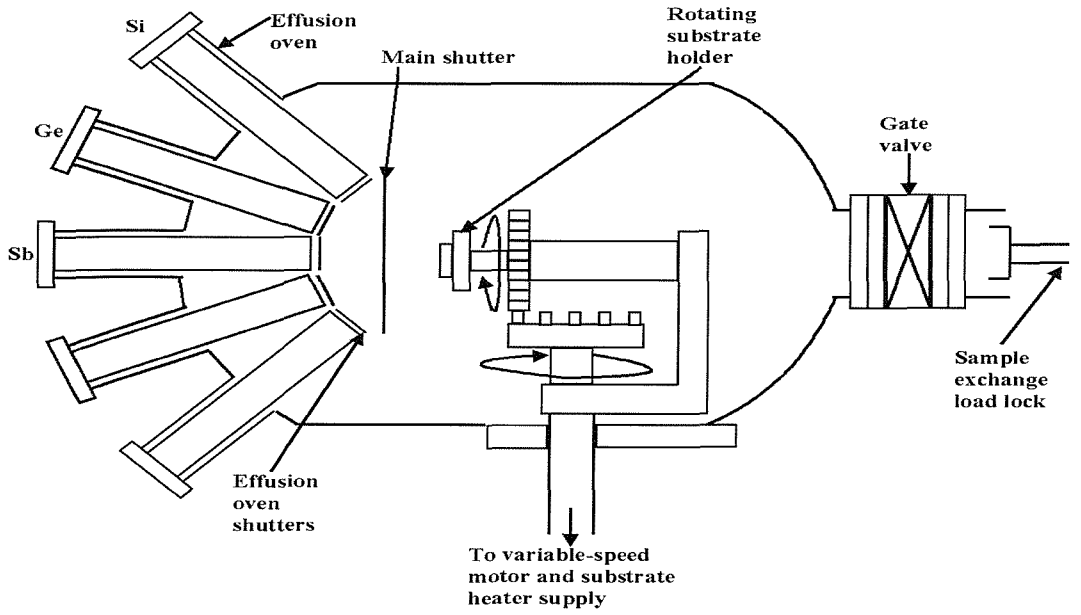


Figure 5.3: Schematic of an MBE system

the kink site is W_s, the equilibrium density of adsorbed atoms is [2, p. 372]:

$$n_{seq} = N_s e^{-W_s/kT} \tag{5.15}$$

N_s is the density of surface positions. Freed from the kink sites, the small energy barrier for the adsorbed atoms to move from site to site is U_s. With a the distance between sites and f the frequency of lattice vibrations, the surface diffusivity D_s is given by [2, p. 372]:

$$D_s = a^2 f e^{-U_s/kT} \tag{5.16}$$

The mean time T for the adsorbed atoms to desorb is given by [2, p. 372]:

$$\frac{1}{T} = f' e^{-E_d/kT} \tag{5.17}$$

Where E_d is the energy barrier for the desorption and f' a pre-exponential factor related to the lattice vibrational frequency f. The desorbing flux F_o can be expressed as [2, p. 373]:

$$F_o = \frac{n_{seq}}{T} = N_s f' e^{-(E_d+W_s)/kT} \tag{5.18}$$

Generally, W_s and E_d are comparable and considerably larger than U_s.

In the BCF theory, adatoms arrive on the surface between the surface steps, diffuse to the edges of the step and along the step edge until they reach a kink site where they are incorporated into the crystal. With j_{inc} the incident flux of adatoms and n_s the density of adsorbed atoms, the net flux j_v of adatoms is given by [2, p. 373]:

$$j_v = j_{inc} - \frac{n_s}{T} \tag{5.19}$$

The surface current density of adatoms is given by [2, p. 373]:

$$j_s = -D_s \nabla n_s \tag{5.20}$$

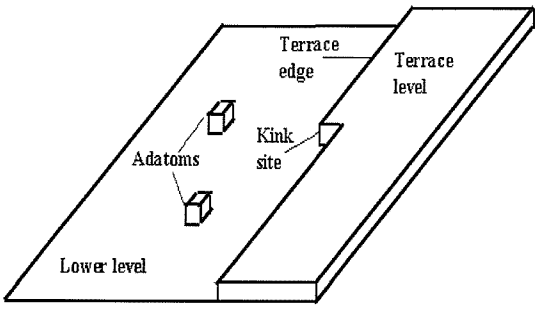


Figure 5.4: MBE Growth in the BCF Theory

For steady state conditions, the divergence of the surface current density equals the net flux of adatoms [2, p. 373]:

$$\nabla j_s + j_v = 0 \tag{5.21}$$

Introducing the supersaturation parameter σ which describes growth ($\sigma > 0$) and evaporation ($\sigma < 0$) [2, p. 374]:

$$\sigma = \frac{j_{inc}}{\frac{n_{seq}}{T}} - 1 \tag{5.22}$$

And the mean migration length before desorption λ_s [2, p. 373]:

$$\lambda_s = \sqrt{D_s T} \tag{5.23}$$

Combining equations (5.20) and (5.21) leads to the following differential equation [2, p. 374]:

$$\lambda_s^2 \nabla^2 n_s = n_{seq}(\sigma + 1) \tag{5.24}$$

BCF solved this differential equation in one dimension with the boundary conditions $n_s(y) = n_{seq}$ and $n(\infty) = (1 + \sigma)n_{seq}$ [2, p. 374]:

$$n_s = n_{seq}(1 + \sigma(1 - e^{-y/\lambda_s})) \tag{5.25}$$

The growth rate can be calculated if one knows the step distribution. However, the BCF theory applies only when the degree of saturation is not too large because 3-D growth dominates the process at high admolecule concentrations.

MBE Growth of Si and SiGe Layers Molecular beam epitaxy (MBE) [1, 2, 15, 16] is a sophisticated method for growing epitaxial films in an ultrahigh vacuum at temperatures between 500^o C and 900^o C. The evaporated species impinge on the heated substrate and condense to form epitaxial layers. MBE has so far received little attention because of its complexity and its low throughput. The low growth temperatures preclude dopant diffusion and the relaxation of strained SiGe layers. Its base vacuum is at least 10^{-10} torr.

Typically, MBE systems include several in situ analysis tools (e.g. electron diffraction and Auger spectroscopy) and use adsorption or turbomolecular pumps in order to achieve the high vacuum (see fig. 5.3). Besides, MBE systems possess a load lock in which a pressure of 10^{-6}

torr is obtained within 15 min.. The wafer is radiantly or resistively heated through the chuck and rotated in order to ensure excellent lateral uniformity across the entire wafer. Shutters in front of the sources are used in order to turn on and off the source elements abruptly.

A bakeout that can take up to a week depending on the level of exposure is performed in order to clean the MBE chamber. Particulation from other surfaces during growth is a major source of defects. Cooling the chambers walls reduces the defect density considerably. A residual metal contamination is a further serious concern in MBE. Incorporated into MBE Si and SiGe films, metal contamination reduces the minority carrier lifetime.

The two most common MBE techniques [17] are the solid source (SSMBE) and gas source (GSMBE) molecular beam epitaxy. Whereas in SSMBE the Si and Ge fluxes are generated by electron evaporation of the solid charges, GSMBE employs the same gases as CVD for epitaxial growth.

SSMBE allows a control of the layer thickness of up to one monolayer and extremely abrupt dopant profiles. The abruptness of dopant profiles and interfaces is limited by segregation [18-25]. The incorporation of dopants and Ge is a complex function of growth environment (temperature, partial pressure, chemicals), Ge and dopant concentration - variables on which also the quality of the SiGe layer depends [24]. Applying a bias voltage to the wafer enhances the dopant incorporation considerably (potential enhanced doping - PED). Growth rates are in the order of 1 to 3 $\AA/s$.

In contrast to SSMBE, GSMBE is selective, but not as controllable. The Ge segregation is suppressed and the planarity of Si/SiGe interfaces is promoted by the presence of hydrogen. The same effect is achieved in surfactant-mediated epitaxy in which growth takes place in the presence of a third atomic species [24-28]. The Ge segregation increases with the Ge concentration, but is independent of temperature and contradicts therefore the two state model that is used in order to describe the surface segregation during epitaxial growth.

Compressively strained SiGe layers with high Ge content [30] ($>$ 30 %) tend to roughen. Performing an anneal after the deposition of the SiGe layer by MBE can lead to the improvement of the crystal quality [20, 24].

In order to fabricate n-MOSFETs with a strained Si channel and p-MOSFETs with a high Ge-content SiGe channel, a virtual substrate has to be grown, i. e. a relaxed SiGe layer. The quality of the virtual substrate determines the quality of the channel which is grown on top of it. A graded Ge profile reduces the defect density [31-37]. Crystal defects in the virtual substrate are the origin of leakage current [37]. The thickness of the virtual substrate layer can be reduced by injection of point defects [38-40] or the local growth on pillars [41, 42]. Furthermore, the local growth on pillars can eliminate cross-hatch [43-48].

For a virtual substrate grown on a standard (100) Si substrate, cross-hatch and misfit dislocations run along the [011] and the [0$\bar{1}$1] directions (see fig. 5.5). Stress relief in virtual substrates on flat surfaces and on large pillars is dominated by the modified Frank-Read (MFR) relaxation

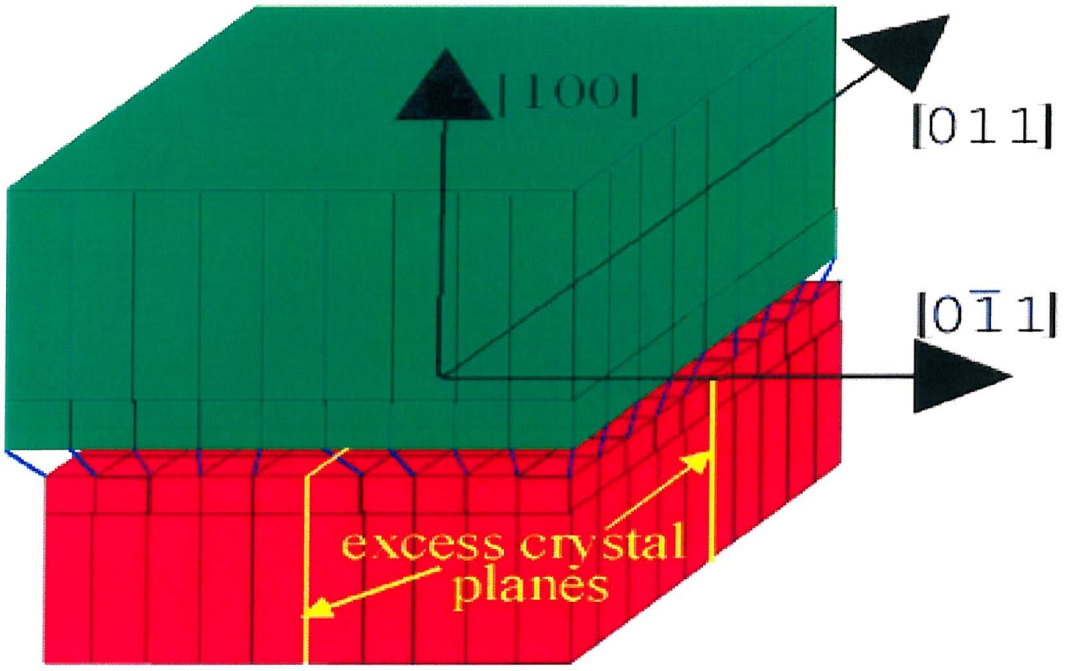


Figure 5.5: Interface of two crystals with different lattice constant. Two excess crystal planes end in dislocation lines along the [011] or [0$\bar{1}$1] directions.

mechanism resulting in a high threading dislocation density, cross-hatch, and residually strained virtual substrates [47]. The two opposite edges of a microscopic pillar induce the formation of misfit dislocations along the [011] and the [0$\bar{1}$1] directions and perpendicular to the edges of the pillar. The closer the distance between two edges along these crystallographic directions is, the more likely it is that a misfit dislocation along these directions is formed [44]. Therefore, the dimensions and orientation have an decisive impact on the relaxation mechanism. Long pillars along the [011] or [0$\bar{1}$1] direction relieve strain asymmetrically, whereas long pillars along the [010] direction can relieve both strain components. Pillars with sufficiently small lateral dimensions ($< 10\ \mu m$) relieve both strain components effectively.

5.2.2 Sputtering

Sputtering is the predominant technique for metal deposition due to the following capabilities [1, p. 379]:

1. high deposition rate
2. deposition of complex alloy compositions
3. deposition of high-temperature and refractory metals
4. deposition of well-controlled, uniform layers on large wafers
5. in situ clean of the wafer surface before deposition

In sputtering, an Ar plasma is generated by a glow discharge. The ionized Ar atoms bombard the cathode (target) and eject some of the target atoms which move to the wafer surface. The ejected atoms are in general neutral. When a magnetic field is applied, the deposition rate increases because the trapped electrons ionize more Ar atoms. A RF-field instead of a dc field prevents the charge accumulation at the target. RF sputtering is therefore used for the deposition of insulating or semiconducting materials.

Ag, Au and Cu are the only three materials with a higher conductivity than Al. Today, Al is still predominant in wiring although wiring with Cu has been introduced recently. Cu can diffuse through SiO_2 and a barrier layer such as Si_3N_4 is needed. More over, copper lacks a dense oxide and is vulnerable to corrosion.

Al films are usually deposited between 200 and 300^oC. Due to the high solubility of Si in Al, Al spikes into the Si substrate leading to source/drain junction leakage. In order to prevent the Si diffusion, Si is often alloyed into Al [2, p. 383]. TiN serves as a barrier or glue metal layer. In reactive sputtering of TiN, nitrogen is introduced into the sputtering chamber. Reactive sputtering is suitable for the deposition of compounds which are difficult to sputter. TiN-Si contacts have a high contact resistance. A Ti layer below the TiN is therefore usually used to ensure good contact resistance [2, p. 388]. CVD is suitable for the deposition of a number of metals and metal compounds. For the deposition of TiN, CVD provides better step coverage than PVD.

5.3 Conclusion

This chapter reviewed the vapor phase deposition methods relevant to conventional processing and this work. CVD methods are used in industry for the deposition of dielectric, polysilicon and epitaxial silicon films. Sputtering is ideal for the deposition of metal and alloy layers.

MBE enables the control of layer thicknesses of up to one monolayer and extremely abrupt dopant profiles. Because of its low throughput, MBE is not used in industry. Besides, MBE is nonselective and requires the use of a deposited field oxide for MOS or CMOS heterochannel applications. The Ge and dopant incorporation are a complex function of the growth environment (temperature, partial pressure, chemicals).

Bibliography

[1] S. A. Campbell, "The Science and Engineering of Microelectronic Fabrication", Oxford University Press, 1996

[2] C. Y. Chang, and S. M. Sze, "ULSI Technology", McGraw-Hill, 1996

[3] M. J. Thwaites and H. S. Reehal, "Growth of Single-crystal Si, Ge and SiGe layers using plasma-assisted CVD", Thin Solid Films, Vol. 294, pp. 76-79, 1997

[4] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors", Semicond. Sci. Technol., Vol. 13, pp. 1225-1246, 1998

[5] E. V. Thomsen, C. Christensen, C. R. Andersen, E. V. Pedersen, P. N. Egginton, O. Hansen, and J. W. Petersen, "Cold-walled UHV/CVD batch reactor for the growth of $Si_{1-x}Ge_x$ layers", Thin Solid Films, Vol. 294, pp. 72-75, 1997

[6] Le Thanh Vinh, V. Aubry-Fortuna, Y. Zheng, D. Bouchier, C. Guedj, and G. Hencelin, "UHV-CVD heteroepitaxial growth of $Si_{1-x}Ge_x$ alloys on Si (100) using silane and germane", Thin Solid Films, Vol. 294, pp. 59-63, 1997

[7] A. V. Potapov, L. K. Orlov, and S. V. Vin, "The growth kinetics of $Si_{1-x}Ge_x$ layers from SiH_4 and GeH_4", Thin Solid Films, Vol. 336, pp. 191-195, 1998

[8] K. L. Wang, S. G. Thomas, and M. O. Tanner, "SiGe band engineering for MOS, CMOS, and quantum effect devices", J. Mat. Sci.: Mat. Electron., Vol. 6, pp. 311-324, 1995

[9] M. Iwaya, T. Takeuchi, S. Yamaguchi, C. Wetzel, H. Amano, and I. Akasaki, "Reduction of Etch Pit Density in Organometallic Vapor Phase Epitaxy-Grown GaN on Sapphire by Insertion of a Low-Temperature-Deposited Buffer Layer between High-Temperature-Grown GaN", Jpn. J. Appl. Phys., Vol. 37, pp. L 316-L 318, 1998

[10] C.-T. Lin, S.-J. Chang, D. K. Nayak, and Y. Shiraki, "Deposition of SiO_2 Films on Strained SiGe Layer by Direct Photo Chemical Vapor Deposition", Jpn. J. Appl. Phys., Part 1, Vol. 34, No. 1, pp. 72-74, 1995

[11] C. T. Lin, S. J. Chang, D. K. Nayak, and Y. Shiraki, " The properties of SiO_2 films using direct photo-chemical vapor deposition on strained SiGe layers", Appl. Surf. Sci., Vol. 92, pp. 193-197, 1996

[12] E. Dufour-Gergam, F. Meyer, F. Delmotte, M. C. Hugon, B. Agius, P. Warren, and D. Dutartre, "Electrical properties of silicon nitride films grown on a SiGe layer by distributed electron cyclotron resonance plasma-enhanced chemical vapor deposition", Thin Solid Films, Vol. 294, pp. 214-216, 1997

[13] B. Gallas, I. Berbezier, and J. Derrien, "Homoepitaxy of silicon at low temperature on clean and Ga-covered substrates", This Solid Films, Vol. 294, pp. 69-71, 1997

[14] K. Noda, T. Tatsumi, T. Uchida, K. Nakajima, H. Miyamoto, and C. Hu, "A $0.1 - \mu m$ Delta-Doped MOSFET Fabricated with Post-Low-Energy Implanting Selective Epitaxy", IEEE Trans. Electron Dev., Vol. 45, No. 4, pp. 809-814 , 1998

[15] T. G. Yugova, V. I. Vdovin, M. G. Mil'vidskii, L. K. Orlov, V. A. Tolomasov, A. V. Potapov, and N. V. Abrosimov, "Dislocation pattern formation in epitaxial structures based on SiGe alloys", Thin Solid Films, Vol. 336, pp. 112-115, 1998

[16] C. Schoellhorn, M. Oehme, M. Bauer, and E. Kasper, "Coalescence of germanium islands on silicon", Thin Solid Films, Vol. 336, pp. 109-111, 1998

[17] T. E. Whall and E. H. C. Parker, "Review Silicon-germanium heterostructures- advanced materials and devices for silicon technology", Journal of Materials Science: Materials in Electronics, Vol. 6, pp. 249-264, 1995

[18] K. Kimura, Y. Endoh, M. Mannami, H.-J. Gossmann, G. H. Gilmer, and L. C. Feldman, "Anomalous surface segregation of Sb in Si during epitaxial growth", Appl. Phys. Lett., Vol. 69, No. 1, pp. 67-70, 1996

[19] H. Jorke, "SURFACE SEGREGATION OF Sb ON Si(100) DURING MOLECULAR BEAM EPITAXY GROWTH", Surface Science, Vol. 193, pp. 569-578, 1988

[20] D. Streit, R. A. Metzger, and F. G. Allen, "Doping of silicon in molecular beam epitaxy systems by solid phase epitaxy", Appl. Phys. Lett., Vol. 44, No. 2, pp. 234-236, 1984

[21] K. Kimura, Y. Endoh, M. Mannami, H.-J. Gossmann, G. H. Gilmer, and L. C. Feldman, "Anomalous surface segregation of Sb in Si during epitaxial growth", Appl. Phys. Lett., Vol. 69, No. 1, pp. 67-70, 1996

[22] S. Fukatsu, S. Kubo, Y. Shiraki, and R. Ito, "Redistribution of Sb in an atomic-layer-doped Si", Appl. Phys. Lett., Vol. 58, No. 11, pp. 1152-1154, 1991

[23] T. E. Whall and E. H. C. Parker, "Growth and characterisation of pseudomorphic Si/SiGe/Si heterostructures for p-channel field-effect transistors", Thin Solid Films, Vol. 306, pp. 338-345, 1997

[24] S.-G. Kim, J. Kim, I.-H. Bae, J. G. Koo, and K.-S. Nam, "Analysis of $Si_{1-x}Ge_x$ on Si heterostructures grown by MBE as a function of substrate temperature", Journal of Crystal Growth, Vol. 186, pp. 375-381, 1998

[25] T. E. Whall and E. H. C. Parker, "SiGe heterostructures for FET applications", J. Phys. D: Appl. Phys., Vol. 31, pp. 1397-1416, 1998

[26] M. Kammler, D. Reinking, K. R. Hoffmann, and M. Horn-von Hoegen, "Surfactant-mediated epitaxy of Ge and Si: progress in growth and electrical characterization", Thin Solid Films, Vol. 336, pp. 29-33, 1998

[27] B. Gallas, I. Berbezier, A. Ronda, and J. Derrien ,"Growth mechanisms of SiGe on (111) and (100) Si substrates", Thin Solid Films, Vol. 294, pp. 22-26, 1997

[28] T. E. Whall and E. H. C. Parker, "Review Silicon-germanium heterostructures- advanced materials and devices for silicon technology", Journal of Materials Science: Materials in Electronics, Vol. 6, pp. 249-264, 1995

[29] G. G. Jernigan, P. E. Thompson, and C. L. Silvestre, "Quantitative measurements of Ge surface segregation during SiGe alloy growth", Surface Science, Vol. 380, pp. 417-426, 1997

[30] S.-G. Kim, J. Kim, I.-H. Bae, J. G. Koo, and K.-S. Nam, "Analysis of $Si_{1-x}Ge_x$ on Si heterostructures grown by MBE as a function of substrate temperature", Journal of Crystal Growth, Vol. 186, pp. 375-381, 1998

[31] A. J. Pidduck, D. J. Robbinson, D. Wallis, G. M. Williams, A. C. Churchill, J. P. Newey, C. Crumpton, and P. W. Smith, "On the growth of high quality relaxed $Si_{1-x}Ge_x$ layers on Si by vapour phase epitaxy", Inst. Phys. Conf. Ser., No. 157, pp. 135-144, 1997

[32] M. Kummer, B. Voegli, and H. Von Kaenel, "STM study of step graded $Si_{1-x}Ge_x/Si(100)$ buffers", Thin Solid Films, Vol. 336, pp. 100-103, 1998

[33] E. Muller, R. Hartmann, C. David, and D. Grutzmacher, "Structural characterization of SiGe step graded buffer layers grown on prestructured Si [100] substrates by molecular beam epitaxy", Thin Solid Films, Vol. 336, pp. 92-95, 1998

[34] S. Martelli, I. Vianey, R. Larciprete, E. Borsella, J. Castro. S. Chiussi, and B. Leon, "X-ray diffraction and x-ray photoelectron spectroscopy study of partially strained SiGe layers produced via excimer laser processing", J. Appl. Phys., Vol. 82, No. 1, pp. 147-153, 1997

[35] E. A. Fitzgerald and S. B. Samavedam, "Line, point, surface defect morphology of graded, relaxed GeSi alloys on Si substrates", Thin Solid Films Vol. 294, pp. 3-10, 1997

[36] M. T. Curie, S. B. Samavedam, T. A. Langdo, C. W. Leitz, and E. A. Fitzgerald, "Controlling threading dislocation densities in Ge on Si using graded SiGe layers and chemical-mechanical polishing", Appl. Phys. Lett., Vol. 72, No. 14, pp. 1718-1720, 1998

[37] S. Kanjanachuchai, T. J. Thornton, J. M. Fernandez, and H. Achmed, "Leakage currents in $Si_{1-x}Ge_x$ virtual substrates: measurements and device implications", Semicond. Sci. Technol., Vol. 13, pp. 1215-1218, 1998

[38] K. Lyutovich, E. Kasper, F. Ernst, M. Bauer, and M. Oehme, "Relaxed SiGe buffer layer growth with point defect injection", Materials Science and Engineering, Vol. B71, pp. 14-19, 2000

[39] E. Kasper, K. Lyutovich, M. Bauer, and Michael Oehme, "New virtual substrate concept for the vertical MOS transistor", Thin Solid Films, pp. 319-322, 1998

[40] M. Bauer, M. Oehme, K. Lyutovich, and E. Kasper, "Ion assisted MBE growth of SiGe nanostructures", Thin Solid Films, Vol. 336, pp. 104-108, 1998

[41] T. Rupp, J. Messarosch, and I. Eisele, "Silicon-germanium molecular beam epitaxy system for high-quality nanostructures and devices", Journal of Crystal Growth, Vol. 183, pp. 99-108, 1998

[42] T. Rupp, F. Kaesen, W. Hansch, E. Hemmerl, D. J. Gravesteijn. R. Schorer, E. Silveira, G. Abstreiter, and I. Eisele, "Defect-free strain relaxation in locally MBE-grown SiGe heterostructures", Thin Solid Films, Vol. 294, pp. 27-32, 1997

[43] G. Wohl, V. Dudek, M. Graf, H. Kibbel, H. J. Herzog, and M. Klose, "Relaxed $Si_{0.7}Ge_{0.3}$ buffer layers grown on patterned silicon substrates for SiGe n-channel HMOSFETs", Thin Solid Films, Vol. 369, pp. 175-181, 2000

[44] E. Muller, R. Hartmann, C. David, and D. Grutzmacher, "Structural characterization of SiGe step graded buffer layers grown on prestructured Si[001] substrates by molecular beam epitaxy", Thin Solid Films, Vol. 336, pp. 92-95, 1998

[45] E. A. Fitzgerald, P. D. Kirchner, R. Proano, G. D. Pettit, J. M. Woodall, and D. G. Ast, "Elimination of interface defects in mismatched epilayers by a reduction in growth area", Appl. Phys. Lett., Vol. 52, No. 18, pp. 1496-1498, 1988

[46] A. R. Powell, R. A. Kubiak, T. E. Whall, E. H. C. Parker, and D. K. Bowen, "Residual Strain and Defect Analysis in As Grown and Annealed SiGe Layers", Mat. Res. Soc. Symp. Proc., Vol. 220, pp. 277-283, 1991

[47] R. Hammond, P. J. Phillips, T. E. Whall, E. H. C. Parker, T. Graf, H. von Kanel, and A. J. Shields, "The elimination of surface cross-hatch from relaxed, limited-area $Si_{1-x}Ge_x$ buffer layers", Appl. Phys. Lett., Vol. 71, No. 17, pp. 2517-2519, 1997

[48] T. Rupp, F. Kaesen, W. Hansch, E. Hemmerl, D. J. Gravesteijn. R. Schorer, E. Silveira, G. Abstreiter, and I. Eisele, "Defect-free strain relaxation in locally MBE-grown SiGe heterostructures", Thin Solid Films, Vol. 294, pp. 27-32, 1997

Chapter 6

Processing

HMOS is compatible with the conventional Si technology and based on the same processes. However, the presence of a strained SiGe layer often imposes additional constraints.

In this chapter, further fabrication processes which are not as fundamental as vapor-phase deposition for HMOS, but as crucial are reviewed. Their effect on a strained SiGe channel is examined. In a very limited number of cases, Ge can be used in a beneficial way. Implantation is an alternative method to fabricate strained SiGe layers. In silicidation, the presence of Ge leads to a better dopant activation and reduces the contact resistance.

6.1 Lithography

Lithography [1, chap. 7-9; 2, chap. 6] is the key technology is semiconductor industry because it determines the device dimensions and affects the device quality. The lithographic methods can be classified into direct write techniques and techniques applying a master mask.

In master mask techniques, a resist that is sensitive to radiation is coated on the wafer surface. The circuit patterns are projected on the resist by different sorts of radiation (optical, electrons, x-rays and ions). After the projection of the circuit patterns on the resist, the resist is developed.

In direct writing, patterns are directly printed on the resist-coated substrate. As direct writing does not suffer from diffraction effects, it is more accurate than master mask techniques. Direct writing can be applied for low-volume production. Whereas X-rays can be solely used for master mask techniques, optical, ion and electron radiation can be used either for direct writing or master mask techniques.

6.1.1 Lithography Techniques

Optical Lithography In optical lithography, the image is formed in a photoresist with visible or ultraviolet radiation. Optical lithography using a photomask is the most common lithography technique as it has so far fulfilled the industrial requirements and allows a high throughput. The

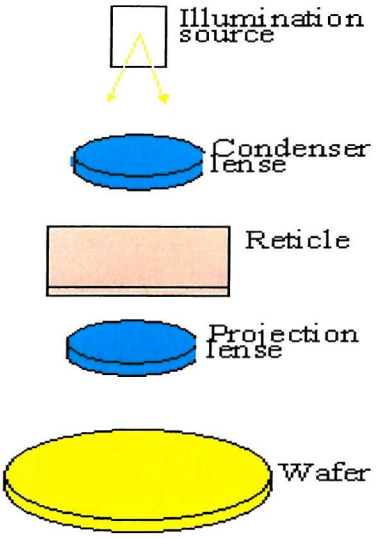


Figure 6.1: Schematic stepper system

master masks are either made be electron beam exposure or laser beam scanning. They consist of a thin chromium layer (0.1 μm) in which the circuit patterns are etched.

Contact and proximity printing do not have any means of image formation between the master mask and photoresist. In contact printing, the master mask is in direct contact with the photoresist, whereas there is a space between $20 - 50$ μm in proximity printing. In reality, the master mask is also not everywhere in direct contact with the photoresist due to the nonuniformity of the wafer surface and mask surface. The direct contact printing produces defects in both mask and wafer. Hence, the masks have a short lifetime.

Projection printing has a higher resolution than proximity printing and uses an optical system (e.g. catadioptric and refraction) between the master mask and the photoresist. The resolution is determined by the optical system and the wave length. The common types of projection printers are scanners and steppers (see fig. 6.1). Scanners have a higher throughput than steppers. The achieved maximum resolution is 0.2 μm for optical systems.

Electron Lithography In electron lithography, a tungsten tip with a radius between $0.5 - 1$ μm heated to about 1900^oC emits current with high stability and density. The emitted electrons are accelerated by a high voltage (10-50 kV). The electron beam is focused with magnetic lenses. The resolution for e-beam lithography is higher than 0.1 μm. It is not limited by wavelength but by electron scattering in the Si and resist, various aberration of electron optics (spherical, chromatic, astigmatism, deflection) and Coulomb repulsion of the electrons. E-beam lithography is used for mask fabrication and direct writing on the wafer surface.

The high energies used in nonoptical lithography can damage oxide, oxide/semiconductor interfaces and SiGe heterostructures [3]. High energy irradiation leads to an increase of the fixed oxide charge density, the neutral trap density, and the interface state density of MOS devices.

The damage is most pronounced in thick oxides and very thin wet oxides are quite radiation hard because the number of ionizing events increases with the oxide thickness. The integrity of the oxide and SiGe heterostructure can be partly restored in an anneal.

X-ray Lithography X-ray lithography was proposed in 1972. In deep UV, all optical materials become opaque. However, transmission increases again in the x-ray region. Compared to optical lithography, the advantages of x-ray lithography are the depth of focus and the fact that light organic contaminants do not print as defects. In general, x-ray lithography uses proximity printing. A masks that is illuminated by x-rays casts a shadow on a resist-covered wafer.

X-ray masks consist of a transmittive membrane on which an absorbing metal layer is patterned. Tungsten and tantalum are used as absorber materials because they can easily be dry etched. The layer is considerably thicker than the chromium layers on photomasks because all materials allow the transmission of x-rays to a certain degree. The projected features have the same size as on the master mask. Hence, the resolution is lower than for e-beam lithography that is used for the mask fabrication. So far, the best achievable resolution has been 0.16 μm.

X-rays are produced by the interaction of electrons on a target material or plasma discharge. Most of the input power is converted to heat. Plasma discharge leads to a higher x-ray flux than the bombardment of a target, but is still very ineffective. The maximum x-ray energy is the energy of the incident electrons.

Ion Beam Lithography Ions are produced by the electrical field in a gas or liquefied metal surrounding a tungsten tip and accelerated by an electrical field. Electrostatic lenses are used to focus the ion beam. These systems have higher aberrations than magnetic systems. However, the resolution of an ion beam is higher than for an electron beam because of less scattering. The best resolution that has ever been attained is 400 $\mathring{A}$.

6.1.2 Photoresist

Photolithography is based on the different solubility between irradiated and unirradiated areas and can be categorized by its polarity. Its most important features are the contrast between exposed and unexposed areas of photoresist, the adhesion to the substrate and the resistance to wet and dry etch. The regions of positive resist which were exposed to irradiation are dissolved more quickly during the development process than the unexposed regions, whereas negative resist has the opposite response to irradiation and the development process. Positive resist is more common than negative resist because of its better resolution. The standard positive photoresists have three components (a resin or base material, a photoactive compound PAC, a solvent) and contains carbon compounds (aromatic rings and polymers). The solvent is responsible for the mechanical properties of the resist. When the photoactive compound that is a dissolution inhibitor is destroyed by the exposure to light, the resin becomes more soluble in an aqueous

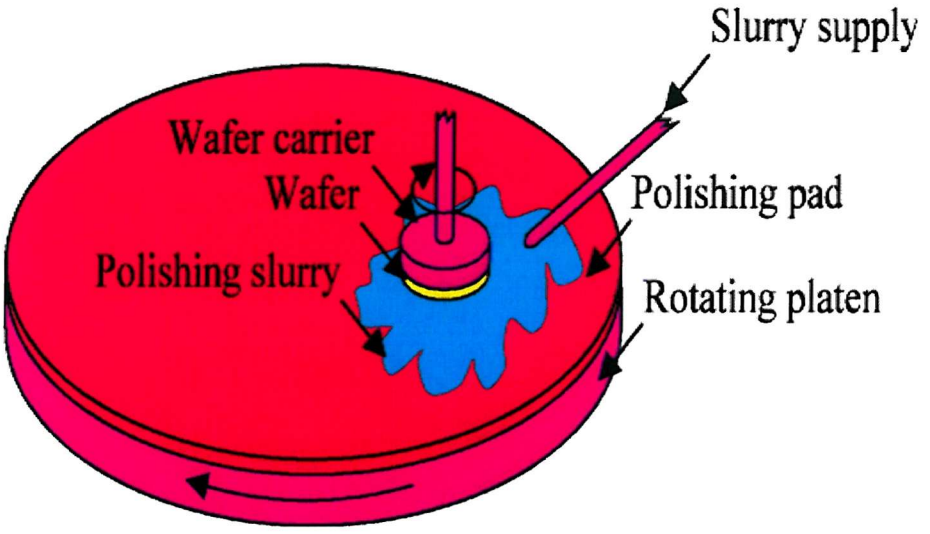


Figure 6.2: Schematic of a CMP polisher

developer solution whereas the unexposed regions do not swell much. In consequence, high resolution is possible with this resist.

The wafers have to receive a pretreatment before they are coated with resist. First of all, dehydration bake at $150 - 200^oC$ is performed. Immediately before the wafer is coated with photoresist, the wafer is typically primed with hexamethyldisilazane to improve the adhesion. Spin coating is the most common method. The wafer is spun at 2000-6000 rpm. After the spin, less than 1 % of the dispensed resist is typically left on the wafer. The resist thickness depends mainly on the spin speed and the viscosity of the resist.

Before exposure, the resist has to undergo a prebake in order to drive off the solvents as the dissolution rate in the developer is strongly dependent on the solvent concentration. The prebake takes typically between 0.5 to 30 min. at $90 - 100^oC$. After exposure, the resist is developed in alkaline developers dissolved in water. The development process is very temperature sensitive and the developer which is consumed during the development cycle has to be replenished. A hardbake is performed after the development in order to harden the resist against further energetic processes (e.g. implantation and plasma etching).

6.2 Chemical/Mechanical Polishing

The planarity of interlevel dielectrics is crucial for the realization of complex circuits requiring several levels of metal wiring. Former approaches were the introduction of dummy oxide/metal features or an oxide etching process. Today, the standard method to achieve planarity is chemical/mechanical polishing (CMP, see fig. 6.2).
In CMP [4], a polishing platen and a wafer holder rotate with different or identical frequency.

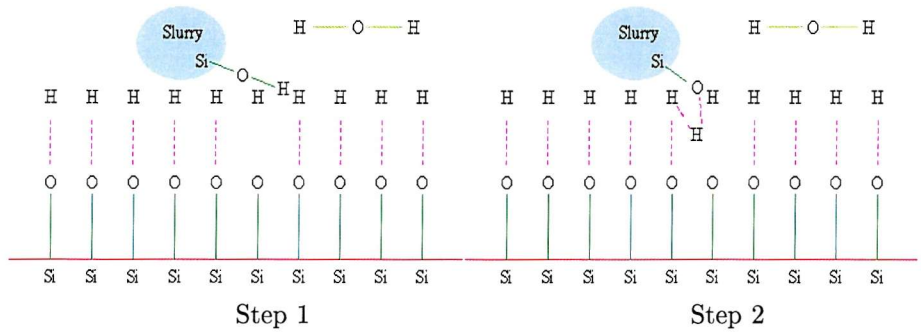


Figure 6.3: Step 1: Formation of hydroxyls in aqueous solution; Step 2: Formation of hydrogen bond between the slurry particle and wafer

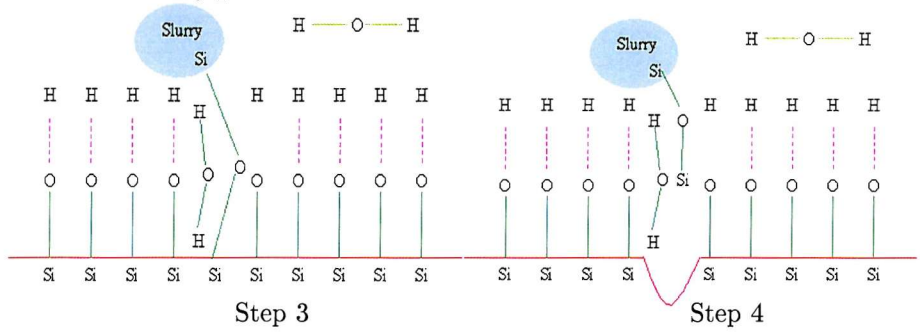


Figure 6.4: Step 3: Release of water molecule and formation of $Si - O$ bond; Step 4: Breaking of $Si - Si$ bond

The Preston equation gives the mechanical removal rate R of glass [2, p. 434]:

$$R = K_p p v \tag{6.1}$$

Where p, v, K_p are respectively pressure, relative velocity and the Preston coefficient. K_p depends on the glass, the polishing pad and the slurry. Planarization occurs because high spots are subject to high pressure and removed preferentially. In order to achieve uniform polishing, the relative velocity between wafer and platen surface has to compensate the effect of the nonuniform polishing slurry that is transported from the edge to the center of the wafer by the surface of the pad. Hence, the wafer should rotate at a higher frequency than the platen.

In SiO_2 polishing, the slurry consists of silica suspended in a KOH solution. The pH determines the formation of hydrogen bonds. Typically, the slurry has a particle size between 10 and 90 nm. Microscopically, polishing is rather chemical in nature than mechanical abrasion. The chemical mechanism can be divided into the following steps [2, p. 436] (see fig.'s 6.3 and 6.4):

1. formation of hydrogen bonds with the oxidized surface of wafer and slurry particles
2. formation of hydrogen bonds between the wafer and slurry particles
3. formation of molecular bonds between wafer and slurry particles
4. breaking of the oxide bonds with the wafer surface/slurry particles surface

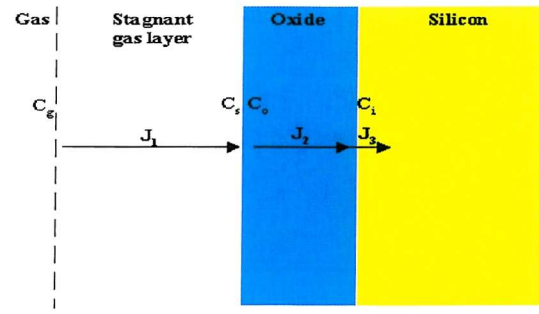


Figure 6.5: Schematic diagram of thermal oxidation

In metal CMP [5], the slurry must contain three different components: fine slurry particles, a corrosion agent and an oxidant. For example, W CMP is extensively used. Its slurry consists of fine alumina powder suspended in a solution of potassium ferricyanide ($K_3Fe(CN)_6$) and potassium dihydrogen phosphate (KH_2PO_4). The ferricyanide serves as both etchant and oxidant, whereas the potassium dihydrogen determines the pH. The etching reaction can be written as [2, p. 439]:

$$W + 6[Fe(CN)_6]^{3-} + 4H_2O \rightarrow (WO_4)^{2-} + 6[Fe(CN)_6]^{4-} + 8H^+ \tag{6.2}$$

The passivation reaction is given by:

$$W + 6[Fe(CN)_6]^{3-} + 3H_2O \rightarrow WO_3 + 6[Fe(CN)_6]^{4-} + 6H^+ \tag{6.3}$$

Effective planarization requires the simultaneous balance of etching, passivation and oxide removal. A pH close to to neutral is favorable for planarization. Etylenediamine can be added in order to adjust the pH close to neutral. Today, potassium ferricyanide that is very toxic is replaced by ferric nitrate $Fe(NO_3)_3$.

6.3 Thermal Oxidation of Si and SiGe

The thermal oxide is widely used as a gate dielectric for MOSFETs and as a field oxide. The excellent properties of thermally grown SiO_2 contributed very much to the popularity of silicon ICs. Thermal oxidation is described by the Deal-Grove model [1, chap. 4] that allows the prediction of oxides larger than 300 $\AA$. The oxide growth is determined by the diffusion of oxygen through the SiO_2 to the Si interface because the diffusivity of oxygen in SiO_2 is several orders of magnitude higher than the diffusivity of Si. As the Si/SiO_2-interface has not come into contact with the ambient it is very clean. The diffusivity increases largely with temperature. The amount of Si consumed during thermal oxidation is roughly 44% of the final oxide thickness.

In dry oxidation, molecular oxygen O_2 reacts with Si according to:

$$Si(s) + O_2(g) \rightarrow SiO_2 \tag{6.4}$$

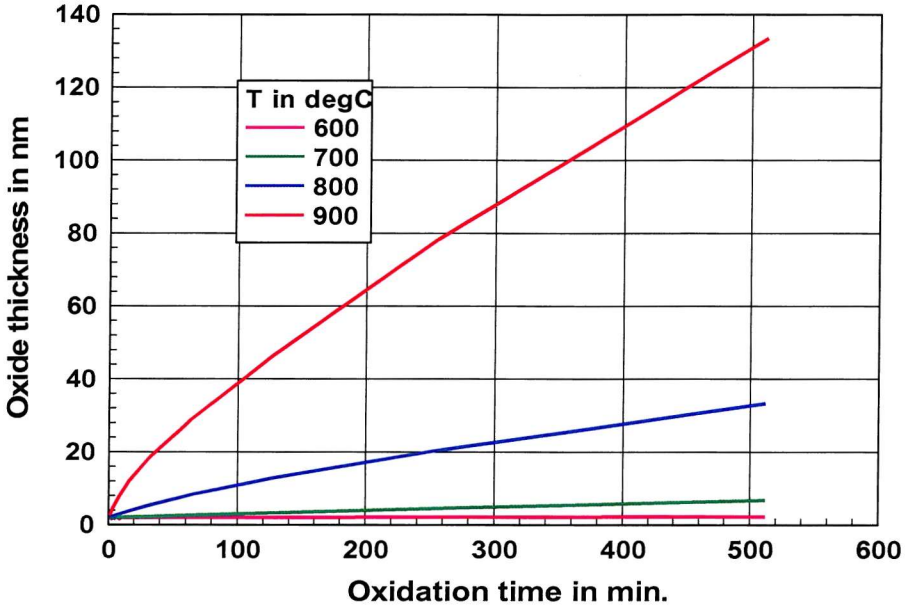


Figure 6.6: Thermal Oxide Growth as simulated with TSUPREM-4 in steps of 100^oC

In the Deal-Grove model, oxygen diffuses from the gas stream through a stagnant gas layer and the silicon oxide to the Si interface (see fig 6.5). The oxygen flow J_1 through the stagnant gas layer is described by [1, p. 69]:

$$J_1 = h_g(C_g - C_s) \tag{6.5}$$

h_g, C_g and C_s are respectively the mass transport coefficient, the oxygen concentration in the gas stream and the oxygen concentration in the gas at the surface of the wafer.

The flow J_2 through the silicon oxide layer can be described by Fick's first law [1, p. 69]:

$$J_2 = D_{O_2}\frac{C_o - C_i}{t_{ox}} \tag{6.6}$$

D_{O_2}, t_{ox}, C_o and C_i are respectively the diffusion coefficient of oxygen in SiO_2, the the thickness of the oxide, the oxygen concentration in the oxide at the wafer surface, and the oxygen concentration at the Si/SiO_2 interface.

The flow of oxygen J_3 reacting with the Si to form SiO_2 is determined by chemical reaction kinetics and proportional to the oxygen concentration [1, p. 70]:

$$J_3 = k_s C_i \tag{6.7}$$

Where k_s is the chemical rate constant for the overall reaction.

In equilibrium, all three flows balance [1, p. 70]:

$$J = J_1 = J_2 = J_3 \tag{6.8}$$

Using equations 6.5 - 6.8 and Henry's law [1, p. 70]:

$$C_o = HP_g = HkTC_g \tag{6.9}$$

The interface concentration of oxygen is found [1, p. 70]:

$$C_i = \frac{HP_g}{1 + \frac{k_s}{h_g} + \frac{k_s t_{ox}}{D_{O_2}}} \tag{6.10}$$

Where H is Henry's gas constant, P_g the partial pressure of oxygen in the gas flow, k the Boltzmann constant, and T the temperature.

In order to obtain the growth rate R, the flux J has to be divided by N, the number of oxygen molecules per unit volume of SiO_2 ($N = 2.2 \cdot 10^{22} cm^{-3}$) [1, p. 70]:

$$R = \frac{J}{N} = \frac{t_{ox}}{dt} = \frac{HK_sP_g}{N(1 + \frac{k_s}{h_g} + \frac{k_s t_{ox}}{D_{O_2}})} \tag{6.11}$$

Assuming that the oxide thickness is t_o at time $t = 0$, the solution of equation 6.11 can be written as [1, p. 70]:

$$t_{ox}^2 + At_{ox} = B(t + \tau) \tag{6.12}$$

Where:

$$A = 2D(\frac{1}{k_s} + \frac{1}{h_g}) \tag{6.13}$$

$$B = \frac{2DHP_g}{N} \tag{6.14}$$

$$\tau = \frac{t_o^2 + At_o}{B} \tag{6.15}$$

Si is usually oxidized at atmospheric pressure when $k_s << h_g$ and the growth rate is nearly independent of the gas phase mass transport and reactor geometry. A and B are well know for the general process conditions. The two limiting forms of equation 6.12 describe the growth of sufficiently thin oxides:

$$t_{ox} = \frac{B}{A}(t + \tau) \tag{6.16}$$

And sufficiently thick oxides:

$$t_{ox}^2 = B(t + \tau) \tag{6.17}$$

B/A is called the linear rate coefficient whereas B is called the parabolic rate coefficient.

The growth of ultrathin oxides that are used as gate oxides for the MOSFETs is not yet fully understood. The oxidation rate of ultrathin oxides follows the following function [1, p. 76]:

$$\frac{dt_{ox}}{dt} = \frac{B}{2t_{ox} + A} + C_1 e^{-t_{ox}/L_1} + C_2 e^{-t_{ox}/L_2} \tag{6.18}$$

Where C_1, C_2, L_1 and L_2 are all constants.

Wet oxidation employs H_2O and oxidizes at a much higher rate than dry oxidation because the H_2O-molecule is considerably smaller than the O_2-molecule. Unfortunately, wet oxides are less dense and do not support as much electrical stress as dry oxides. Often, the growth chamber also contains a small amount of HCl gas (1 to 3 %) that reacts with most heavy metal atoms and is used to clean the gas ambient of these impurities. Moreover, HCl increases the oxidation rate and can improve the Si/SiO_2-interface properties. However, HCl should be avoided for thin

gate oxides as it can roughen their Si/SiO_2-interface. An $HCl-$purge to clean up the furnace before oxidation is advisable before the growth of thin gate oxides.

A heavy substrate dopant concentration tends to enhance the oxidation rate by increasing the diffusivity in SiO_2. Fluorine implanted into the gate can lead to an increase [6] and degradation [7] of the gate oxide.

The oxidation of SiGe can be described by a formalism based on Wagner's work on the oxidation of metallic alloys. Ge acts as a catalyst in wet oxidation whereas no enhanced oxide growth is reported for dry oxidation [8-11]. However, capacitance measurements indicate an increased gate oxide growth for the SiGe MOSFETs as discussed in chap. 8.

Germanium and silicon are oxidized simultaneously and the growth of pure SiO_2 prevails as long as sufficiently Si is supplied towards the reaction interface. The GeO_2 is subsequently reduced by free silicon [8-11]:

$$Si + GeO_2 \rightarrow SiO_2 + Ge \tag{6.19}$$

The liberated Ge is rejected by the growing SiO_2, piles up at the interface and is finally incorporated into the oxide when a critical Ge concentration is surpassed. The critical Ge concentration is temperature dependent. The Ge pile up results in a high oxide charge, a high interface state density and poor break down characteristics [12]. In general, the interface between the Si cap and SiGe channel is not abrupt and Ge is piles up when the Si cap is thermally oxidized [13]. The interface can be improved in an anneal in H_2O [14] or N_2 [15] ambient. As for thermal oxidation, Ge segregates at the $SiO_2/Si_{1-x}Ge_x-$interface of a buried oxide formed by oxygen implantation (SIMOX, see chap. 5.7, [16, 17]).

At low growth temperature [18-24], the diffusion of Ge is prevented which is therefore incorporated in the oxide instead of piling up at the growth interface. The use of anodic oxide [25-28] as an alternative to thermal oxide was investigated at the University of Southampton [29]. The major draw back of anodic oxide is that it is very leaky when grown in small active area windows as required for MOSFETs than over the whole wafer.

Thermal oxidation is supposed to reduce the mobility in the SiGe channel due to strain relaxation and diffusion of dopants and Ge even when no SiGe is oxidized [30]. Alternative dielectrics such as high dielectric constant (high-K) materials (e. g. ZrO_2, HFO_2, Al_2O_3, Ta_2O_5) are under investigation as prospective candidates in conventional CMOS. Deposited oxides [31, 32] or high-K dielectrics [33, 34] might be beneficial to HMOSFETs.

Silicon nitride is seen as a natural alternative to SiO_2 because deposited nitrides are already in use. The dielectric constant of silicon nitride is roughly twice that of SiO_2. Furthermore, nitride layers are excellent diffusion barriers. But, silicon nitride films have three major set backs:

- Silicon nitride has a high interface state density because it is formed by the diffusion of Si to the top surface and not at the Si/SiN_4-interface.

- Thermal nitridization requires very elevated temperatures and leads to considerable strain because the thermal expansion coefficient for silicon nitride is about twice that of silicon.
- Thick nitrides films (> 40 $\AA$) cannot be grown due to the low diffusivity of Si.

Oxynitride SiO_xN_y can be formed by exposing a thermal oxide to a high temperature NH_3 anneal. The exact composition depends on the process conditions. Nitrided films are more resistant to oxide breakdown, radiation damage and have a higher dielectric constant than SiO_2 [35, 36]. Nitrided oxides display a high concentration of electron traps that can be reduce by the reoxidation of the nitrided film.

6.4 Rapid Thermal Processes

In conventional thermal processes, a whole batch of wafers is processed at the same time. The batch is always in thermal equilibrium with the environment during processing. The processes typically take place in a clean, high-purity fused quartz tube surrounded by heating elements.

Rapid thermal processing methods (RTP, [1, chap. 6; 2, chap. 4; 37]) which originate from single wafer processing were originally designed for tight control of contamination and process parameters and try to compensate the loss of time due to single wafer processing [1, chap. 6; 2, chap. 4]. The wafer is either heated by a transient lamp or has to be loaded very quickly in a heated process chamber. The first case where the wafer is not in thermal equilibrium with the surroundings is more common than the second case where the wafer is in thermal equilibrium with the surroundings. The temperature and its uniformity are in general difficult to control and RTP is difficult to model [38].

RTP requires a lower thermal budget than conventional thermal processes because the wafer is heated up rapidly. Due to the high temperature ramp rate, RTP is seldom an isothermal process and thermal stress can produce slip lines. After implantation, annealing leads to the creation of excess point defects and in consequence transient enhanced diffusion (TED). Although the rapid temperature ramp of RTP can reduce the number of point defects quickly, the diffusion during the short RTP is often determined by TED. In comparison with conventional furnace annealing, TED is suppressed.

The merger of single-wafer RTP and single-wafer CVD has lead to continuous improvements to the deposition of thin Si films. In RTCVD Si epitaxy, the incorporation of dopants is mainly controlled by ramping the temperature instead of regulating the gas flow. This method is called limited reaction processing (LRP, [2, chap. 4.3]).

High quality silicides and oxides can be grown by RTCVD [2, chap. 4.3]. Deposited oxides such as TEOS (tetraethoxysilane) and TMCTS (tetramethylcyclotrasiloxane) display a good step coverage. In rapid thermal oxidation (RTO), oxidation is accelerated by UV radiation. The oxide quality is comparable or better than furnace oxide. For the deposition of other films, RTP has

met only limited success. Sometimes, the advantages of RTP cannot compensate the loss of time due to single wafer processing.

The role of RTP for SiGe technology is not yet investigated and far from obvious. On the one hand, the relatively low thermal budget of RTP might prevent strain relaxation. On the other hand, the high thermal stress during the ramp of the temperature might cause strain relaxation.

6.5 Etching

The patterns in the photoresist on the wafer surface can be transferred into the layer underneath the resist by etching [1, 2]. The selectivity, rate, and anisotropy are important figures of merit of an etch technique. Etching techniques consist either of dry etch or wet etch methods. In contrast to wet etch, dry etch can be anisotropic, but damages the wafer surface more than wet etching.

6.5.1 Wet Etching

Wet etching is a purely chemical process. Except for a possible dependence of the etch rate on the crystal orientation, it is therefore isotropic. It can be highly selective and harmless to the substrate. The etch rate depends on the transfer of the etchant species to the wafer, the chemical reaction with the exposed surface, and the transfer of the reaction products away from the wafer. The slowest step, the so called rate limiting step, determines the etch rate. Agitating the wet etch solution can therefore enhance the etch rate.

As wet etch does not harm the underlying substrate excessively, it is especially used for etching oxide which protects the substrate during processing [1, p. 255]:

$$SiO_2 + 6HF \rightarrow H_2 + SiF_6 + 2H_2O \tag{6.20}$$

The selectivity of HF solutions is usually higher than 100:1 with respect to oxide over silicon. Deposited oxides usually have a higher etch rate than thermal oxides. Dopant impurities further increase the etch rate. A $HNO_3 : H_2O : HF-$mixture can be used in order to remove $Si_{1-x}Ge_x$ selectively from Si [39] and a $KOH : K_2Cr_2O_7 : Propanol : H_2O-$ in order to remove Si selectively from $Si_{1-x}Ge_x$ [40].

6.5.2 Dry Etching

Dry etching can be anisotropic and also be much better controlled. Dry etching is synonymous with plasma assisted etch because plasma reactive and ion etch are the most popular methods. In general, dry etching is more accurate, but damages the wafer surface more than wet etching. The etch damage induced in the $Si/Si_{1-x}Ge_x$ is reduced for low power reactive ion etching [41, 42].

The plasma is an ionized gas with equal amount of positively and negatively charged particles. The gas does not have to be completely ionized. In ULSI, typically less than 10 % of the particles

are ionized. The plasma is formed between two electrodes at low pressure. Electrons accelerated by an RF field collide with atoms and molecules. After dissociation, they become extremely reactive and are absorbed at the wafer surface where they form volatile products. These products are pumped out of the reactor. Dry etching can be further categorized into physical and chemical methods.

Plasma etch techniques include the following standard methods [2, p. 344]:

- barrel etching (chemical)
- down stream plasma etching (chemical)
- reactive ion etching (RIE, chemical+physical)
- magnetic enhanced RIE (chemical+physical)
- magnetic confinement triode RIE(chemical+physical)
- electron cyclotron resonance etching (chemical+physical)
- inductively coupled plasma etching (ICP, chemical+physical)

In chemical etching, a plasma generates neutral reactive species which interact with the wafer surface to form volatile products. Chemical etching is associated with high etch rate, good selectivity, low damage and is always isotropic. CF_4 plasmas etches selectively Si on SiO_2 and SiO_2 on Si [2, p. 336]:

$$Si + F \rightarrow SiF_4 \tag{6.21}$$

$$SiO_2 + 4F \rightarrow SiF_4 + O_2 \tag{6.22}$$

The addition of small concentrations of O_2 increase the etch rate of both Si and SiO_2, but enhances the selectivity of Si with respect to SiO_2. In contrast, the addition of moderate concentrations of H_2 decreases the etch rate of both Si and SiO_2, but enhances the selectivity of SiO_2 with respect to Si. The same chemistry can be used for the etch of $Si_{1-x}Ge_x$ whose etch rate increases with x [43].

In physical etching, positive ions bombard the surface at high speed. In practice, physical etching techniques exist mostly in combination with chemical etching. Physical etching exhibits low selectivity, high damage and can be anisotropic. Pure physical etching consumes the mask at the same etch rate as the underlying layer and has a low throughput. Chlorine-, bromine-, and fluorine-based chemistry etches Si. Chlorine based chemistry is used in order to etch aluminium and tungsten, oxygen based chemistry in order to etch resist.

6.6 Wafer Cleaning Technology

The requirements for ultraclean and smooth Si surfaces increase with device scaling. Particles, native oxide, microroughness, metals, and organic contamination are detrimental to the wafer

surface. In general, cleaning leads to the degradation of the wafer surface and is never 100 % efficient. Thus, the first priority is to prevent the wafer contamination by maintaining a clean environment. The contamination of the wafer surface occurs from the equipment, ambient, gas, chemicals, and DI water. Wet cleans are critical for HMOSFETs because they can lead to the removal of Si which is part of the Si cap on top of the SiGe channel of p-MOSFETs. The control of the Si cap layer thickness is crucial in order to ensure the confinement of the carriers (see chap. 3, [44]).

The RCA was invented by Kern and Puotinen in 1960 and is still the predominant clean. It consists of the two sequential cleaning solutions SC1 (NH_4OH-H_2O_2-H_2O, 1:1:5 to 1:2:7, $70-80^oC$) and SC2 (HCL-H_2O_2-H_2O, 1:1:6 to 1:2:8, $70-80^oC$). SC means standard clean.

The SC1 removes effectively organic contaminants and particles. Particle contamination [2, p. 63] is critical because their size can be close to the feature size of ICs. Photoresist consists of organic material and is the main source of contamination apart the ambient [2, p. 71]. Organic contamination can change the oxidation rate and is not very dangerous. On the wafer surface, organic contamination prevents the removal of other contaminants. The etch rate of SC1 [45] depends exponentially on the Ge fraction [40]. Hence, residual Ge (see section 6.3) in the Si cap might enhance the etch rate considerably.

The adhesion of particles and organic contaminants is a strong function of the pH of the solution. Particles adhere to the wafer surface due to Van der Waals forces, electrical double layers, capillary forces and chemical bonds. The hydrogen peroxide oxidizes the silicon surface while the OH^- from the NH_4OH provides a negative charge on the silicon surface and the particle. The particles are repelled from the wafer surface. The megasonic clean further enhances the effectiveness of the SC1. Sonic pressure waves wet the particle at high power (300 W) and high frequency (850-900 kHz). The solvent can diffuse into the interface and the particle is easily removed.

The SC2 desorbs metal contaminants (except noble metals) due to its low pH value by forming soluble complexes while oxidizing the wafer surface. According to ref. [46], SC2 oxidizes blank Si, but does not etch SiO_2. A good controllability of the Si cap of HMOSFETs can therefore be achieved if native oxide is cleaned only by SC2.

Metal contamination has a deleterious impact even underneath its detection limit by TRXF and AAS. It results in structural defects at the interface, stacking faults during later oxidation or epitaxial growth, increased leakage current and reduced carrier lifetime. Noble metals adhere to the wafer surface due to their high electronegativity and are extremely difficult to remove, whereas other metals are incorporated into the oxide (Ca, Fe...).

An immersion in a mixture of HF-H_2O-H_2O_2 [2, p. 92] is proposed for the removal of the native oxide and noble metals. However, a complete removal of noble metals is impossible. The only secure means against contamination with noble metals is an ultraclean environment.

Cleaning and especially the combined oxidation and etching processes of the SC1 lead to the

Process	Etch Rate
$NH_4OH + SC1 + SC2 + HF$	4.33 $nm/stage$
$SC1 + HF$	3.0 $nm/stage$
$SC1 + HF + H_2O$	3.0 $nm/stage$
$H_2O_2 + HF$	1.8 $nm/stage$
$SC1$	0.144 $nm/min.$

Table 6.1: Etch rate of different cleans

increase of the surface roughness. The surface roughening effect decreases by reducing:

- the portion of the NH_4OH
- the temperature
- the cleaning time

Today, there are some new dry cleaning techniques (HF/H_2O vapor cleaning, ultraviolet-ozone cleaning [UVOC], laser cleaning [47]) and new cleaning sequences applying the same chemicals as the RCA (Advanced wet cleaning processes, IMEC), HF and ozone [48].

The Si etch rate of the cleaning sequences given in tab. 6.1 was investigated by Glyn Braithwaite. The cleaning sequence were repeated many times and the etch depth was measured by $\alpha-$stepper, AFM and Sloane DekTak 3030. The standard SUMC compositions, bath temperatures and batch times were used:

- SC1 ($NH_4OH : H_2O_2 : H_2O$): 1.2 : 1 : 21 for 10 min. at 72^oC
- SC2 ($HCL : H_2O_2 : H_2O$): 1.1 : 1 : 21.7 for 10 min. at 72^oC
- Fuming nitric acid clean (70 %): 10 min. at room temperature
- HF dip (2.4 %) : a few seconds to hydrophobic the wafer at room temperature
- H_2O_2 (30 %): 10 min. at room temperature

However, the etch rate for the Si cap of HMOSFETs may be considerably higher than the ones given in tab. 6.1 due to residual Ge in the Si cap.

6.7 Ion Implantation

Originally, dopants were introduced into semiconductors by diffusion and the dopant profile was difficult to control. The first commercial implanters [1, chap. 5] were sold in 1973 and soon became ubiquitous despite initial reluctance. Apart from doping, implantation can also be employed for Ge implantation, preamorphization and device isolation. Device isolation is not described in more detail and shortly mentioned here. Silicon on insulator uses either a high energy implant of N^+ (Separation by Implanted Nitrogen, SIMNI) or O^+ (Separation

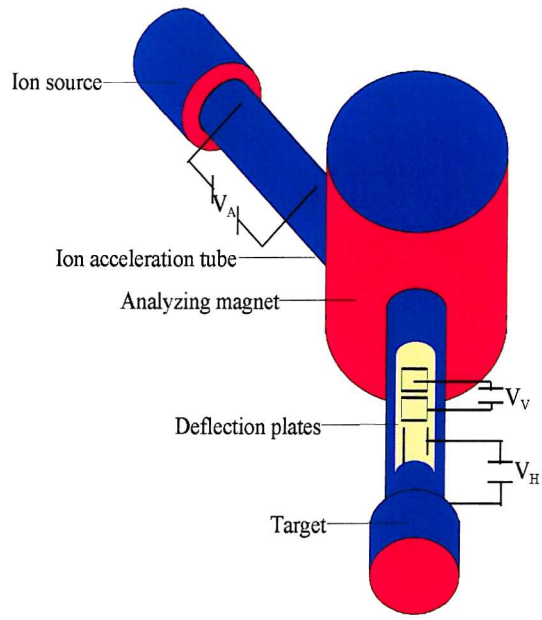


Figure 6.7: Schematic of an ion implanter

by Implanted Oxygen, SIMOX). In order to minimize the surface damage, the wafer must be heated during implantation. This is so important because the underlying oxide layer prevents a recrystallization from the substrate. A very high temperature step is performed in order to form the oxide.

6.7.1 Ion Implanter

The major components of an ion implanter can be seen in fig. 6.7. The ion source contains a feed gas. In Si technology, the most common feed gases are: BF_3, AsH_3, PH_3. Alternatively, a solid charge that is heated can be used. The resultant vapor is used as the source gas.

In a low pressure chamber, the feed gas is broken up into a variety of atomic and molecular species. The feed gas passes between a hot filament and a metal plate. The filament is highly negative with respect to the plate. Electrons that boil off the hot filament are accelerated towards the metal plate. On their way, they collide with feed gas molecules and molecular dissociation occurs if the energy is high enough. BF_3 breaks up into B, B^+, BF_2, BF_2^+, F^+, and some other species. The positive ions pass through a slit at the exit side of the source chamber which is negatively biased with respect to the filament. The ion beam is typically a few millimeters by one or two centimeters wide and consists of a variety of species.

The desired implant species is normally selected with an analyzing magnet. The analyzing magnet typically has a radius of up to 1 m. The ions can be accelerated either before of after the mass analysis. If the ions are accelerated before analysis, a larger magnet is required. However, the ions are much more likely not to lose their energy before reaching the sample surface. In general, the acceleration tube is several meters long and kept at a high vacuum in order to

prevent collisions during acceleration. In high current machines, the charge of the ions causes the beam to expand.

In many implanters, the ion beam is directed by both horizontal and vertical deflection plates on to the target. The current which is measured with a Faraday cup at the end station serves as a direct measurement of the implantation dose.

6.8 Implantation Profile

When an energetic ion enters the solid, it starts losing energy through interactions with electrons and lattice atoms. The range R is the distance that ions travel in the semiconductor, whereas the projected range R_p is the distance along the depth axis for an uniform beam. The ions are also scattered laterally and penetrate past the edges of a mask. In an amorphous solid, a Gaussian distribution can approximate the resulting impurity concentration $N(x)$ as a function of depth [1, p. 106]:

$$N(x) \simeq \frac{\phi}{\sqrt{2\pi}\Delta R_p} e^{(x-R_p)^2/2\Delta R_p^2} \tag{6.23}$$

Where ΔR_p and ϕ are respectively the standard deviation of the projected range and the dose.

In order to describe the boron profile, a Person type IV distribution is usually used. The profile is described by four moments [1, p. 108; 51, p. 94]:

1. The first moment or projected range R_p:

$$R_p = \int_0^\infty xN(x)dx \tag{6.24}$$

   R_p is the average penetration depth.

2. The second moment or straggle ΔR_p:

$$\Delta R_p = \int_0^\infty (x - R_p)^2 N(x)dx \tag{6.25}$$

   It is a measure of the width of the profile

3. The skewness γ is based on the third moment:

$$\gamma = \frac{\int_0^\infty (x - R_p)^3 N(x)dx}{(\Delta R_p)^3} \tag{6.26}$$

   It is a measure of the asymmetry of the profile. Negative values represent an increased concentration on the surface side.

4. The kurtosis β is based on the fourth moment:

$$\beta = \frac{\int_0^\infty (x - R_p)^4 N(x)dx}{(\Delta R_p)^4} \tag{6.27}$$

   The kurtosis increases with the flatness of the top of the implantation profile.

The Pearson type IV distribution is given by [1, p. 110]:

$$n(x) = n(R_p)\exp\frac{\ln\left[b_0 + b_1(x-R_p) + b_2(x-R_p)^2\right]}{2b_2} - \frac{b_1 + 2b_1b_2}{\sqrt{4b_0b_2^3 - b_1^2b_2^2}}\tan^{-1}\left[\frac{2b_2(x-R_p)+b_1}{\sqrt{4b_0b_2 - b_1^2}}\right] \tag{6.28}$$

Where:

$$b_o = -\frac{\Delta R_p^2(4\beta - 3\gamma^2)}{10\beta - 12\gamma^2 - 18} \tag{6.29}$$

$$b_1 = -\gamma\Delta R_p\frac{\beta+3}{10\beta - 12\gamma^2 - 18} \tag{6.30}$$

And:

$$b_2 = -\frac{2\beta - 3\gamma^2 - 6}{10\beta - 12\gamma^2 - 18} \tag{6.31}$$

In single crystals, channeling is possible when the ion velocity is parallel to a major crystal orientation. In a channel, nuclear stopping is not very effective and the electron density is low. The ion makes many glancing internal collisions that are nearly elastic. The ion continues either in the same direction until it loses all its energy or dechannels as a result of crystal defects or impurities. The incident ion can be redirected by a scattering event inside the target along a crystallographic orientation. However, as the probability for such an event is very low, this effect produces no substantial distortion of the implant profile for ion velocities that are not in parallel with a major crystallographic orientation.

The critical angle Ψ for channeling is given by [1, p. 111]:

$$\Psi = 9.73^\circ\sqrt{\frac{Z_iZ_t}{Ed}} \tag{6.32}$$

Where E is the incident energy in keV, d the atomic spacing along the crystallographic orientation in Å, Z_i the number of proton for the ion and Z_t the number of protons for the target. Submicron CMOS requires shallow junctions. A tilt angle or preamorphization with Si or Ge is normally used in order to prevent channeling. The preamorphization implant also serves for the prevention of transient enhanced diffusion (TED) that is extremely critical for boron. TED is due to the large number of interstitial atoms after implantation.

Shallow boron junctions are especially difficult to form because boron is very light and has therefore a deep projected range. Besides, it is difficult to lower the implant energy below 10 keV because of concerns about the beam stability. Therefore, BF_2 is often implanted instead of B. For a BF_2-implant, the kinetic energy associated with boron that is believed to dissociate upon impact is given by [1, p. 118]:

$$E_B = E_{BF_2}\frac{M_B}{M_{BF_2}} = E_{BF_2}\frac{11}{49} \tag{6.33}$$

Unfortunately, F has a negative effect on the reliability of ultrathin gate oxides although it can prevent boron diffusion (see chap. 6.3, [7, 49, 50]).

6.8.1 Implantation Damage and Solid Phase Epitaxial Growth

When ions are implanted into a crystal wafer, they can create point defects like interstitials and vacancies (primary defects). Some of the substrate atoms even gain a sufficient energy to eject some further atoms. Vacancies in the atomic lattice can occur if the energy loss per collision is high enough to overcome the binding energy of the Si target atoms ($E_b \sim 15\ eV/atom$). When a critical dose of ions is implanted, no evidence of long range order exists any more and the substrate is rendered amorphous. The critical dose is determined by the implant energy, the implant species, the target material and the substrate temperature during implantation. The critical dose becomes very large at high temperature because the substrate self-anneals during implantation. Heavy ions lose a higher fraction of their energy by interactions with lattice atoms and cause therefore more damage than light atoms. Strained SiGe heterojunctions can support a low dose implant whereas a high dose implant is detrimental [52-54].

In solid phase epitaxial growth (SPEG) [1, p. 112], the crystallinity is restored in the high temperature anneal (typically between $550 - 1000^oC$) during regrowth from the crystalline substrate. Besides, the anneal moves a large fraction of the implanted impurities to lattice sites. Depending on the species, the implanted atoms can act as dopants or form a heterojunction. During the annealing of the implanted crystal, the high potential energy of the defects is reduced by recombination or agglomeration.

However, microislands of single crystal material can also serve as nucleation centers during SPEG. When disparate growth fronts meet, defects are created. Therefore, layers regrown by SPEG can contain 1-D defects like dislocations loops, divacancies, but also 2-D and 3-D defects like twins and stacking faults.

SPEG enables the activation of dopants at much lower temperature than conventional dopant activation. The thermal recrystallization rate decreases exponentially with inverse temperature. For amorphous Si onto a $< 100 >$ Si substrate, the thermal recrystallization rate is $10^{-2}\ cm/s$ for 1000^oC and $10^{-7}\ cm/s$ for 600^oC [55].

Defects formed during the recrystallization (secondary defects) can be categorized into five different groups [56, p. 8], [1, p. 177]:

- Type I defects arise for doses exceeding a critical value while no amorphous layer is formed. They are typically located close to the projected range $\bar{R}_p$ of the implanted species and consist mainly of dislocations loops formed during the annihilation of a supersaturated number of interstitial point defects.

- Type II defects or end-of-range (EOR) defects occur after solid phase epitaxial growth (SPEG) of an amorphous layer and are formed below the amorphous/crystalline (a/c) interface in the implanted material.

- Type III defects are due to the imperfect SPEG of the amorphous layer. They extend from

the a/c interface to the surface. Among typical type III defects (microtwins, segregation related defects...), hairpin dislocations are the most common defects.

- When the threshold damage density (TDD) for the formation of an amorphous layer is only slightly surpassed, the amorphous layer is buried and the regrowth starts at both a/c interfaces. Type IV defects are formed where the two advancing a/c interfaces intercept and consist of perfect or faulted dislocation loops known as "clamshell" defects.

- Type V defects arise when the solid solubility of the implanted species in the crystalline silicon is exceeded and a second chemical phase is formed.

For SiGe heterojunctions [57-64], only type II and III defects are expected as Si and Ge are completely miscible and the implant concentrations are high.

The EOR defects are the result of the agglomeration of interstitial Si atoms that are shaped as dislocation loops at the a/c interface. Due to their interstitial nature, EOR defects introduce compressive strain. The density of EOR defects increases slightly with dose, but dramatically with implantation energy and implantation temperature. For low annealing temperatures ($700-800^oC$), the number of loops decreases, while the mean radius of the loops augments and the number of atoms bound by the loops stays constant. A high-temperature anneal ($> 900^oC$) reduces the number of atoms bound by the loops and finally dissolves the dislocation loops. The EOR defects can be spatially separated from the SiGe heterojunction by applying a Si implant which has a higher energy than the Ge implant (EPIFAB process [65-69]).

Hairpin dislocations nucleate at the advancing a/c interface during the recrystallization of the amorphous layer. The number of "hairpin defects" is reduced for low implantation temperature and low implantation energy. A high mass and dose of the implanted ions results in severe hairpin dislocations. An annealing step at very high temperatures ($\sim 1150^oC$) can completely eliminate hairpin dislocations.

6.9 Silicidation

Silicide films [70-86] are used for the gate, source, and drain of MOSFETs to reduce the series resistance and therefore increase the switching speed. Silicide films are formed either by codepositing a metal and Si (polycide process) or in a solid-state reaction between a metal and Si (salicide process). Polycides are often used on gates in memory chips whereas salicides is preferred for logic chips.

6.9.1 Metal-Semiconductor Contact

In general, the work function of the metal Φ_m is not equal to the work function of the semiconductor Φ_s ($\Phi_m \neq \Phi_s$). Hence, there is a potential barrier E_b at the interface of a metal and a

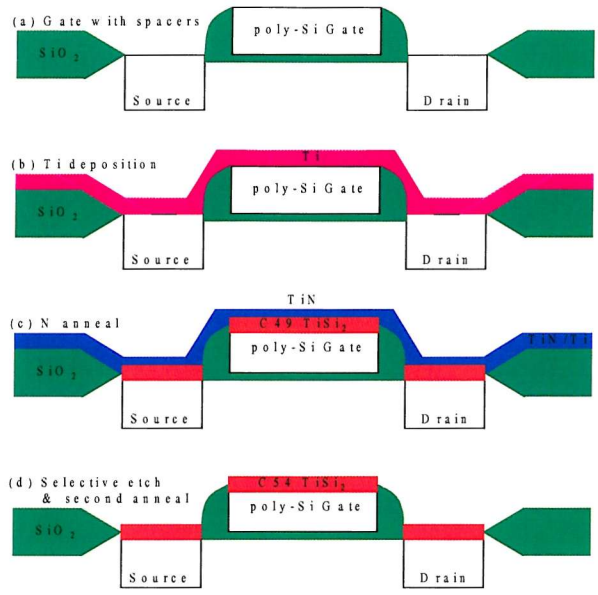


Figure 6.8: Schematic diagram of Ti salicide process

semiconductor analog to the flatband voltage in a MOS structure (see chap. 3):

$$E_b = q(\Phi_m - \Phi_s) = q\phi_B \tag{6.34}$$

Where ϕ_B is the barrier height.

The potential barrier results in a contact resistance R_c defined as [71, p. 304]:

$$R_c = \left[\frac{\partial j}{\partial V}\right]^{-1}_{V=0} \tag{6.35}$$

Where j is the current density and V the voltage across the metal-semiconductor interface.

The contact is rectifying when a depletion region exists at the interface in the semiconductor and ohmic when an accumulation region exists at the interface in the semiconductor (see tab. 6.2).

The thermionic emission current density across a rectifying metal-semiconductor contact is given by [71, p. 262]:

$$j = J_s(\exp(qV/kT) - 1) \tag{6.36}$$

The saturation current density J_s is given by:

$$J_s = AT^2 \exp(-q\phi_B/kT) \tag{6.37}$$

Where A and V are the Richardson's constant and the voltage at the metal-Si interface respectively.

The width of the depletion region W in the semiconductor at the rectifying metal-semiconductor contact is given by [72, p. 254]:

$$W = \sqrt{\frac{2\epsilon_s}{qN_b}\phi_B} \tag{6.38}$$

Semiconductor	R_c
$\Phi_m > \Phi_s$	
n	rectifying
p	ohmic
$\Phi_m < \Phi_s$	
n	ohmic
p	rectifying

Table 6.2: Four different cases of the contact resistance

The size of the depletion region decreases as ϕ_B decreases and N_b increases. Apart from the thermionic emission current, there are still the diffusion and the tunneling currents. Doping the semiconductor heavily and choosing a metal for which ϕ_B is small as for silicides, the width of the depletion layer decreases and tunneling is easy. As only the active dopant concentration contributes to the lowering of the contact resistance, the activation step is crucial for reducing the contact resistance. An RTA anneal is ideal because the fast ramp down rates freeze in the high temperature dopant activation while diffusion is minimized. SiGe leads to a better dopant activation and therefore an increase in the tunnel current [73, 74].

The contact resistance R_c between a metal and a heavily doped semiconductor is proportional to [70]:

$$R_c \sim \exp\left(\frac{4\pi\phi_B}{h}\right)\sqrt{\frac{\epsilon_s m^*}{N_b}} \tag{6.39}$$

Where h is the Planck's constant and m^* the effective mass.
Ultra-shallow junction can be achieved by the outdiffusion from doped silicide which is implanted with Ge^+ [75, 76] in order to enhance the outdiffusion. The decrease in yield, degradation of the gate oxide and oxide interface, and counterdoping have been reported as negative side effects of silicide [70].

6.9.2 Silicidation Process

Salicide $TiSi_2$ has a relatively high thermal stability and a low resistivity. It is the most common silicide for salicide application, but $CoSi_2$ is under investigation as an alternative [77-79]. A Ti salicide (self-aligned silicide) process [70] is schematically shown in fig. 6.8. First of all, the gate stack is created and source, drain, and gate are implanted (a). Then, Ti is deposited by sputtering (b). During a nitrogen anneal at a temperature between 600 and 700^oC, Ti reacts with exposed Si to form $C49$ $TiSi_2$ (c). The $C49$ phase has a base-centered orthorhombic structure and a high resistivity ($60-300$ $\mu\Omega cm$). 2.3 nm of Si are consumed per nm of Ti in the solid-state reaction.

Nucleation limited reactions as it is the case for $C49$ $TiSi_2$ are expected when the free energy change ΔG for the reaction is small with respect to the increase in surface energy σ. The free

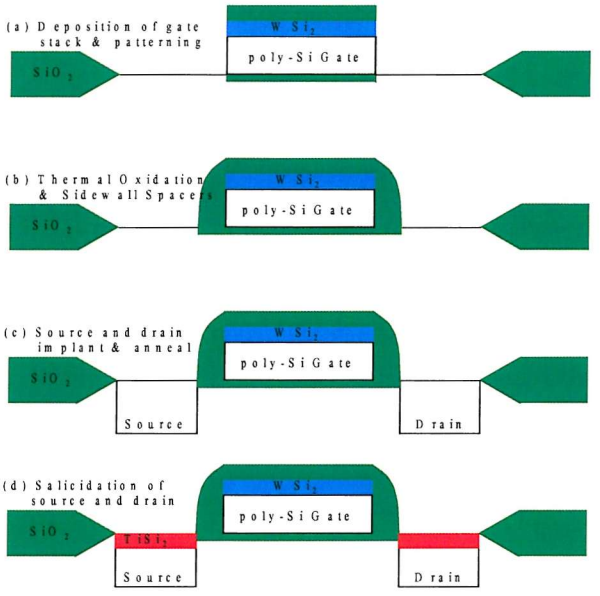


Figure 6.9: Schematic diagram of W polycide process

energy per unit volume ΔG_n for a nucleus of radius r is given by:

$$\Delta G_n = \sigma \cdot 4\pi r^2 - \Delta G \cdot \frac{4}{3}\pi r^3 \tag{6.40}$$

Only nuclei surpassing a critical size $r^* = 2\sigma/\Delta G$ can grow. The number of nuclei is a strong function of temperature. Hence, the metal-Ti reaction takes place in a very limited temperature window and can produce rough films.

A low oxygen level ensures that the metal is not oxidized and can be achieved more easily in an RTA chamber than in a furnace. Besides, RTA is ideal for nucleation limited reactions requiring a a high temperature for a short time. The metal-Si reaction is very sensitive to contamination and all the native oxide has to be removed before Ti deposition.

The nitrogen diffuses into the Si, forms TiN in the field regions and over the spacers, and blocks the lateral movement of Si. Thus, the shorting of adjacent contacts (bridging) due to the diffusion of Si and lateral silicide formation can be prevented by the use of a relatively low annealing temperature and an nitrogen ambient. The TiN formation is diffusion limited and reduces the amount of Ti available for the formation of $C49$ $TiSi_2$.

The TiN and unreacted Ti are removed in a preferential wet etch, typically $NH_4OH : H_2O_2 : H_2O$ $(1:1:5)$. The etch mechanism is not yet understood. Then, $C49$ $TiSi_2$ is transformed into less resistive face-centered orthorhombic $C54$ $TiSi_2$ $(15-35\ \mu\Omega cm)$ in a second anneal above $(800^oC$, (d)). This transformation is as well nucleation limited and dependent on the thickness of the $C49$ $TiSi_2$ layer. Grain boundary triple points are the preferred nucleation sites for thin films (55 nm or less), whereas grain boundaries are favored for films exceeding 100 nm. The surface-to-volume ratio, the nucleation barrier and in consequence the temperature required for the transformation of the $C49$-phase to the $C54$-phase increase with decreasing film thickness.

Polycide WSi_2 is the most common silicide for polycide application because of its high thermal stability, low resistivity, and ease of patterning by dry etching processes. A W polycide process [70] used for the gate stack is shown in fig. 6.9. The deposition of WSi_2 on top of the doped poly-Si is followed by the deposition of oxide. The stoichiometry of the deposited silicide must be well controlled in order to ensure low resistivity. The removal of the native oxide before WSi_2 deposition is essential for reducing the contact resistance. The deposited amorphous or polycrystalline film with a hexagonal structure transforms to a tetragonal phase at $600°C$.

After the etching of the gate stack, a brief thermal oxidation is performed to restore the integrity of the gate oxide (a). Then, the sidewalls are formed (b), and the source and drain are implanted and annealed (c). The source and drain can be salicided in a salicide process (d).

6.10 Conclusion

An overview over essential process steps already known from conventional Si technology was given. The existence of a strained SiGe layer channel is especially critical for process steps affecting the substrate surface, ie cleaning and oxidation. The cleaning of SiGe is not yet investigated in detail. However, the SC1 can etch SiGe at a considerable rate and might have to be omitted.

High temperature steps can relax strained SiGe layers, lead to the outdiffusion of Ge, and therefore be detrimental for HMOS devices. RTP reduces the thermal budget. It might be beneficial to HMOS depending on the effect caused by the increase of thermal stress.

Thermal oxidized HMOSFETs typically display a high interface state density (see sections 4.3 and 6.3). Besides, thermal oxidation can lead to strain relaxation, Ge and dopant diffusion. Deposited high-K dielectrics which are supposed to replace the thermal gate oxide eventually may be advisable for HMOSFETs.

Etch processes have to be adapted, whereas other steps such as lithography and chemical/mechanical polishing are completely identical for SiGe and Si technology. Alternatively to CVD and PVD, implantation can be used for the fabrication of SiGe heterojunctions. A high dose implant into a SiGe layer should be prevented. As an enhanced dopant activation is achieved in SiGe, Ge can be used in conjunction with silicide for the reduction of the contact resistance.

Bibliography

[1] S. A. Campbell, "The Science and Engineering of Microelectronic Fabrication", Oxford University Press, 1996

[2] C. Y. Chang and S. M. Sze, "ULSI Technology", McGraw-Hill, 1996

[3] D. J. Paul, J. M. Ryan, M. Pepper, A. N. Broers, T. E. Whall, J. M. Fernandez, and B. A. Joyce, "Electron Beam Induced Damage of Silicon Germanium", Microelectron. Eng., Vol. 35, pp. 59-62, 1997

[4] K. Rajan, R. Singh, J. Adler, U. Mahajan, Y. Rabinovich, and B. M. Moudgil, "Surface interaction forces in chemical-mechanical planarization", Thin Solid Films, Vol. 308-309, pp. 529-532, 1997

[5] C. G. Kallingal, D. J. Duquette, and S. P. Murarka, "An Investigation of Slurry Chemistry Used in Chemical Mechanical Planarization of Aluminium", J. Electrochem. Soc., Vol. 145, No. 6, pp. 2074-2081, 1998

[6] J.-Y. Tsai, Y. Shi, S. Prasad, S. W.-C. Yeh, and R. Rakkit, "Slight Gate Oxide Thickness Increase in PMOS Devices with BF_2 Implanted Polysilicon Gate", IEEE Electron Dev. Lett., Vol. 19, No. 9, pp. 348-350, 1998

[7] C. Lin, A. I. Chou, K. Kumar, P. Choudhury, and J. C. Lee, "Effect of BF_2 implantation on ultrathin gate oxide reliability", Appl. Phys. Lett., Vol. 69, No. 11, pp. 1591-1592, 1996

[8] S. J. Kilpatrick, R. J. Jaccodine, and P. E. Thompson, "A diffusional model for the oxidation behavior of $Si_{1-x}Ge_x$ alloys", J. Appl. Phys., Vol. 81, No. 12, pp. 8018-8028, 1997

[9] P.-E. Hellberg, S.-L. Zhang, F.-M. d'Heurle, and C. S. Petersson, "Oxidation of silicon-germanium alloys. I. An experimental study", J. Appl. Phys., Vol 82, No. 11, pp. 5773-5778, 1997

[10] P.-E. Hellberg, S.-L. Zhang, F.-M. d'Heurle, and C. S. Petersson, "Oxidation of silicon-germanium alloys. II. A mathematical model", J. Appl. Phys., Vol 82, No. 11, pp. 5779-5787, 1997

[11] F. K. LeGoues, R. Rosenberg, and B. S. Meyerson, "Kinetics and mechanism of oxidation of SiGe: dry versus wet oxidation", Appl. Phys. Lett., Vol. 54, No. 7, pp. 644-646, 1989

[12] M. Mukhopadhyay, S. K. Ray, T. B. Ghosh, M. Sreemany, and C. K. Maiti, "Interface properties of thin oxide layers grown on strained SiGe layers at low temperatures", Semicond. Sci. Technol., Vol. 11, pp. 360-365, 1996

[13] D. J. Norris, A. G. Cullis, T. J. Grasby, and E. H. C. Parker, "Investigation of nanoscale Ge segregation in p-channel SiGe/Si field effect transistor structures by electron energy loss imaging", J. Appl. Phys., Vol. 82, No. 12, pp. 7183-7185, 1999

[14] D. Tchikatilov, Y. F. Yang, and E. S. Yang, "Improvement of SiGe oxide grown by electron cyclotron resonance using H_2O vapor annealing", Appl. Phys. Lett., Vol. 69, No. 17, pp. 2578-2580, 1996

[15] C.-G. Ahn, H.-S. Kang, Y.-K. Kwon, and B. Kang, "Effects of Segregated Ge on Electrical Properties of $SiO_2/SiGe$ Interface", Jpn. J. Appl. Phys., Vol. 37, pp. 1316-1319, 1998

[16] J. Min, P. K. Chu, Y. C. Cheng, J. B. Liu, S. Im, S. Iyer, and N. W. Cheung, "Buried oxide formation by plasma immersion ion implantation", Mat. Chem. Phys., Vol. 40, pp. 219-222, 1995

[17] J. P. Zhang, P. L. F. Hemment, J. E. Castle, H. D. Liu, J. F. Watts, R. A. Kubiak, S. M. Newstead, T. E. Whall, and E. H. C. Parker,"Thermodynamic behaviour of GeO_2 formed by oxygen implantation into relaxed $Si_{0.5}Ge_{0.5}$ alloy", Nucl. Instr. and Meth., Vol. B 96, pp. 281-285, 1995

[18] C. Tetelin, X. Wallart, J. P. Nys, and L. Vescan, "Germanium Behaviour During the Low-temperature Plasma-assisted Oxidation of SiGe Alloys", Surf. Interface Anal., Vol. 23, pp. 363-366, 1995

[19] L. K. Bera, M. Mukhopadhyay, S. K. Ray, D. K. Nayak, N. Usami, Y. Shiraki, and C. K. Maiti, " Oxidation of strained Si in a microwave electron cyclotron resonance plasma", Appl. Phys. Lett., Vol. 70, No. 2, pp. 217-219, 1997

[20] L. K. Bera, S. K. Ray, D. K. Nayak, N. Usami, Y. Shiraki, and C. K. Maiti, "Electrical properties of oxides grown on strained Si using microwave N_2O plasma", Appl. Phys. Lett., Vol. 70, No. 1, pp. 66-68, 1997

[21] L-P. Chen, Y.-C. Chan, S.-J. Chang, G.-W. Huang, and C.-Y. Chang, "Direct oxidation of $Si_{1-x}Ge_x$ Layers Using Vacuum-Ultra-Violet Light radiation in Oxygen", Jpn. J. Appl. Phys., Vol. 37, Part 2, No. 2A, pp. L122-L124, 1998

[22] C. Tetelin, X. Wallart, L. Vescan, and J. P. Nys, "Plasma assisted oxidation of SiGe layers at 500^oC: interface characterization", Appl. Surf. Sci., Vol. 104/105, pp. 385-391, 1996

[23] I. S. Goh, J. F. Zhang, S. Hall, W. Eccleston, and K. Werner, " Plasma Oxidation of Si and SiGe", Microelectron. Eng., Vol 28, pp. 221-224, 1995

[24] A. Sareen, P. Lundgren, and S. Bengtsson, "Low Thermal Budget Gate Oxides on $Si/Si_{1-x}Ge_x/Si$ Substrates", ULIS'2001 Workshop, pp. 49-52, Grenoble, 2001

[25] J. C. Yeoh, P. W. Green, T. J. Thornton, S. Kaya, K. Fobelets, and J. M. Fernandez, "MOS gated Si:SiGe quantum wells formed by anodic oxidation", Semicon. Sci. Technol., Vol. 13, pp. 1442-1445, 1998

[26] J. Rappich, I. Sieber, A. Schoepke, and W. Fuessel, "Influence of a silicon cap on SiGe passivation by anodic oxidation", Mat. Res. Soc. Symp. Proc., Vol. 451, pp. 215-220, 1997

[27] I. S. Go, J. F. Zhang, S. Hall, W. Eccleston, and K. Werner, "Electrical Properties of plasma-grown Oxide on MBE-grown SiGe", Semicond. Sci. Technol., Vol. 10, pp. 818-828, 1995

[28] M. Glueck, J. Hersener, H. G. Umbach, J. Rappich, and J. Stein, "Implementation of Low Thermal Budget Techniques to Si and SiGe MOSFET Device Processing", Solid-State Phenomena, Vols. 57-58, pp. 413-418, 1997

[29] R. M. Sidek, U. N. Straube, A. M. Waite, A.G. R. Evans, C. Parry, P. Phillips. T. E. Whall, and E. H. C. Parker, "SiGe CMOS fabrication using SiGe MBE and anodic/LTO gate oxide", Semicond. Sci. Technol. 15, pp. 135-138, 2000

[30] R. S. Prasad, T. J. Thornton, S. Kanjanachuchai, J. Fernandez, and A. Matsumura, "Mobility degradation in gated Si:SiGe quantum wells with thermally grown oxides", Electronics Lett., Vol. 31, No. 21, pp. 1876-1878, 1995

[31] C. T. Lin, S. J. Chang, D. K. Nayak, and Y. Shiraki, "The properties of SiO_2 films using direct photo-chemical vapor deposition on strained SiGe layers", Appl. Sur. Sci., Vol. 92, pp. 193-197, 1996

[32] C.-T. Lin, S.-J. Chang, D. K. Nayak, and Y. Shiraki, "Deposition of SiO_2 Films on Strained $SiGe$ Layer by Direct Photo Chemical Vapor Deposition", Jpn. J. Appl. Phys., Vol. 34, Part. 1, No. 1, pp. 72-72, 1995

[33] J. C. Lee, B. Lee, L. Kang, R. Nieh, K. Onishi, C. Kang, R. Choi, E. Dharmarajan, S. Goplan, and Y. Leon, "Ultra-thin High-K Gate Dielectrics", ULIS'2001 Workshop, pp. 67-72, Grenoble, 2001

[34] E. Aderstedt and P. Lundgren, "Refined DC model of the Tunnel Emitter Transistor", ULIS'2001 Workshop, pp. 85-88, Grenoble, 2001

[35] M. Mukhopadhyay, S. K. Ray, D. K. Nayak, and C. K. Maiti, "Ultrathin oxides using N_2O on strained $Si_{1-x}Ge_x$ layers", Appl. Phys. Lett., Vol. 68, No. 9, pp. 1262-1264, 1996

[36] E. Dufour-Gergam, F. Meyer, F. Delmotte, M. C. Hugon, B. Agius, P. Warren, and D. Dutartre, "Electrical properties of silicon nitride films grown on a SiGe layer by distributed electron cyclotron resonance plasma-enhanced chemical vapor deposition", Thin Solid Films, Vol. 94, pp. 214-216, 1997

[37] J. Salzer, "Rapid Thermal processing market", Solid-State Technology, pp. 92-97, July 1998

[38] H. Habuka, T. Otsuka, M. Mayusumi, M. Shimada, and K. Okuyama, "A Direct Approach for Evaluating the Thermal Condition of a Silicon Substrate under Infrared Rays and Specular Reflectors", J. Electrochem. Soc., Vol. 146, No. 2, pp. 713-718, 1999

[39] D. J. Godbey, A. H. Krist, K. D. Hobart, and M. E. Twigg, "Selective Removal of $Si_{1-x}Ge_x$ from (100) Si using HNO_3 and HF", J. Electrochem. Soc., Vol. 139, No. 10, pp. 2943-2945, 1992

[40] F. S. Scott, D. S. Miles, D. T. Grider, and J. J. Wortman, "Selective Chemical Etching of Polycrystalline $SiGe$ Alloys with Respect to Si and SiO_2", J. Electronic Mat., Vol. 21, No. 8, pp. 805-810, 1992

[41] R. G. van Veen, M. J. Teepen, E. van der Drift, T. Zijlstra, E. J. M. Fakkeldij, K. Werner, A. H. Verbruggen, and S. Radelaar, "Chemical vs. physical factors in dry etching induced damage in the Si/Ge_xSi1-x system", Microelectronic Engineering, Vol. 35, pp. 55-58, 1997

[42] K. Y. Lee, S. J. Koester, K. Ismail, and J. O. Chu, "Electrical characterization of $Si/Si_{0.7}Ge_{0.3}$ quantum well wires fabricated by low damage CF_4 reactive ion etching", Microelectron. Eng., Vol. 35, pp. 33-36, 1997

[43] N. S. Lloyd, "Interactions between molecules and surfaces", Ph.D. thesis, Southampton, 1998

[44] S. Verdonckt-Vandebroek, E. F. Crabbe, B. S. Meyerson, D. L. Harame, P. J. Restle, J. M. C. Stork, and J. B. Johnson, "SiGe-Channel Heterojunction p-MOSFET's", IEEE Trans. Electron Dev., Vol. 41, No. 1, pp. 90-101, 1994

[45] H. Kobayashi, J. Ryuta, T. Shingyouji, and Y. Shimanuki, "Study of Si Etch Rate in Various Composition of SC1 Solution", Jpn. J. Appl. Phys. Vol. 32, Part 2, No. 1A/B, L45-L47, 1993

[46] C. R. Inomata, H. Ogawa, K. Ishikawa, and S. Fujimura, "Infrared Spectroscopy Study of Chemical Oxides Formed by a Sequence of RCA Standard Cleaning Treatments", J. Electrochem. Soc., Vol. 143, No. 9, pp. 2995-3000, 1996

[47] Y.-F. Lu, W.-D. Song, and T.-S. Low, "Laser cleaning of micro-particles from a solid surface - theory and applications", Mat. Chem. Phys., Vol. 54, pp. 181-185, 1998

[48] T. Ohmi, "Total Room Temperature Wet Cleaning for Si Substrate Surface", J. Electrochem. Soc., Vol. 143, No. 9, pp. 2957-2964, 1996

[49] F. Boussaid, M. Benzohra, F. Olivie, D. Alquier, and A. Martinez, "Electrically active defects in BF_2^+ implanted and germanium preamorphized silicon", Nucl. Instrum. Methods Phys. Res. B, Vol. 134, pp. 195-201, 1998

[50] L. Z. Wang, M. S. C. Luo, H. H. Tseng, and S. A. Ajuria, "The influence of fluorine on boron-enhanced diffusion in silicon by BF_2^+ implantation through oxide during high temperature rapid thermal anneal", J. Electrochem. Soc., Vol. 144, pp. L298-L301, 1997

[51] E. Rimini, "Ion implantation: basics to device fabrication", Kluwer Academic Publishers, 1995

[52] D. Y. C. Lie, "Doping and Processing Epitaxial Ge_xSi_{1-x} Films on $Si(100)$ by Ion Implantation for Si-Based Heterojunction Devices Applications", J. Electron Mat., Vol. 27, No. 5, pp. 377-401, 1998

[53] M. S. Oh, M. H. Joo, S. Im, H. B. Kim, H. K. Kim, and J. H. Song, "Effects of BF_2^+ implantation on the strain-relaxation of pseudomorphic metastable $Ge_{0.06}Si_{0.94}$ alloy layers", Nucl. Instrum. Methods Phys. Res. B, Vol. 147, pp. 49-55, 1999

[54] S. Im, F. Eisen, M. A. Nicolet, M. O. Tanner, K. L. Wang, and N. D. Theodore, "Strain-conserving doping of a pseudomorphic metastable $Ge_{0.06}Si_{0.94}$ layer on Si(100) by low-dose BF_2^+ implantation", J. Appl. Phys., Vol. 81, pp. 1695-1699, 1997

[55] Olson and Roth, Mater. Sci. Rept., Vol. 3, 1998

[56] F. Christiano, "Extended defects in SiGe device structures formed by ion implantation", Ph.D. Thesis, University of Surrey, 1998

[57] H. Jiang and R. G. Elliman, "Electrical Properties of GeSi Surface- and Buried-Channel p-MOSFET's Fabricated by Ge Implantation", IEEE Transactions on Electron Devices, Vol. 43, No. 1, pp. 97-103, 1996

[58] R. G. Elliman and W. C. Wong, "The fabrication of epitaxial Ge_xSi_{1-x} layers by ion implantation", Nuc. Instr. and Meth. in Phys. Res., Vol. B80/81, pp. 768-772, 1993

[59] Z. Xia, "Solid Phase epitaxial regrowth of $Si_{1-x}Ge_x$ layers formed by single-energy Ge^+, double-energy Si^+ and Ge^+, and Ge^+ and Ge^{2+} ion implantation", Mat. Sci. Eng , Vol. B 34, pp. 42-46, 1995

[60] R. G. Elliman, W.-C. Wong, and P. Kringhoj, "Solid-Phase Epitaxial Crystallisation of Ge_xSi_{1-x} Alloy Layers", Mat. Res. Soc. Symp. Proc., Vol. 321, pp. 375-385, 1994

[61] K. C. Liu, S. K. Ray, S. K. Oswal, and S. K. Banerjee, "A deep submicron $Si_{1-x}Ge_x/Si$ Vertical PMOSFET Fabricated by Ge Ion Implantation", IEEE Electron Dev. Lett., Vol. 19, No. 1, pp. 13-15, 1998

[62] S. Kal and N. B. Chakrabarti, "Strained Silicon-SiGe Devices using Germanium Implantation", IETE J. Res., Vol. 43, No. 2 & 3, pp. 185-192, 1997

[63] Z. Xia, E. O. Ristolainen, and P. H. Holloway, "$SiGe/Si$ heterostructures produced by double-energy Si^+ and Ge^+, and Ge^+ and Ge^{2+} ion implantations", J. Vac. Sci. Technol. B, Vol. 16, No. 5, pp. 2629-2632, 1998

[64] C. J. Patel and J. B. Butcher, "Germanium redistribution phenomena in the synthesis of SiGe layers", Mat. Res. Soc. Symp. Proc., Vol. 439, pp. 215-220, 1997

[65] Y. Suprun-Belevich, F. Christiano, A. Nejim, P. L. F. Hemment, and B. J. Sealy, "Evolution of mechanical strain and extended defects in annealed (1 0 0) silicon samples implanted with Ge^+ ions ", Nucl. Instr. and Meth. in Phys. Res. B 140, pp. 91-98, 1998

[66] A. Nejim, N. P. Barradas, C. Jeynes, F. Cristiano, E. Wendler, K. Gaertner, and B. J. Sealy, "Residual post anneal damage of Ge and C co-implantation of Si determined by quantitative RBS-channeling", Nucl. Instr. and Meth. in Phys. Res. B 139, pp. 244-248, 1998

[67] F. Christiano, A. Nejim, B. de Maudit, A. Claverie, and P. L. F. Hemment, "Characterization of extended defects in SiGe alloys formed by high dose Ge^+ implantation into Si", Nucl. Instr. and Meth. in Phys. Res. B 120, pp. 156-160, 1996

[68] F. Christiano, A. Nejim, Y. Suprun-Belevich, A. Claverie, and P. L. F. Hemment, "Formation of extended defects and strain relaxation in ion beam synthesised SiGe alloys", Nucl. Instr. and Meth. in Phys. Res. B 147, pp. 35-42, 1999

[69] F. Christiano, C. Bonafos, A. Nejim, S. Lombardo, M. Omri, D. Alquier, A. Martinez, S. U. Campisano, P. L. F. Hemment, and A. Claverie, "Interstitial trapping efficiency of C^+ implanted into preamorphised silicon - control of EOR defects", Nucl. Instr. and Meth. in Phys. Res. B 127 / 128, pp. 22-26, 1997

[70] J. P. Gambino and E. G. Colgan, "Silicides and ohmic contacts", Mat. Chem. Phys., Vol. 52, pp. 99-146, 1998

[71] S. M. Sze, "Physics of Semiconductor Devices", John Wiley & Sons, 1981

[72] Henry Mathieu, "Physique des semiconducteur et des composants électroniques", Masson, 1996

[73] I. Mizushima, A. Murakoshi, K. Suguro, N. Aoki, and J. Yamauchi, "Ultra high dose boron ion implantation: super-saturation of boron and its application", Mat. Chem. Phys., Vol. 54-59, 1998

[74] W. Freiman, R. Beserman, and K. Dettmer, "Titanium metallization of Si/Ge alloys and superlattices", Thin Solid Films, Vol. 294, pp. 217-219, 1997

[75] C. M. Osburn and K. R. Bellur, "Low parasitic resistance contacts for scaled ULSI devices", Thin Solid Films, Vol. 332, pp. 428-436, 1998

[76] C.-H. Chu, K.-J. Ho, C.-T. Huang, S.-H. Shvu, and T.-F. Lei, "Controlling the diffusion of implanted boron in Si and silicide by multiple implants", Mat. Chem. Phys., Vol. 54, pp. 6-66, 1998

[77] K. Inoue, K. Mikagi, H. Abiko, S. Chikaki, and T. Kikkawa, "A New Cobalt Salicide Technology for $0.15-\mu m$ CMOS Devices", IEEE Trans. Electron Devices, Vol. 45, No. 11, pp. 2312-2318, 1998

[78] H. Sumi, K. Tajima, J. Suenga, H. Harifuchi, and Y. Okamoto, "Comparative Evaluation of Ti Salicide and Co Salicide Processes for CMOS Devices", Electronics and Communications in Japan, Part 2, Vol. 81, No. 8, pp. 26-31, 1998

[79] J. A. Kittl, Q. Z. Hong, H. Yang, N. Yu, S. B. Samavedam, and M. A. Gribelyuk, "Advanced Salicides for 0.10 μm CMOS: Co salicide processes with low diode leakage and Ti salicide processes with direct formation of low resistivity C54 $TiSi_2$", Thin Solid Films, Vol. 332, pp. 404-411, 1998

[80] J. P. Snyder, C. R. Helms, and J. Nishi, "Experimental investigation of a PtSi source and drain field emission transistor", Appl. Phys. Lett., Vol. 67, No. 10, pp. 1420-1422, 1995

[81] S. J. Koester, R. Hammond, and J. O. Chu, "Extremely High Transconductance $Ge/Si_{0.4}Ge_{0.6}$ p-MODFET's Grown by UHV-CVD", IEEE Electron Dev. Lett., Vol. 21, No. 3, pp. 110-112, 2000

[82] M. H. Juang, M. J. Tsai, M. C. Hu, and C. J. Yang, "FORMATION OF NISI-SILICIDED SHALLOW P^+N JUNCTIONS BY BF_2^+ IMPLANTATION INTO THIN NI OR NISI FILMS ON SI SUBSTRATES AND SUBSEQUENT ANNEAL", Solid-St. Electron., Vol. 42, No. 11, pp. 1953-1958, 1998

[83] P. S. Lee, K. L. Pey, D. Mangelinck, J. Ding, A. T. S. Wee, and L. Chan, "Improved NiSi Salicide Process Using Presilicide N_2^+ Implant for MOSFET's", IEEE Electron Dev. Lett, Vol. 21, No. 12, pp. 566-568, 2000

[84] J. F. Liu, J. Y. Feng, and J. Zhu, "Comparison of the thermal stability of NiSi films in Ni/Pt/(111)Si and Ni/Pt/(100)Si systems", J. Appl. Phys., Vol. 90, No. 2, pp. 745-749, 2001

[85] P. Wang, J. Yin, W. Sheng, Y. Zheng, L. Zhao, and D. Xu, "Effect of deposition and treatment conditions on growth of nanometer PtSi heterostructure", J. Vac. Sci. Technol., Vol. B 18, No. 5, pp. 2406-2410, 2000

[86] H. Iwai, T. Ohguro, and S. Ohmi, "NiSi salicide technology for scaled CMOS", Microelectronic Engineering, Vol. 60, pp. 157-169, 2002

Chapter 7

Pseudomorphic SiGe CMOS Process

Problems in the fabrication of SiGe CMOS chips may be as severe as those arising during heterolayer growth. A remarkable performance improvement of SiGe p-MOSFETs in comparison with Si reference p-MOSFETs was demonstrated in ref. [1]. However, the process was tailored for the low thermal budget constraints imposed by SiGe. In contrast, the investigated pseudomorphic SiGe CMOS process addresses the requirements for an industrial type of CMOS process. A thermal gate oxide is grown and the MBE layer is counterdoped for the n-MOSFETs. The devices experience a considerable thermal budget and the substrate is doped at a level typical of conventional Si MOSFETs.

The aim of the presented batch is to approach gently a competitive industrial process. The standard cleaning procedures are therefore employed. As mentioned in chapter 4 and 6, the control of the Si cap layer is critical. The estimation of the removed Si cap layer thickness is based on the cleaning experiments given in section 6.6. The calculations assume that the Si cap layer thickness has the properties of pure Si. However, latest research suggests (see section 6.3, [2]) that the interface between the SiGe channel and Si cap layer is not abrupt. The considerable amount of Ge in the Si cap can enhance the etch rate during cleaning (see section 6.6).

No difference in the electrical characteristics of the fabricated SiGe and Si reference p-MOSFETs were found. Extensive x-ray, TEM, and SIMS analysis showed that this was due to a lack of germanium in the active channel. Instead of discussing the electrical results in detail, it is investigated why the expected $Si_{1-x}Ge_x$ layers are not present. The MOSFETs suffer from processing problems, which are the results of required adaptations (LTO field oxide, thermal budget, layer architecture, substrate doping). The batch was useful in that it highlights problems in the manufacturing of competitive SiGe MOSFETs.

7.1 Fabrication Process

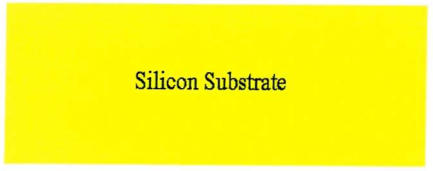
The process is depicted in fig. 7.1. A complete process listing in given in appendix B. The process sequence starts with the growth of a thin thermal oxide (20 nm @ 900^oC) that protects the substrate against contamination. Consecutively, alignment marks are etched in the 4-inch wafers (p-type, CZ, $17-33\ \Omega cm$ resistivity, step 1). N-well and p-well are respectively formed by the implantation of phosphorus ($2.0 \cdot 10^{13}\ at./cm^2$ @ 160 keV, $4.0 \cdot 10^{12}\ at./cm^2$@ 70 keV) and boron ($2.0 \cdot 10^{13} at./cm^2$ @ 160 keV, $4.0 \cdot 10^{12}\ at./cm^2$@ 25 keV, step 2). An RTA ($10''$ @ 1100^oC, step 3) anneals the implantation damage and activates the dopants.

At Warwick university, the protective oxide is wet etched. For different wafers, different MBE layers are grown over the whole wafer surface by solid-source molecular beam epitaxy (MBE) in a VG Semicon V90S MBE system as MBE is non-selective (step 4). The thick Si buffer ensures a good crystal quality of the SiGe heterostructures and in consequence good transport properties. Uniform and linearly graded Ge profiles are used (0 % , 20 % , 36 % and 0-40 % Ge). The side with the low Ge content of the linearly graded $0-40$ % SiGe layer is close to the substrate.

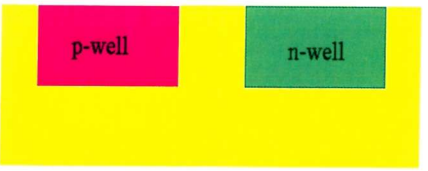
The process was simulated with TSUPREM-4 and MEDICI. In order to achieve a threshold voltage of $-0.5\ V$ for the SiGe p-MOSFETs with a linearly graded $0-40$ % Ge channel for undepleted poly-Si gates and good turn-off characteristics, the MBE layer is insitu doped. An implant could destroy the integrity of the SiGe heterostructure. The intended antimony profile and the layer structure are depicted in fig. 7.2. The red line is the technologically realizable antimony profile which tails off much slower than the ideal dopant profile (broken blue line) because of surface segregation during MBE growth. Besides, the shape of the Sb profile allows us to counterdope it easily as required for the n-MOSFETs.

After the MBE growth, a 20 nm low-pressure CVD oxide (LTO) is deposited in order to protect the substrate against contamination (step 5). In the p-well, the n-type doping of the MBE layer is compensated by a boron implant ($2.5 \cdot 10^{13}\ at./cm^2$ @ 20 keV, step 6) and is expected to result in a threshold voltage of 0.5 V for the n-MOSFETs. Another 500 nm of LTO are deposited (step 7) for device isolation. Consecutively, the active area is wet etched (step 8). Wet etch is very selective and removes only a tiny fraction of a Si substrate. However, HF can etch SiGe at a much higher rate (chap. 6.5, [3]). The wet etch turned out to be rather inhomogeneous. Therefore, the field oxide was not etched completely in some active areas. This explains a reduced performance of some MOSFETs.

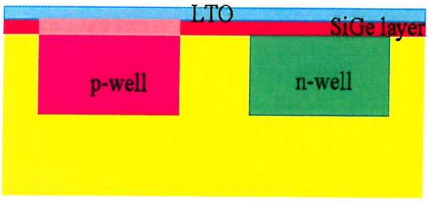
A 6 nm gate oxide is grown at 800^oC (step 9). Then, 200 nm of undoped amorphous silicon is deposited as gate material (step 10). The Si surface under the field oxide receives a full RCA clean between MBE growth and field oxide deposition, whereas the Si surface in the active areas receives two RCAs and an HF dip. The etch rate of the SC-1 increases exponentially with the Ge content (chap. 6.6, [4]), whereas the SC-2 removes only a negligible amount of Si [5]. According to an etch experiment carried out by Glyn Braithwaite (see chap 6.6), the remaining Si cap



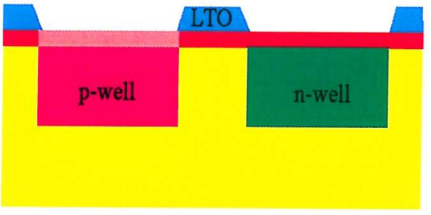
Step 1: Alignment Masks



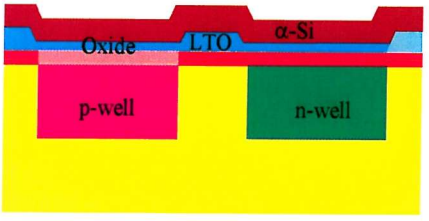
Step 2+3: n- and p-well Implantation and Furnace Anneal



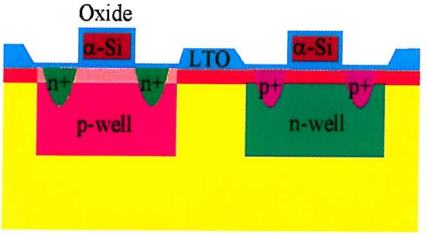
Step 4+5+6: Growth of the SiGe layer, LTO Deposition (thin), and Counterdoping Implant



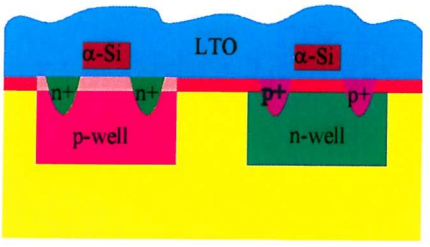
Step 7+8: LTO Deposition (thick) and Etch of Active Area



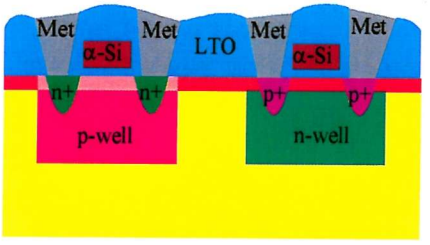
Step 9+10: Gate Oxidation and Amorphous Silicon Deposition



Step 11+12+13: Polysilicon Etch, Reoxidation Stage, and Implantation



Step 14+15: LTO Deposition, RTA



Step 16+17:Contact Windows and Metallisation

Figure 7.1: Schematic representation of the pseudomorphic SiGe CMOS process, the dimensions are not to scale

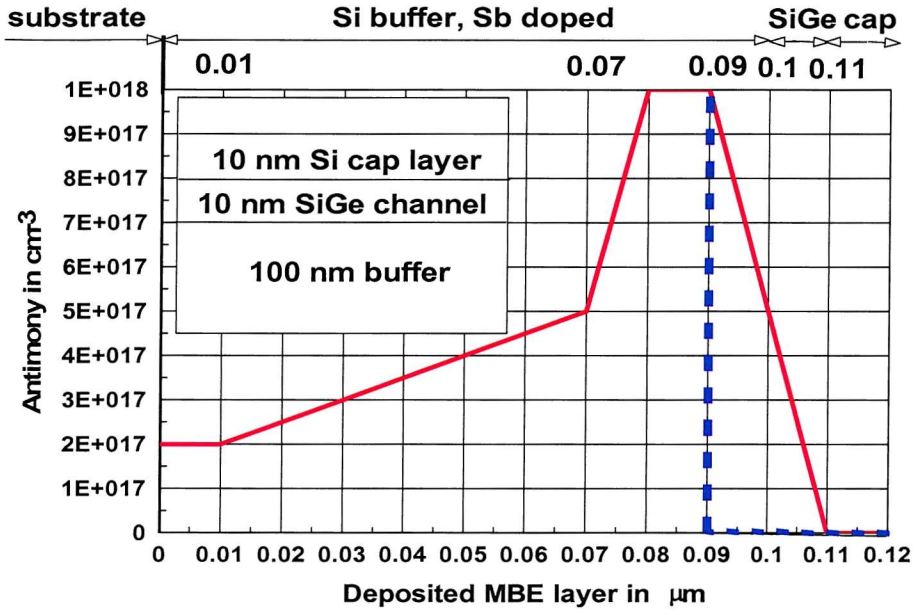


Figure 7.2: Substrate doping profile and layer structure

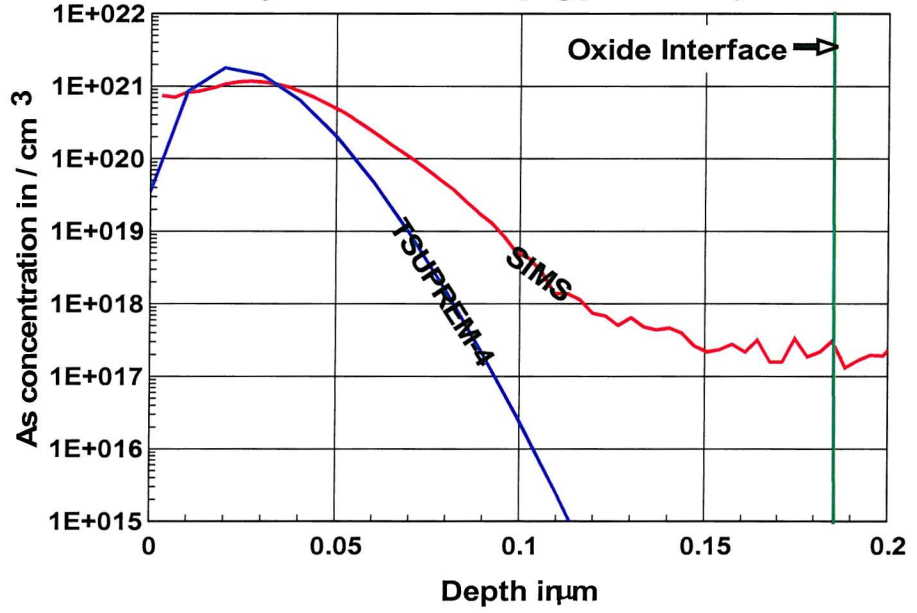
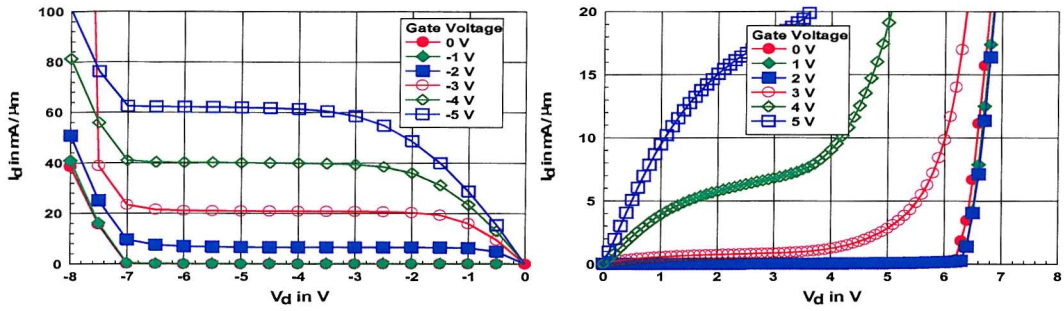


Figure 7.3: Gate doping profile for an As^+ implant at a dose of $5 \cdot 10^{15}/cm^2$ at 40 keV



(a) (b)

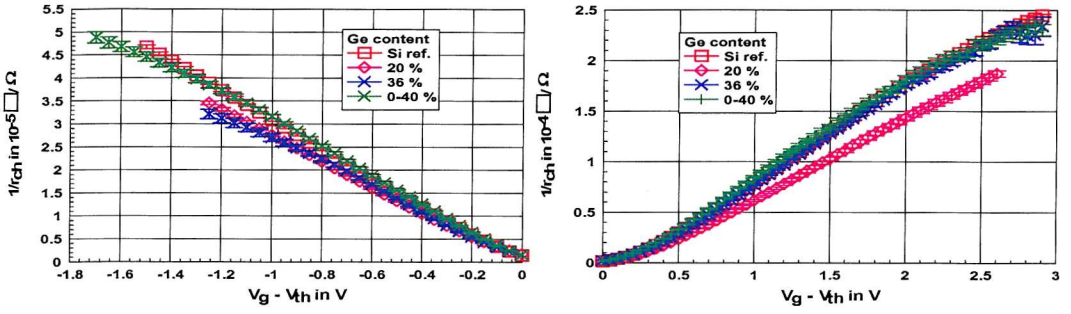
Figure 7.4: Si MOSFET characteristics ($L = 1\ \mu m$): (a) p-MOSFET annealed for 1 min. at 900^oC (b) n-MOSFET annealed for 1 min. at 850^oC

thickness is estimated as approximately 2 nm.

The gates are defined by e-beam lithography (step 11). A poly reoxidation stage is performed in order to anneal the damage due to the polysilicon etch and in order to grow a thin protective oxide (6 nm @ 800^oC, step 12). $5 \cdot 10^{15}\ at./cm^2$ @ $40\ keV$ of arsenic and $5 \cdot 10^{15}\ at./cm^2$ @ $45\ keV$ of boron difluoride are implanted into gate, source and drain respectively of n-MOSFET and p-MOSFET in a selfaligning process. The implant energies and the thickness of the amorphous Si layer to be used were determined in a previous implantation experiment.

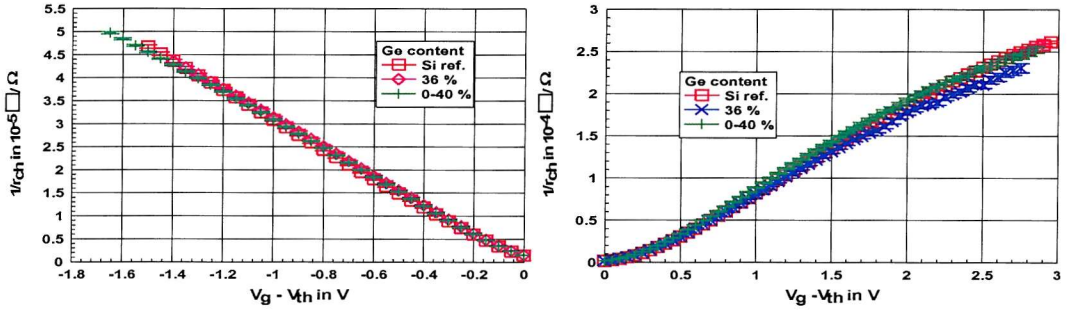
Fig. 7.3 shows that the implantation profile is deeper than simulated by TSUPREM-4 and has a much wider tail. Due to the wide tail, it is impossible to achieve a high dopant concentration close to the gate oxide without a considerable penetration in the substrate. However, this would lead to a high off-state current and might destroy the integrity of the SiGe channel. As the thermal budget has to be low and the implantation of the SiGe channel has to be prevented, a sufficient concentration of dopants may not diffuse to the gate oxide in order to avoid poly depletion. Insitu doped poly-Si is not available, but will be used for limited area/virtual substrate PMOS batch.

Furthermore, 400 nm of LTO (step 14) is deposited for planarization. Three alternative RTAs (1 min. @ 850^oC, 3 min. @ 850^oC, 1 min. @ 900^oC) are used for the diffusion of the dopants in the gate and the activation of source, drain and gate dopants (step 15). After the creation of the contact windows (step 16), a thin (100 nm) titanium layer and an aluminium layer ($1\ \mu m$) are deposited (step 17). The titanium layer lowers the contact resistance between metal and semiconductor. The aluminium contains 1 % Si in order to prevent the spiking of aluminium. The metallic wires are created in a dry etch. Finally, an alloy anneal is performed for 30 min. at 420^oC. The process is completed by several cleaning steps.



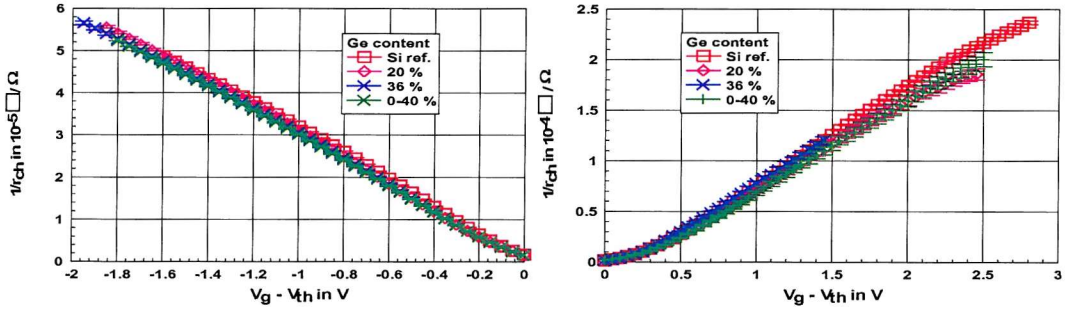
(a) (b)

Figure 7.5: $1/r_{ch}$ for 1 min. anneal at 850^oC: (a) p-MOSFETs (b) n-MOSFETs



(a) (b)

Figure 7.6: $1/r_{ch}$ for 3 min. anneal at 850^oC: (a) p-MOSFETs (b) n-MOSFETs



(a) (b)

Figure 7.7: $1/r_{ch}$ for 1 min. anneal at 900^oC: (a) p-MOSFETs (b) n-MOSFETs

7.2 Electrical Results

On each wafer, n- and p-MOSFETs with a channel width of 20 μm and a channel length of 20, 5, 2, and 1 μm were measured for $V_d = 50mV$. The gate voltage was in general varied from 0 to - 3 V for the p-MOSFETs and from 0 to 5 V for the n-MOSFETs. The routine discussed in section 4.5 was used for parameter extraction. For the n-MOSFETs, the whole measured voltage range was not always used due to gate leakage. The mobility is not extracted because the low values for the HF capacitance indicate gate depletion (see also fig. 7.8). Typically, very dispersed results were found for MOSFETs with identical design on all wafers.

Poly depletion can explain slight variations in the MOSFET performance. A variation in the thickness of the amorphous layer combined with a low annealing time and temperature can lead to a different dopant concentration in the gate at the gate oxide. This effect should be more severe for decreasing annealing temperature and time.

The etch of the active area was an extremely critical stage for this process. On the one hand, the Si cap layer on top of the SiGe channel had to be extremely thin in order to assure the confinement of the carriers in the SiGe channel. On the other hand, the cap layer should not be completely removed during processing. The etch of the field oxide in the active area was nonuniform and had to be stopped before the field oxide was completely removed in all active areas in order to ensure that enough Si was left for further processing on top of the SiGe layer and in order to to prevent the undercutting of the field oxide.

In general, it was observed across all wafers that some active areas were perfectly etched whereas in some active areas an important fraction of the active area remained covered by field oxide. The incomplete etch of the field oxide in the active area should reduce the device performance considerably and the MOSFETs affected by an incomplete etch of the active area should easily be identified.

In order to eliminate the error introduced by the dispersion in the device characteristics, the six best devices among eighteen with relatively good device characteristics and the same dimensions were evaluated. It has to be stressed that there is an interdependence between the yield and the device performance. A low yield can cause a low extracted device performance because not many devices with a good performance are found.

Fig. 7.4 shows that the p-MOSFETs have good device characteristics whereas the n-MOSFETs have poor punchthrough characteristics. The counterdoping of an n-doped substrate in order to produce n-MOSFETs is definitively not ideal. This route was chosen as an alternative to the removal of the thick MBE layer. It stresses the importance to grow MOSFETs with a thin buffer layer which does not have to be insitu doped as further alternatives would require a far more complicated process.

The n- and p-MOSFETs for all splits have a similar channel sheet resistance, respectively (fig.'s 7.5- 7.7). The large dispersion in the measurements has an important impact on the

Ge content in %	R_{sd} in $k\Omega\mu m$	ΔL in μm	V_{th} $(L = 1\ \mu m)$ in V	S $(L = 1\ \mu m)$ in mV/dec
	Split annealed for 1 min. at 850^oC			
0	6.18 ± 1.88	0.16 ± 0.06	-1.50 ± 0.08	130.6 ± 8.6
20	7.28 ± 2.7	0.03 ± 0.07	-1.60 ± 0.02	133.7 ± 4.1
36	25.28 ± 5.04	0.18 ± 0.11	-1.50 ± 0.02	121.6 ± 9.8
0-40	5.68 ± 2.08	0.21 ± 0.05	-1.20 ± 0.01	109.5 ± 3.0
	Split annealed for 3 min. at 850^oC			
0	6.00 ± 0.76	0.22 ± 0.02	-1.45 ± 0.01	118.1 ± 1.3
36	7.14 ± 1.96	0.01 ± 0.05	-1.10 ± 0.01	126.6 ± 2.0
0-40	4.3 ± 1.26	0.12 ± 0.04	-1.40 ± 0.07	156.8 ± 9.8
	Split annealed for 1 min. at 900^oC			
0	5.84 ± 2.52	-0.12 ± 0.08	-0.90 ± 0.09	151.7 ± 25.6
20	2.06 ± 1.60	-0.03 ± 0.05	-1.05 ± 0.02	173.5 ± 5.8
36	2.42 ± 0.96	0.10 ± 0.03	-0.95 ± 0.01	141.5 ± 1.2
0-40	3.30 ± 1.10	0.04 ± 0.03	-1.20 ± 0.04	141.9 ± 1.6

Table 7.1: R_{sd}, ΔL, V_{th}, and S for the p-MOSFETs

Ge cont. in %	R_{sd} in $k\Omega\mu m$	ΔL in μm	V_{th} $(L = 1\ \mu m)$ in V	S $(L = 1\ \mu m)$ in mV/dec
	Split annealed for 1 min. at 850^oC			
0	1.70 ± 0.24	0.14 ± 0.02	2.05 ± 0.02	239.5 ± 3.3
20	1.51 ± 0.07	0.13 ± 0.04	2.10 ± 0.05	272.1 ± 6.1
36	1.41 ± 0.29	-0.01 ± 0.02	1.95 ± 0.04	270.0 ± 6.8
0-40	0.47 ± 0.62	-0.04 ± 0.05	2.05 ± 0.01	279.4 ± 31.5
	Split annealed for 3 min. at 850^oC			
0	2.44 ± 0.15	0.16 ± 0.01	1.75 ± 0.04	229.5 ± 3.0
36	1.54 ± 0.47	-0.08 ± 0.04	2.15 ± 0.02	244.7 ± 9.5
0-40	0.67 ± 0.32	-0.02 ± 0.03	2.10 ± 0.02	236.5 ± 5.1
	Split annealed for 1 min. at 900^oC			
0	2.98 ± 0.21	0.26 ± 0.02	1.90 ± 0.06	225.8 ± 4.3
36	0.24 ± 0.83	-0.04 ± 0.06	1.95 ± 0.03	219.6 ± 6.9
20	1.94 ± 0.77	0.22 ± 0.05	2.15 ± 0.02	221.3 ± 5.6
0-40	1.42 ± 1.49	-0.24 ± 0.10	2.35 ± 0.02	274.9 ± 9.6

Table 7.2: R_{sd}, ΔL, V_{th}, and S for the n-MOSFETs

extraction of R_{sd} and ΔL (see tab.'s 7.1 and 7.2) and results in unreasonable values in some cases (E.g., 36 % Ge, split annealed for 1 min. at $850°C$. The same wafer has a pretty high channel sheet resistance). Whereas the threshold voltage for the n-MOSFETs is in general far higher than intended, the magnitude of the p-MOSFET threshold voltage decreases with the activation temperature and time and is the only parameter clearly depending on the annealing conditions. A substantial percentage of the MOSFETs display gate leakage which can be attributed to poly-Si fillets left in the corners of the active areas. These fillets are a common problems of processes with active areas etched into the field oxide because the edges of the active area are very steep.

7.3 Analysis of the Layer Structure

The wafers were further analyzed by SIMS, TEM, and x-ray. Fig. 7.8 (a) and (b) show a SIMS profile of a completely processed capacitor annealed for 1 min. at $900°C$ with the substrate dopant profile of an p-MOSFET and gate dopant profile of a n-MOSFET and with the substrate dopant profile of a n-MOSFET and gate dopant profile of an p-MOSFET, respectively. The profiles themselves are distorted by a thin film left in the contact windows after the removal of the metal and field oxide which covered a part of the profiled area on top of the poly-Si gate. The step at the surface of the poly gates results in the twofold measurement of the actual dopant profile shifted by a distance given by the thickness of the layer left in the active areas, but does not hamper the detection of a Ge peak.

The low dopant concentration in the gate at the gate oxide interface suggests that the MOSFETs are susceptible to poly depletion. An even lower dopant concentration is expected for the MOSFETs annealed at $850°C$. A 36 % Ge peak was expected between the oxide interface (red line) and the peak in the antimony profile (yellow curve). However, only residual Ge was found in the active area and under the field oxide as it was confirmed by further analysis for all SiGe wafers.

It is questionable why the SiGe MOSFETs do not have the expected SiGe layer. The SiGe layer must either not have been grown by MBE or have been removed during cleaning. The literature about the etch of SiGe is very sparse. The data given in ref. [4] and the data of the internal etch experiment carried out by G. Braithwaite is used in order to investigate the removal of SiGe during SC-1. It is assumed the the SC-2 etches a negligible amount of SiGe as it is the case for Si [5]. The Ge profile as a function of depth $x(d)$ takes into account that the cap layer can contain a substantial percentage of Ge (see fig. 7.9):

$$x(d) = \begin{cases} 0.36 \cdot \exp\left(d-10\right)^2/b & : \quad d < 10\ nm \\ 0.36 & : \quad 10\ nm \leq d \leq 20\ nm \\ 0.36 \cdot \exp\left(d-20\right)^2/b & : \quad 20\ nm < d \end{cases} \tag{7.1}$$

The parameter b determines the width of the Ge tail outside the channel.

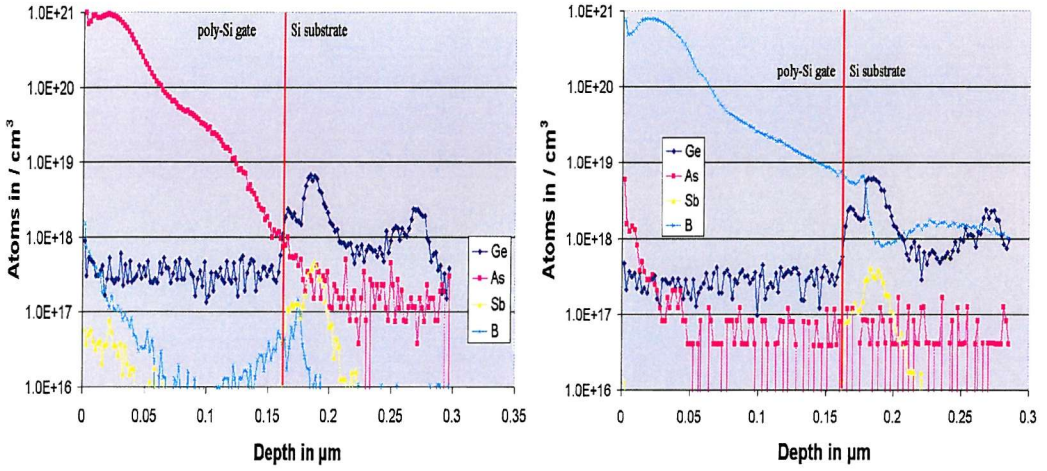


Figure 7.8: SIMS profile of capacitor: (a) p-MOSFET substrate (b) n-MOSFET substrate

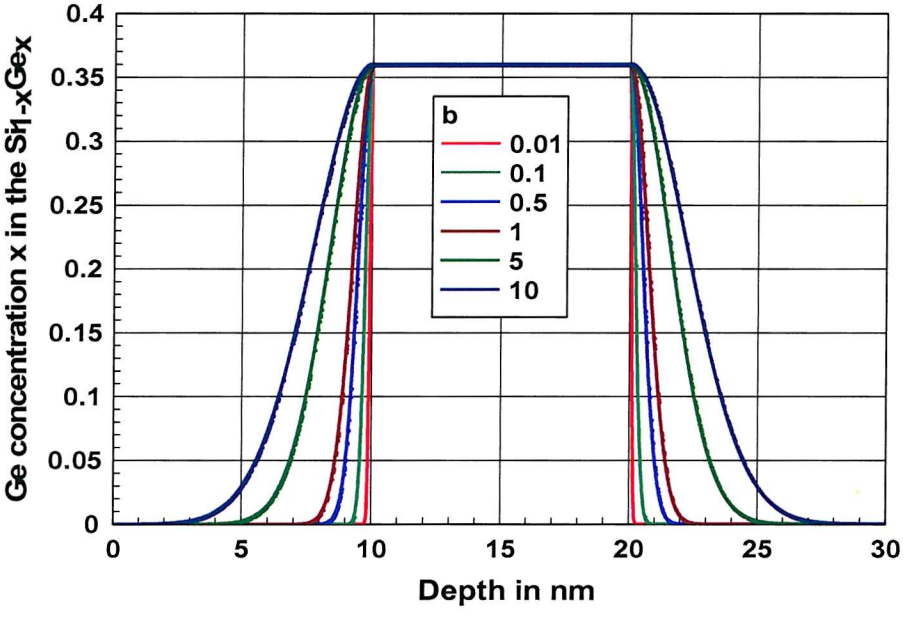


Figure 7.9: Ge profile of the $Si_{1-x}Ge_x$ channel

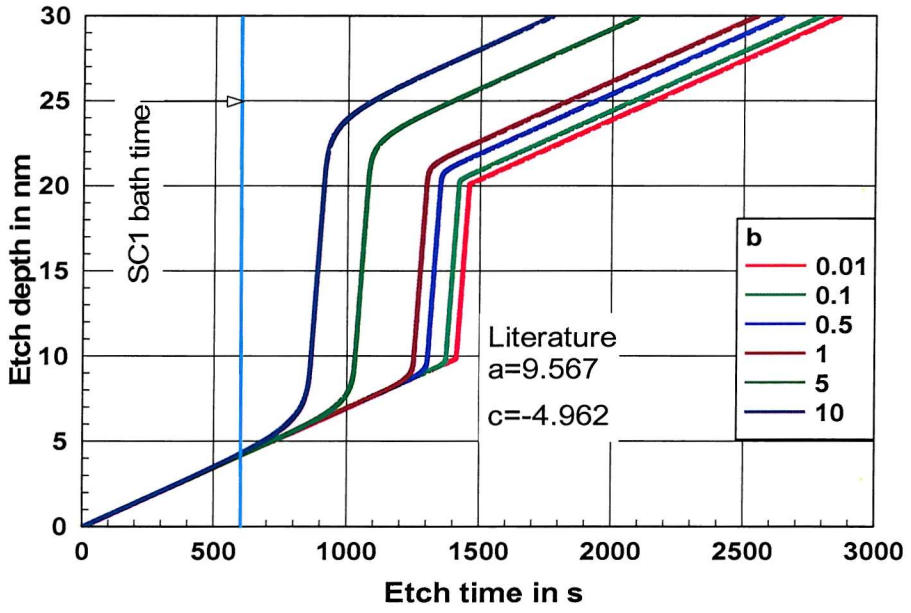


Figure 7.10: Etch depth as a function of SC-1 bath time using values found in literature

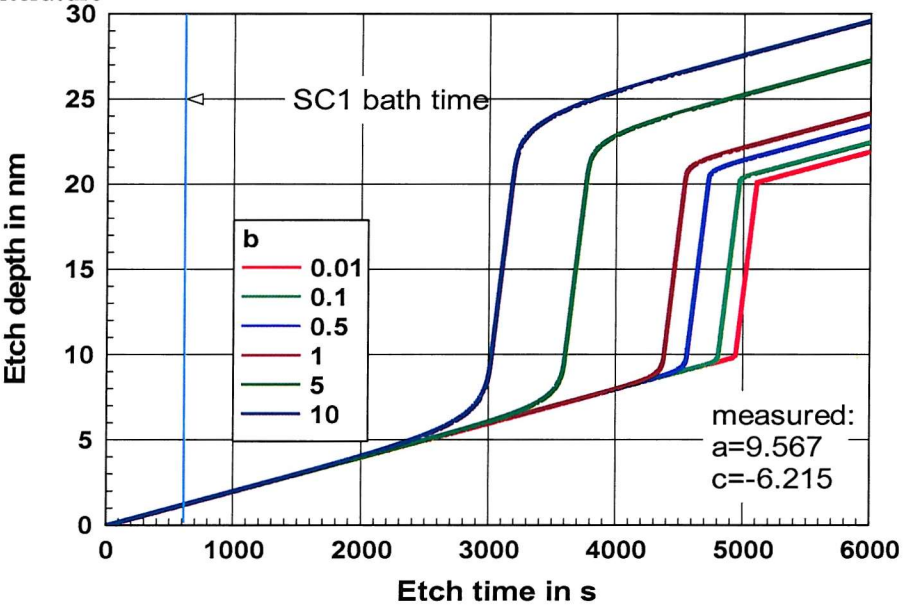
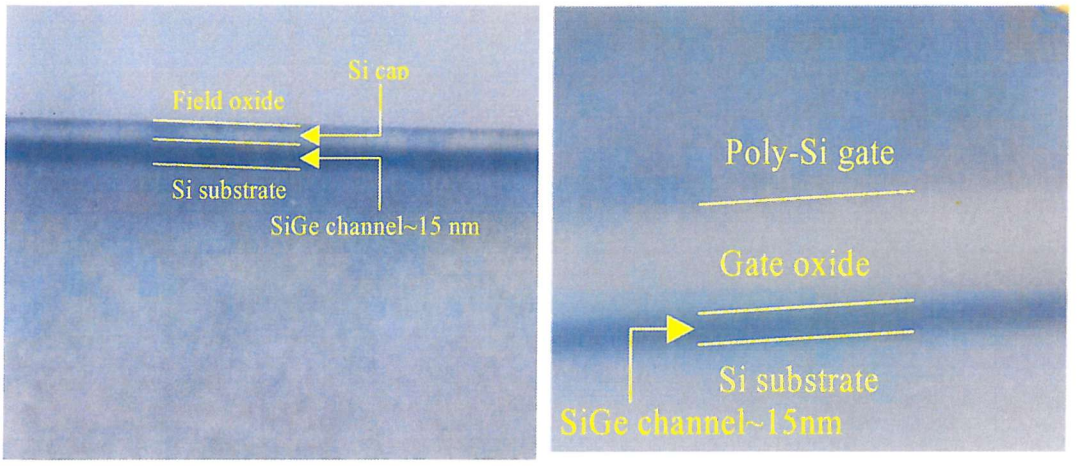


Figure 7.11: Etch depth as a function of SC-1 bath time using own the Si etch rate determined in the internal cleaning experiment



(a) (b)

Figure 7.12: MBE layer for the previous SiGe MBE CMOS batch (a) under the field oxide (b) under the gate oxide

The etch rate $r(x)$ is exponentially dependent on the Ge content [4]:

$$r(x) = \exp(a \cdot x + c)(nm/s) \tag{7.2}$$

Fig.'s 7.10 and 7.11 show the etch depth as a function of the etch time based on the data given in [4] and on c determined from the internal cleaning experiments combined with a determined from [4], respectively. A single SC-1 as it is seen by the wafer surface under the field oxide should definitively not remove the SiGe layer. This suggests that the SiGe layers may not have been grown. However, the limited data available is valid for poly-$Si_{1-x}Ge_x$ and we can not be sure that it is also the case for $Si_{1-x}Ge_x$.

For a previous SiGe MBE batch [1], the strained SiGe MBE layer was found in the processed devices. Initially, a 15 nm Si cap layer was grown on top of a 15 nm SiGe channel. Underneath the field oxide, 1.5 nm of Si is expected to be removed by an RCA clean followed by a fuming nitric acid clean. Fig. 7.12 a shows that the thickness of the remaining cap layer is comparable to the thickness of the SiGe layer. 6 nm of Si is expected to be removed by the cleaning sequence RCA-fuming nitric acid clean-HF-RCA before the deposition of the LTO gate oxide.

In contrast, no Si cap layer was found on top of the SiGe channel under the gate oxide (see fig. 7.12 (b)). This indicates that the results of the internal cleaning experiments may not completely apply to the Si cap on top of the SiGe channel.

A definite answer to the question whether the SiGe layer could have been removed by a single RCA as it was experienced by the wafer surface under the field oxide can only be obtained from a cleaning experiment. Therefore, an MBE layer with a 10 nm thick $Si_{0.8}Ge_{0.2}$ layer covered by a 10 nm Si cap was grown over a wafer as a replica of the specified layer architecture for batch presented in this chapter. The wafer underwent an RCA, the deposition of 20 nm of LTO,

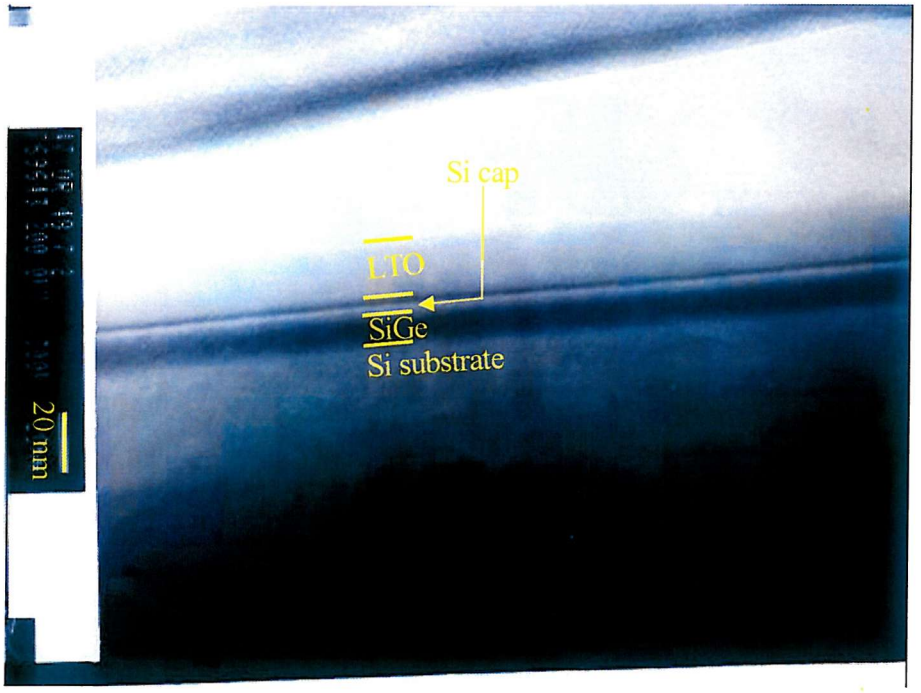


Figure 7.13: TEM photograph of the sample used for the cleaning experiment

and a further RCA like the processed wafers after MBE growth and before the deposition of the complete field oxide. The further cleans could not have had an effect on the substrate underneath the thick field oxide. TEM analysis confirmed the existence of a Si cap on top of the Si channel after the clean (see fig. 7.13). Only a mistake during growth can explain the non-existence of the SiGe heterostructure.

7.4 Conclusion

Due to an error in MBE growth, none of the wafers possessed the expected SiGe heterostructures. This raises the question whether the monitoring of the MBE growth should be improved.

Furthermore, this batch highlights how difficult the preservation of the SiGe heterostructure during cleaning is. As yield is not the critical issue, it is suggested to omit the SC-1 and to use solely the SC-2 for the cleaning of the exposed wafer surface after MBE growth for future SiGe batches.

Bibliography

[1] R. M. Sidek, U. N. Straube, A. M. Waite, A. G. R. Evans, C. Parry, P. Phillips, T. E. Whall, and E. H. C. Parker, "SiGe CMOS fabrication using SiGe MBE and anodic/LTO gate oxide", Semicond. Sci. Technol., Vol. 15, pp. 135-138, 2000

[2] D. J. Norris, A. G. Cullis, T. J. Grasby, and E. H. C. Parker, "Investigation of nanoscale Ge segregation in p-channel SiGe/Si field effect transistor structures by electron energy loss imaging", J. Appl. Phys., Vol. 82, No. 12, pp. 7183-7185, 1999

[3] D. J. Godbey, A. H. Krist, K. D. Hobart, and M. E. Twigg, "Selective Removal of $Si_{1-x}Ge_x$ from (100) Si using HNO_3 and HF", J. Electrochem. Soc., Vol. 139, No. 10, pp. 2943-2945, 1992

[4] F. S. Scott, D. S. Miles, D. T. Grider, and J. J. Wortman, "Selective Chemical Etching of Polycrystalline $SiGe$ Alloys with Respect to Si and SiO_2", J. Electronic Mat., Vol. 21, No. 8, pp. 805-810, 1992

[5] C. R. Inomata, H. Ogawa, K. Ishikawa, and S. Fujimura, "Infrared Spectroscopy Study of Chemical Oxides Formed by a Sequence of RCA Standard Cleaning Treatments", J. Electrochem. Soc., Vol. 143, No. 9, pp. 2995-3000, 1996

Chapter 8

SiGe CMOS/Ge^+ Implantation Process

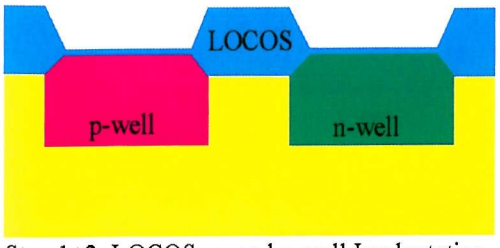
High quality strained SiGe heterostructures can be fabricated by UHCVD and MBE. However, the growth processes are time-consuming and expensive. In contrast, implantation is an industrial standard process and cheap. The species can be implanted into a selected area defined by resist. Although, only low Ge concentrations can be expected by implantation and no abrupt interfaces can be achieved, the above mentioned advantages render Ge^+ implantation a potential candidate for the fabrication of HMOSFETs on an industrial scale [1-7; chap. 6.8.1].

In this chapter, a CMOS process with Ge^+ implanted p-MOSFETs is presented. The process is based on a conventional CMOS process and employs the EPIFAB process developed at the University of Surrey [8-13; chap. 6.8.1]. The parameter extraction routines presented in chapter 4.4 and 4.5 were used in order to assess the MOSFET device performance. The Ge implanted MOSFETs were compared to Si reference MOSFETs.

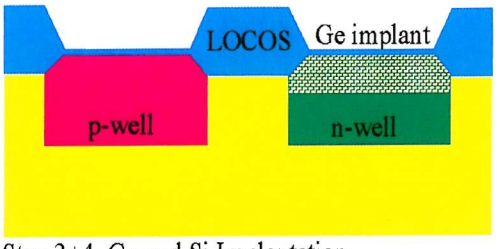
8.1 Fabrication Process

The process is based on an industrial CMOS process and depicted in fig. 8.1. A complete process listing is given in appendix B. The reference Si MOSFETs and SiGe MOSFETs are on the same wafer. The process sequence starts with the growth of a thin thermal oxide (20 nm @ 900^oC) that protects the substrate against contamination. Alignment marks are etched in the 4-inch wafers (p-type, CZ, $17-33\ \Omega cm$ resistivity, orientation 100). Consecutively, the whole oxide is stripped off and regrown. Then, LOCOS is performed for device isolation (400 nm @ 1000^oC) (step 1). In order to ensure a clean substrate, the oxide is stripped off, regrown and stripped off again in the active area. Then, either 20 nm or 10 nm of oxide are grown as a pad oxide (see tab. 8.1).

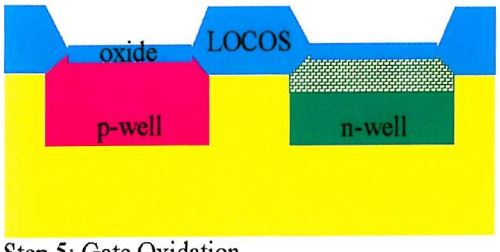
A spilt is made in the p-well and n-well triple implants (step 2) for the two different gate



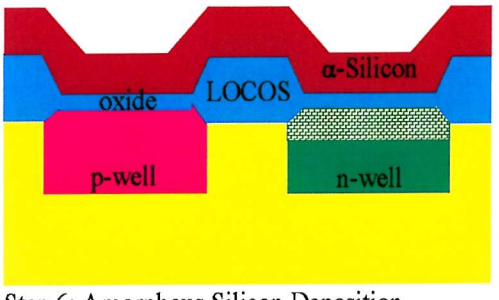
Step 1+2: LOCOS, n- and p-well Implantation



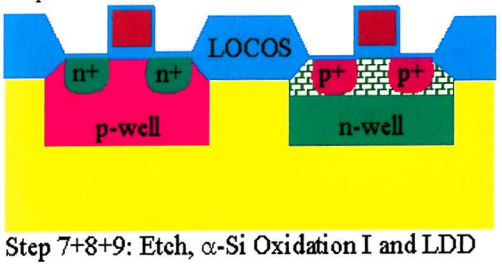
Step 3+4: Ge and Si Implantation



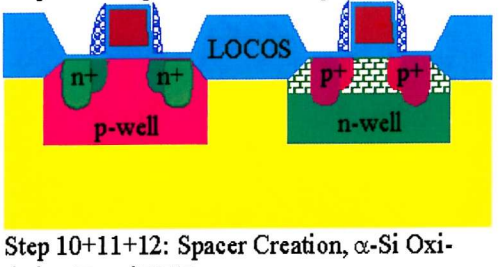
Step 5: Gate Oxidation



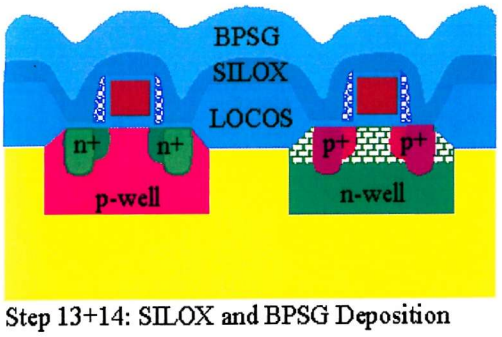
Step 6: Amorphous Silicon Deposition



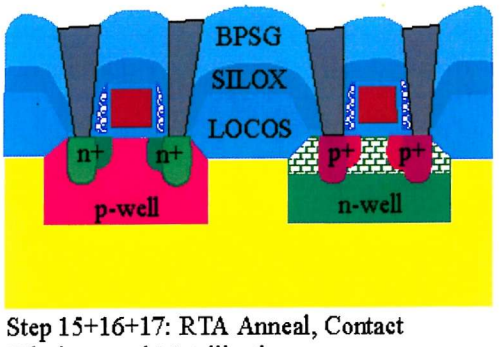
Step 7+8+9: Etch, α-Si Oxidation I and LDD



Step 10+11+12: Spacer Creation, α-Si Oxidation II and HDD



Step 13+14: SILOX and BPSG Deposition



Step 15+16+17: RTA Anneal, Contact Windows and Metallisation

Figure 8.1: Schematic representation of the SiGe CMOS/Ge^+ implantation process

wafer number	pad oxide in nm	Ge conc. in $atoms/cm^2$	gate oxide in nm
1	10	$5 \cdot 10^{15}$	12
2	10	$5 \cdot 10^{15}$	6
3	10	$1 \cdot 10^{16}$	12
4	10	$1 \cdot 10^{16}$	6
5	10	$2 \cdot 10^{16}$	12
6	10	$2 \cdot 10^{16}$	6
7	20	$5 \cdot 10^{15}$	12
8	20	$5 \cdot 10^{15}$	6
9	20	$1 \cdot 10^{16}$	12
10	20	$1 \cdot 10^{16}$	6
11	20	$2 \cdot 10^{16}$	12
12	20	$2 \cdot 10^{16}$	6

Table 8.1: Splits for the SiGe CMOS/Ge^+ implantation batch

oxide thicknesses. The implant doses and energies for the 6 nm gate oxides are overtaken from a conventional process (batch k1543s). They satisfy threshold voltage requirements and are adapted for the 12 nm gate oxide devices [1]:

- 6 nm gate oxide:
  - p-well: $4.0 \cdot 10^{12}$ B^+ $at./cm^2$ @ 200 keV, $3.0 \cdot 10^{13}$ B^+ $at./cm^2$ @ 120 keV, $4.0 \cdot 10^{12}$ B^+ $at./cm^2$ @ 25 keV
  - n-well: $4.0 \cdot 10^{12}$ P^+ $at./cm^2$ @ 400 keV, $2.0 \cdot 10^{13}$ P^+ $at./cm^2$ @ 280 keV, $8.0 \cdot 10^{12}$ P^+ $at./cm^2$ @ 70 keV
- 12 nm gate oxide:
  - p-well: $2.0 \cdot 10^{12}$ B^+ $at./cm^2$ @ 200 keV, $1.5 \cdot 10^{13}$ B^+ $at./cm^2$ @ 120 keV, $2.0 \cdot 10^{12}$ B^+ $at./cm^2$ @ 25 keV
  - n-well: $2.0 \cdot 10^{12}$ P^+ $at./cm^2$ @ 400 keV, $1.0 \cdot 10^{13}$ P^+ $at./cm^2$ @ 280 keV, $4.0 \cdot 10^{12}$ P^+ $at./cm^2$ @ 70 keV

The EPIFAB process [8-13] developed at the University of Surrey is used in order to create defect free SiGe heterojunctions. Ge^+ is implanted at room temperature and 30 keV into the n-well of each second row whereas the rest of the wafers is protected by resist (step 3). The splits are denoted in tab. 8.1. The chosen implant doses are all within the critical thickness (see section 2.2). However, the experimental values cannot be applied exactly because some of the Ge is removed when the oxide is stripped off and and the Ge concentration increase during gate oxidation. The protected p-MOSFETs serve as Si reference devices.

After the resist strip, the wafers are implanted with Si ($6 \cdot 10^{15}$ Si^+/cm^2 @ 500 keV) while they are cooled with liquid nitrogen (step 4). The Si^+ implant under low temperature conditions

[1] Accidentally, twice the specified dose was used for the 70 keV n-well implant

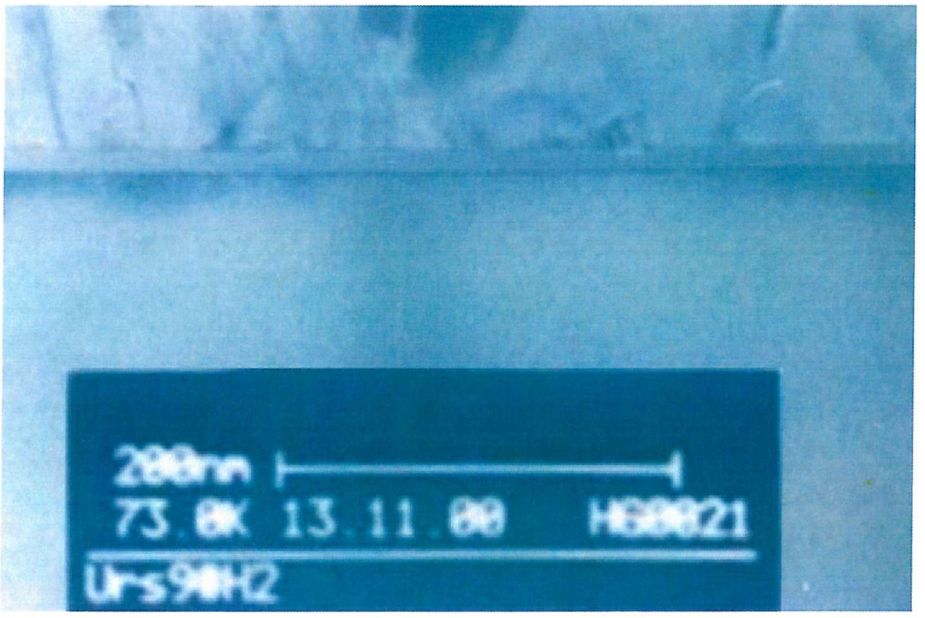


Figure 8.2: TEM photograph underneath the gate region for samples with a 12 nm gate oxide, a 20 nm pad oxide implanted at a Ge^+ dose of $5 \cdot 10^{15}/cm^2$

leads to a well amorphized layer and a good crystal quality of the regrown epilayer. The wafers are not cooled during the Ge^+ implant because the resist would get brittle. Besides, the Si^+ implant drives the end of range defects out of the active device regions about 1 μm deep into the substrate. The TEM photograph of a gate stack shows no defects in the active device region (see fig. 8.2).

Either a 6 nm (800^oC) or 12 nm (900^oC) gate oxide is grown (step 5). The recrystallization of the amorphous layer takes place during the growth of the gate oxide. The gate oxidation is followed by a 15 min. N_2 anneal at the same temperature the gates are grown. Then, 250 nm of amorphous Si is deposited as gate material (step 6) and defined by e-beam lithography (step 7). In the next step, a 6 nm oxide is grown in order to protect the wafer during subsequent processing and anneal the damage due to the etch (step 8). Then, a LDD implant that consists of four implantation steps at the same energy and doping concentration is performed. Between each implantation step, the wafers are rotated by 90^oC (p-MOSFETs: BF_2^+, $4 \times 2.5 \cdot 10^{13}$ $at./cm^2$ @ 33 keV; n-MOSFETs: P^+, $4 \times 4.0 \cdot 10^{13}$ $at./cm^2$ @ 50 keV, step 9) to ensure that the source and drain are uniformly implanted.

The creation of the LTO spacers (step 10) includes an etch step and, hence, requires a further oxidation stage (step 11). Then, the HDD implant is performed (p-MOSFETs: $5.0 \cdot 10^{15}$ BF_2^+/cm^2 @ 45 keV, n-MOSFETs: $5.0 \cdot 10^{15}$ As^+/cm^2 @ 60 keV, step 12). After the gate, source, and drain implant, an anneal (20 min. at 650^oC) removes the fluorine.

100 nm of SILOX and 500 nm of BPSG are deposited (step 13) for planarization of the wafer surface. SILOX is undoped and does therefore not contaminate the substrate, whereas BPSG

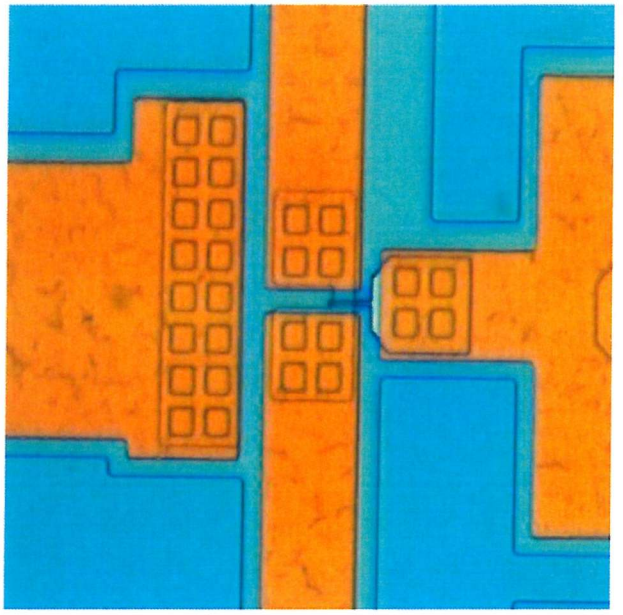


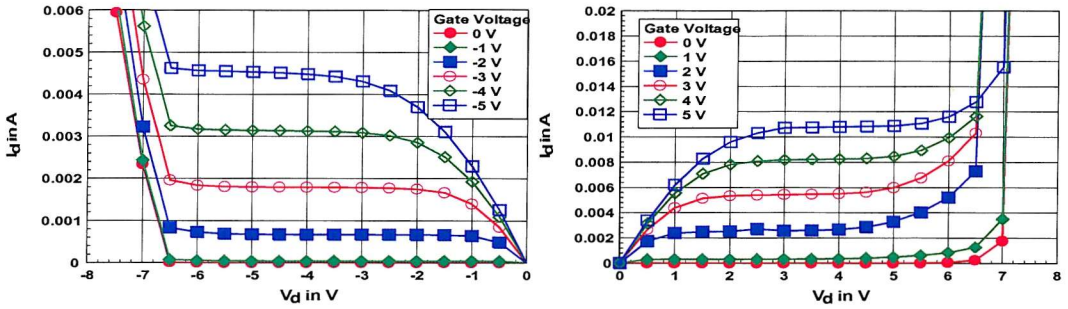
Figure 8.3: MOSFET with $L = 1\ \mu m$ and $W = 5\ \mu m$

is easy to reflow and smoothes the surface. In an RTA (temperature ramped to 1050^oC within 20 s and then kept at this temperature for another 20 s, step 14), the gate, source, and drain dopants are activated and the SILOX and BPSG are reflowed. Then, the contact windows are dry etched (step 15). A dip etch in BHF before the metallization stage is used in order to make good source, drain and gate contacts.

25 nm of Ti, 75 nm of TiN, 1 μm of Al, and 30 nm TiN are deposited and defined (step 16). The Ti layer reduces the contact resistance between metal and semiconductor. The first TIN layer prevents the spiking of Si and the second reflections during optical lithography. The metallic wires are created by dry etching the aluminium and titanium layer. Finally, the wafers are annealed at 420^oC for 30 min. in an H_2/N_2 ambient. Several cleaning steps are implemented when ever necessary.

8.2 Results

Six identical n- and p-MOSFETs with a channel width of 20 μm and gate lengths of either 32, 16, 8, or 1 μm were chosen for parameter extraction. The drain voltage was fixed at $-50\ mV$ for the p-MOSFETs and 50 mV for the n-MOSFETs. The gate voltage was varied between 0 and -3 V for the p-MOSFETs and between 0 and 3 V for the n-MOSFETs. Additionally, the n- and p-MOSFETs with $L = 1\ \mu m$ were respectively measured for $V_d = 2.5\ V$ and $V_d = -2.5\ V$. The parameters extracted with the parameter extraction routine described in chap. 4.5 are given in appendix C. The given error is one standard deviation.



(a) (b)

Figure 8.4: Device characteristics for MOSFET without Ge implant, with $L = 1\ \mu m$ and $W = 20\ \mu m$ on wafer # 2 (6 nm gate oxide) (a) p-MOSFET (b) n-MOSFET

Wafer No.	R_{sd} in $k\Omega\mu m$	ΔL in μm	Θ_o in $1/V$	μ_o in cm^2/Vs
6 nm gate oxide				
2	0.89 ± 0.04	0.30 ± 0.01	0.171 ± 0.001	76.14 ± 0.38
4	0.94 ± 0.10	0.20 ± 0.00	0.184 ± 0.001	80.75 ± 0.14
6	0.79 ± 0.07	0.19 ± 0.01	0.169 ± 0.001	76.22 ± 0.14
8	0.94 ± 0.07	0.40 ± 0.03	0.162 ± 0.001	73.79 ± 0.20
10	1.00 ± 0.03	0.23 ± 0.01	0.171 ± 0.001	76.65 ± 0.21
12	1.42 ± 0.06	0.21 ± 0.01	0.140 ± 0.001	66.49 ± 0.41
12 nm gate oxide				
1	1.37 ± 0.10	0.23 ± 0.01	0.102 ± 0.000	96.65 ± 0.18
3	1.80 ± 0.26	0.29 ± 0.02	0.096 ± 0.002	94.12 ± 0.86
5	1.46 ± 0.17	0.39 ± 0.02	0.105 ± 0.001	98.74 ± 0.24
7	1.17 ± 0.06	0.22 ± 0.01	0.089 ± 0.002	94.48 ± 0.80
9	1.24 ± 0.14	0.28 ± 0.02	0.089 ± 0.002	88.96 ± 0.80
11	1.26 ± 0.01	0.23 ± 0.01	0.103 ± 0.000	96.26 ± 0.08

Table 8.2: R_{sd}, ΔL, Θ_o and μ_o for the p-MOSFETs without Ge implant

8.2.1 P-MOSFETs

P-MOSFETs without Ge^+ implant The devices with the same gate oxide thickness distinguish from each other only by the thickness of the pad oxide which was grown before well implantation (see tab. 8.1). As the pad oxide should not have a pronounced effect, devices with the same gate oxide thickness display very similar device characteristics. Slight variations are of a merely random nature.

Devices with a 6 nm gate oxide The parameters extracted for various L are given in tab. 8.2. R_{sd}, ΔL, Θ_o, and μ_o range respectively from 0.79 to 1.42 $k\Omega\mu m$, 0.19 to 0.40 μm, 0.140 to 0.184 V^{-1}, and 66.49 to 80.75 cm^2/Vs.

For $L = 1\ \mu m$ (see tab. 8.3), I_{on} ($V_g = V_d = -2.5\ V$), I_{off} ($V_g = 0\ V; V_d = -2.5\ V$), S, and

wafer number	I_{on} in 10^{-5} $A/\mu m$	I_{off} in 10^{-11} $A/\mu m$	S in mV/dec	V_{th} in V
6 nm gate oxide				
2	4.16 ± 0.38	70.6 ± 17.3	135.4 ± 4.1	-0.913 ± 0.016
4	3.89 ± 0.39	4.82 ± 0.58	134.7 ± 7.3	-0.858 ± 0.037
6	2.98 ± 0.27	3.31 ± 1.31	145.3 ± 2.7	-1.024 ± 0.020
8	3.86 ± 0.17	5.78 ± 1.50	123.4 ± 0.4	-0.928 ± 0.003
10	4.06 ± 0.06	4.19 ± 0.62	141.1 ± 2.7	-0.798 ± 0.003
12	2.91 ± 0.04	8.60 ± 1.32	129.2 ± 0.2	-1.006 ± 0.008
12 nm gate oxide				
1	2.12 ± 0.09	5.67 ± 1.62	139.4 ± 7.8	-1.002 ± 0.010
3	1.74 ± 0.15	4.39 ± 2.09	136.2 ± 3.8	-1.197 ± 0.048
5	1.81 ± 0.27	16.6 ± 2.8	137.0 ± 3.1	-1.168 ± 0.007
7	1.73 ± 0.03	3.02 ± 0.70	140.1 ± 1.9	-1.191 ± 0.007
9	1.72 ± 0.06	2.69 ± 0.45	141.1 ± 2.7	-1.237 ± 0.010
11	1.98 ± 0.07	2.01 ± 0.23	129.2 ± 0.2	-1.073 ± 0.002

Table 8.3: Parameters for the p-MOSFETs without Ge implant and with $L = 1$ μm

V_{th} range respectively from 29.1 to 41.6 $\mu A/\mu m$, 33.1 to 706 $pA/\mu m$, 123.4 to 141.1 $mV/dec.$, and -0.798 to -1.024 V. The drain currents and transconductances as a function of V_g for $V_d = -2.5$ V and $L = 1$ μm are given in fig. 8.5. The drain current as a function of the drain voltage is shown in fig. 8.4 (a) for a device with $L = 1$ μm and different gate voltages.

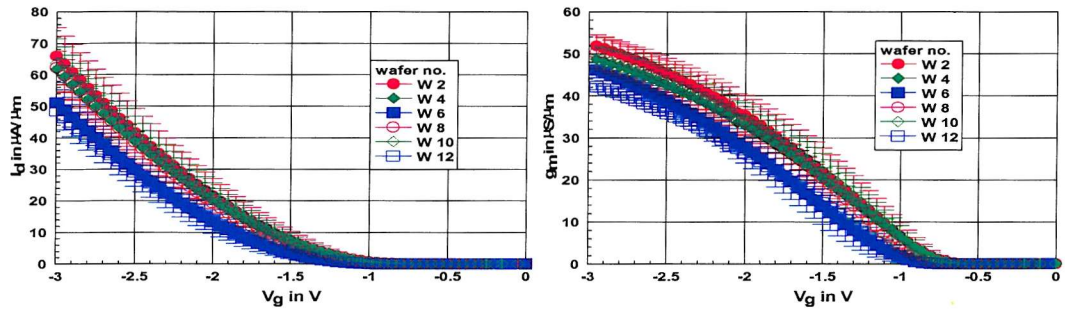
The p-MOSFETs are slightly superior to the ones of a similar process (batch k1543s, # 10) for which the following parameters have been found: $R_{sd} = 2.64 \pm 0.1$ $k\Omega\mu m$, $\Delta L = 0.00 \pm 0.01$ μm, $\Theta_o = 0.170 \pm 0.002$ V^{-1}, $\mu_o = 69.40 \pm 0.50$ cm^2/Vs, $S = 119.9 \pm 0.8$ mV/dec, $I_{on} = 21.8 \pm 3.6 \cdot$ $\mu A/\mu m$, $I_{off} = 1.65 \pm 2.52$ $pA/\mu m$, and $V_{th} = -0.913 \pm 0.008$ V.

Devices with a 12 nm gate oxide The parameters extracted for various L are given in tab. 8.2. R_{sd}, ΔL, Θ_o, and μ_o, range respectively from 1.17 to 1.80 $k\Omega\mu m$, 0.23 to 0.39 μm, 0.089 to 0.105 V^{-1}, and 88.56 to 98.74 cm^2/Vs.

For $L = 1$ μm (see tab. 8.3), I_{on} ($V_g = V_d = -2.5$ V), I_{off} ($V_g = 0$ V; $V_d = -2.5$ V), S, and V_{th} range respectively from 17.2 to 21.2 $\mu A/\mu m$, 20.1 to 166 $pA/\mu m$, 129.2 to 141.1 $mV/dec.$, and -1.002 to -1.237 V. The drain currents and transconductances as a function of V_g for $V_d = -2.5$ V and $L = 1$ μm are given in fig. 8.6.

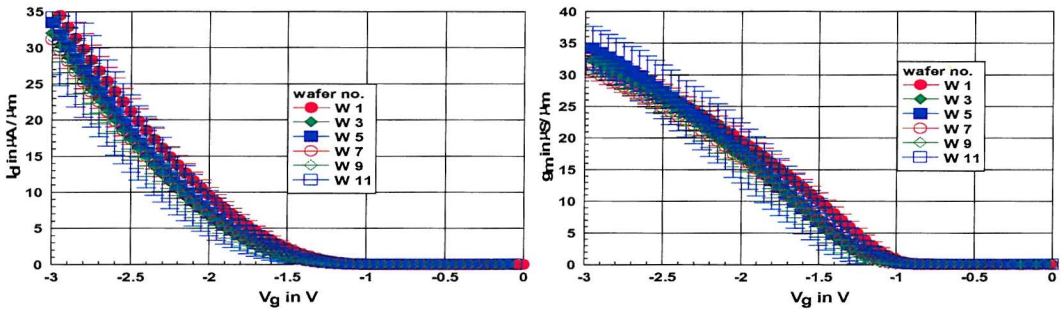
Compared to the split with the 6 nm oxide, the low field mobility and the magnitude of the threshold are higher, the mobility degradation factor and the on state current are lower. The higher mobility is due to a lower n-well dopant concentration and a smaller systematic error in the mobility extraction because of the neglect of the capacitance associated with the inversion layer. The higher threshold voltage is a consequence of the thicker gate oxide.

The mobility degradation factor typically decreases with increasing oxide thickness because



(a) (b)

Figure 8.5: P-MOSFETs with a 6 nm gate oxide and without Ge ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m



(a) (b)

Figure 8.6: P-MOSFETs with a 12 nm gate oxide and without Ge ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m

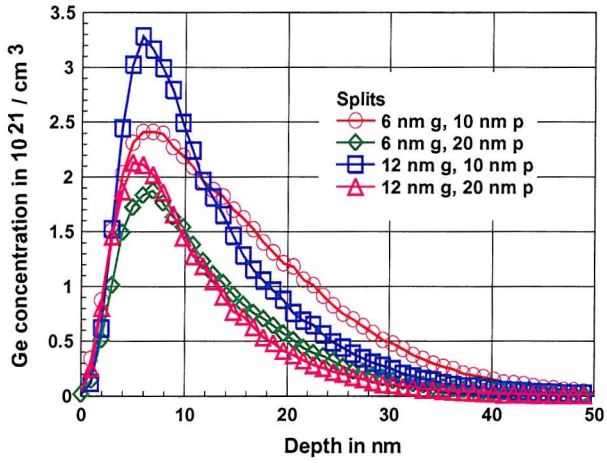


Figure 8.7: Germanium profile for the dose of $5 \cdot 10^{15} at./cm^2$

	Interface states in $10^{11}/cm^2$			
gate oxide	12 nm		6 nm	
pad oxide	10 nm	20 nm	10 nm	20 nm
Ge^+ dose in $at./cm^2$				
Si ref.	1.25 ± 0.09	0.95 ± 0.12	0.84 ± 0.01	1.20 ± 0.04
$5 \cdot 10^{15}$	3.07 ± 0.02	3.13 ± 0.07	4.04 ± 0.19	2.63 ± 0.04
$1 \cdot 10^{16}$	6.65 ± 0.34	6.84 ± 1.14	14.10 ± 0.20	9.35 ± 0.30
$2 \cdot 10^{16}$	22.60 ± 0.39	11.18 ± 0.14	40.70 ± 0.31	31.1 ± 0.37

Table 8.4: Number of interface states per cm^2

Ge^+ dose in $at./cm^2$	R_{sd} in $k\Omega\mu m$	ΔL in μm	Θ_o in $1/V$	μ_o in cm^2/Vs
6 nm gate oxide				
	10 nm pad oxide			
$5 \cdot 10^{15}$	0.90 ± 0.08	0.26 ± 0.02	0.117 ± 0.001	68.47 ± 0.36
$1 \cdot 10^{16}$	0.87 ± 0.14	0.10 ± 0.01	0.077 ± 0.002	59.36 ± 0.58
$2 \cdot 10^{16}$	15.61 ± 1.71	0.11 ± 0.06	-0.171 ± 0.004	14.23 ± 0.45
	20 nm pad oxide			
$5 \cdot 10^{15}$	0.80 ± 0.05	0.24 ± 0.01	0.112 ± 0.001	64.87 ± 0.19
$1 \cdot 10^{16}$	0.72 ± 0.14	0.15 ± 0.02	0.072 ± 0.002	58.08 ± 0.54
$2 \cdot 10^{16}$	18.96 ± 0.99	0.29 ± 0.03	-0.275 ± 0.006	13.30 ± 00.41
12 nm gate oxide				
	10 nm pad oxide			
$5 \cdot 10^{15}$	1.05 ± 0.02	0.20 ± 0.01	0.064 ± 0.000	87.12 ± 0.24
$1 \cdot 10^{16}$	1.34 ± 0.09	0.23 ± 0.01	0.049 ± 0.001	83.54 ± 0.36
$2 \cdot 10^{16}$	1.89 ± 0.57	0.26 ± 0.04	-0.174 ± 0.002	27.02 ± 0.62
	20 nm pad oxide			
$5 \cdot 10^{15}$	2.51 ± 0.36	-0.08 ± 0.05	0.061 ± 0.001	82.34 ± 0.34
$1 \cdot 10^{16}$	1.04 ± 0.06	0.20 ± 0.01	0.034 ± 0.001	73.64 ± 0.16
$2 \cdot 10^{16}$	1.23 ± 0.25	0.20 ± 0.01	-0.056 ± 0.003	49.82 ± 0.77

Table 8.5: R_{sd}, ΔL, Θ_o and μ_o for the Ge implanted p-MOSFETs

mobility degradation depends on the variation of the electrical field. [2] The electrical field varies less for thick than for thin gate oxides as a function of the gate voltage. The on state current is lower due to the thicker gate oxide and the higher magnitude of the threshold voltage.

Ge^+ implanted p-MOSFETs In fig. 8.11 - 8.13, device parameters of the Ge^+ implanted p-MOSFETs are displayed as a function of the Ge^+ dose. The interface state density has been extracted from the peak of the Elliot charge pumping curve (see tab. 8.4) measured at 100 kHz for $\Delta V_A = 4\ V$. The values are subject to the restrictions discussed in chap. 4.3. The

[2] If the space charge is neglected, the variation of the gate voltage is given by $\Delta V_g = -\Delta E_{SiO_2} \cdot t_{SiO_2}$. Hence, the variation of the electrical field is proportional to the inverse of the oxide thickness for a given variation of the gate voltage ($\Delta E_{SiO_2} \propto 1/\cdot t_{SiO_2}$). The mobility can be modeled by $\mu = \mu_o/(1 + E_{SiO_2}/E_c)$ and be expressed as $\mu = \mu_o/(1 + \Delta V/(t_{SiO_2} \cdot E_c)) = \mu_o/(1 + \Theta_o \Delta V)$. Where E_c is a constant. The mobility degradation factor $\Theta_o = 1/(t_{SiO_2} \cdot E_c)$ decreases with the oxide thickness.

Ge^+ dose in $at./cm^2$	I_{on} in $10^{-5}A/\mu m$	I_{off} in $10^{-11}A/\mu m$	S in mV/dec	V_{th} in V
6 nm gate oxide				
	10 nm pad oxide			
$5 \cdot 10^{15}$	3.59 ± 0.20	2.12 ± 1.23	165.6 ± 2.0	-1.000 ± 0.027
$1 \cdot 10^{16}$	1.93 ± 0.12	4.58 ± 1.17	232.4 ± 9.7	-1.248 ± 0.034
$2 \cdot 10^{16}$	0.16 ± 0.02	0.88 ± 0.08	372.3 ± 2.0	-1.762 ± 0.008
	20 nm pad oxide			
$5 \cdot 10^{15}$	2.89 ± 0.07	3.30 ± 0.68	150.7 ± 2.0	-1.056 ± 0.025
$1 \cdot 10^{16}$	1.87 ± 0.13	5.07 ± 1.28	216.9 ± 5.0	-1.164 ± 0.017
$2 \cdot 10^{16}$	0.14 ± 0.01	7.35 ± 0.56	473.0 ± 7.8	-1.862 ± 0.007
12 nm gate oxide				
	10 nm pad oxide			
$5 \cdot 10^{15}$	1.63 ± 0.15	4.00 ± 3.23	160.0 ± 0.5	-1.125 ± 0.004
$1 \cdot 10^{16}$	1.10 ± 0.17	0.35 ± 0.11	211.6 ± 7.6	-1.378 ± 0.029
$2 \cdot 10^{16}$	0.10 ± 0.02	1.58 ± 0.37	380.8 ± 7.7	-1.946 ± 0.005
	20 nm pad oxide			
$5 \cdot 10^{15}$	1.13 ± 0.02	4.06 ± 1.39	171.6 ± 0.5	-1.375 ± 0.003
$1 \cdot 10^{16}$	5.58 ± 0.60	0.05 ± 0.00	272.4 ± 4.8	-1.653 ± 0.012
$2 \cdot 10^{16}$	0.23 ± 0.02	2.54 ± 0.58	380.4 ± 2.4	-1.751 ± 0.009

Table 8.6: Parameters for the Ge implanted p-MOSFETs with $L = 1\ \mu m$

parameters for the MOSFETs without Ge are extracted on the MOSFETs with the lowest Ge^+ dose. The parameters extracted for various L are given in tab. 8.5. There is no remarkable difference between the devices which had a 10 nm thick and a 20 nm thick pad oxide.

The source and drain series resistance is in general independent of the Ge^+ dose. The mobility degradation factors decreases slightly for the devices with the Ge^+ doses of 5e15 $at./cm^2$ and 1e16 $at./cm^2$ (see fig. 8.12 (b)), and gets even negative for the devices with a Ge^+ dose of 2e16 $at./cm^2$. The trapping of carriers in the channel region might be responsible for the physically unacceptable negative mobility degradation factor. More carriers are trapped at a small magnitude of the gate drive than at a high magnitude of the gate drive because the variation of the surface potential is more important for a small magnitude of the gate drive than for a large magnitude of the gate drive. The number of interface traps increases with the Ge^+ dose (see tab. 8.4, fig.'s 8.19 and 8.16).

The low-field mobility decreases as well slightly for the Ge^+ doses of 5e15 $at./cm^2$ and 1e16 $at./cm^2$ (see fig. 8.11 (a)), whereas there is a considerable mobility reduction for a Ge^+ dose of 2e16 $at./cm^2$. Interface traps can as well be blamed for the decrease of the extracted mobility. They contribute to Coulomb scattering and reduces the number of carriers. The extracted mobility does not necessarily correspond to the physical mobility as the number of interface traps becomes actually comparable to the number of expected carriers for a high Ge^+ dose (see tab. 8.4).

Besides, a slight increase of the oxide thickness has been found for the Ge^+ implanted samples

by quasistatic CV measurements for MOSFETs with $L = 6\ \mu m$ and $W = 1000\ \mu m$ on wafer # 8 and wafer # 10 which have a nominally 6 nm gate oxide as it can be seen from the gate-to-channel capacitance C_{gc}. The split-CV equipment was found to be close to the detection limit for all MOSFETs and could be used only for the wafers with a 6 nm thick gate oxide. An increase of the oxide thickness with the Ge^+ dose reduces the extracted mobility further as the nominal value (6 nm) for the oxide thickness is used in the calculations. The mobility extracted with the routine described in chap. 4.5 is shown in fig. 8.9 and fig. 8.10.

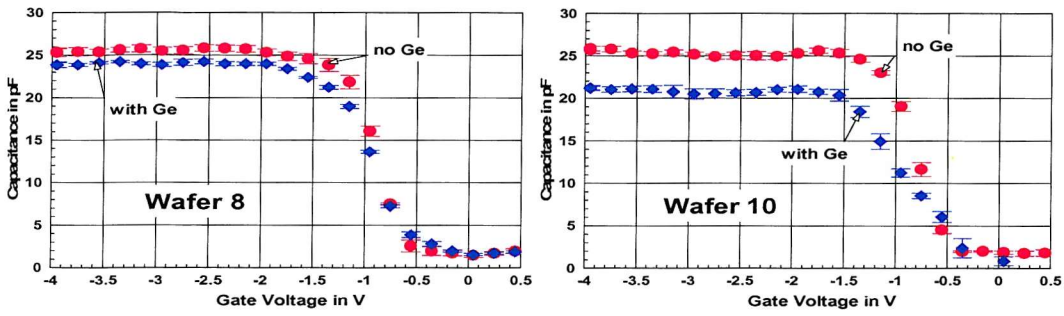
I_{on} ($V_g = V_d = -2.5\ V$), I_{off} ($V_g = 0\ V; V_d = -2.5\ V$), S, and V_{th} are given as a function of the Ge^+ dose for $L = 1\ \mu m$ in tab. 8.6. The magnitude of the threshold voltage increases slightly for the Ge^+ doses of 5e15 $at./cm^2$ and 1e16 $at./cm^2$ and considerably for a Ge^+ dose of 2e16 $at./cm^2$ (see fig. 8.12 (a)) due to a fixed interface charge and an increase of the oxide thickness. Whereas I_{off} is largely independent of the Ge^+ dose, I_{on} is the parameter which is most sensitive to the Ge^+ dose because it is influenced by the mobility reduction and the increase of the magnitude of the threshold voltage as a function of the Ge^+ dose (see fig. 8.13(a)). The subthreshold slope (see fig. 8.13 (b)) increases as a function of the Ge^+ implant because the interface state density increases.

The drain currents and transconductances for $V_d = -2.5\ V$ as a function of the gate voltage are displayed in fig. 8.14 and fig. 8.15 for the devices with a 6 nm gate oxide and in fig. 8.17 and fig. 8.18 for the devices with a 12 nm gate oxide and compared to the Si reference devices on the wafer with the lowest Ge^+ implant.

The Ge profile was measured on samples with all different gate and pad oxide thicknesses and a Ge^+ dose of $5 \cdot 10^{15}\ at./cm^2$ by SIMS (see fig. 8.7). The approximate position of the oxide interface is $x = 0$. The exact position can not be determined because SIMS broadens the correct profile due to the limited space resolution. The influence of the Si cap and gate oxide thickness manifests in the width and height of the Ge profile. A thicker oxide and thicker cap layer reduce the width of the Ge profile. An increase of the oxide thickness results in a higher peak Ge concentration due to Ge segregation. The peak values are in general lower for a 20 nm pad oxide than for a 10 nm pad oxide. It is important to note that the real Ge peak concentrations may be considerably higher, but are broadened and reduced in height because they are not wide.

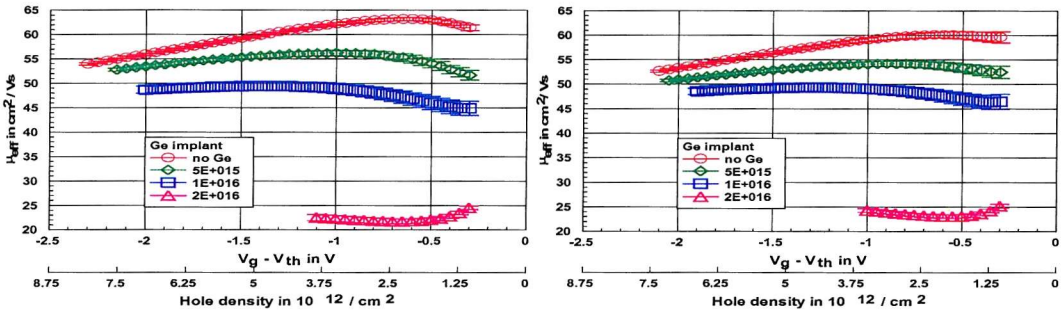
Devices with a 6 nm gate oxide The source and drain series resistance (see fig. 8.11 (b)) increases suddenly for the devices with a Ge^+ dose of $2 \cdot 10^{16}\ at./cm^2$. However, the high source and drain resistance might be the consequence of a systematic error due to parameter extraction because the equation given in chap. 4.4 may not be appropriate to fit the drain current of the devices with a Ge^+ dose of $2 \cdot 10^{16}\ at./cm^2$.

Devices with a 12 nm gate oxide Except for R_{sd} which is not particularly higher for a Ge^+ dose of $2 \cdot 10^{16}\ at./cm^2$ than for the two other Ge^+ doses, the results for the Ge^+ implanted



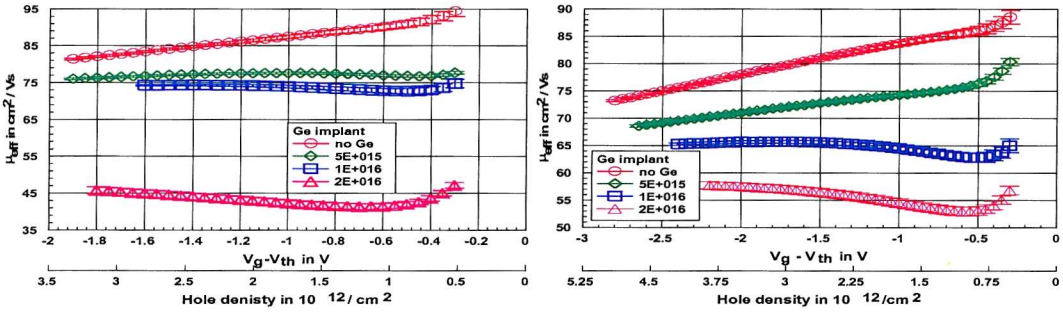
(a) (b)

Figure 8.8: Comparison of C_{cg} measured quasi-static without Ge (red) and with Ge (blue) ($L = 6\ \mu m$, $W = 1000\ \mu m$, $t_{ox} = 6\ nm$) (a) $5 \cdot 10^{15}/cm^2$ (b) $1 \cdot 10^{16}/cm^2$



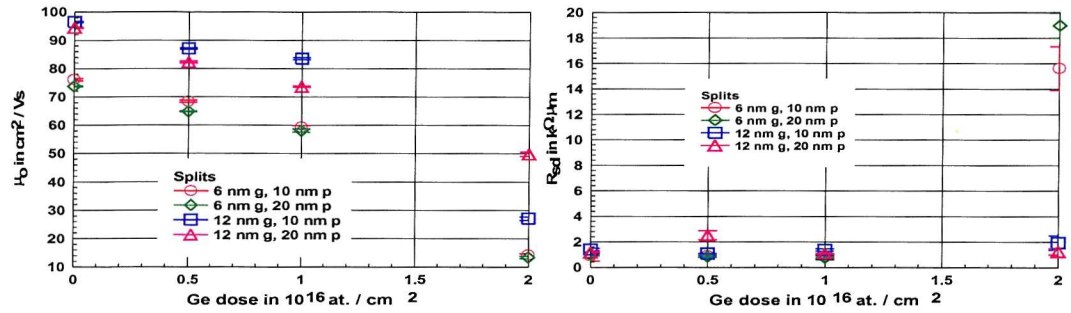
(a) (b)

Figure 8.9: Mobility for p-MOSFETs with a 6 nm gate oxide (a) 10 nm pad oxide (b) 20 nm pad oxide



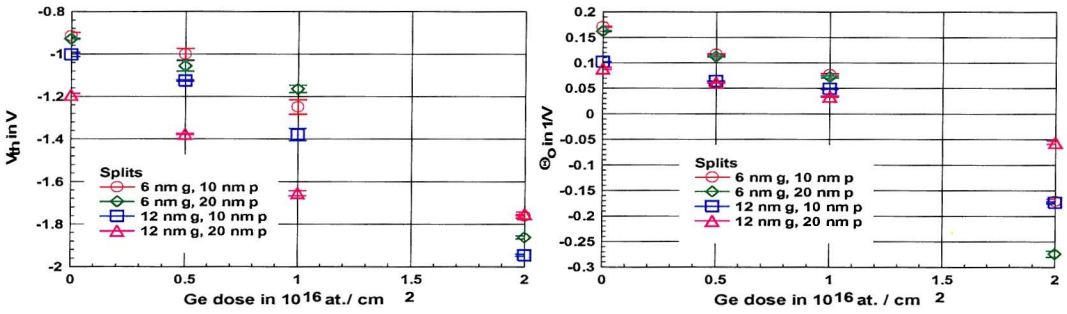
(a) (b)

Figure 8.10: Mobility for p-MOSFETs with a 12 nm gate oxide (a) 10 nm pad oxide (b) 20 nm pad oxide



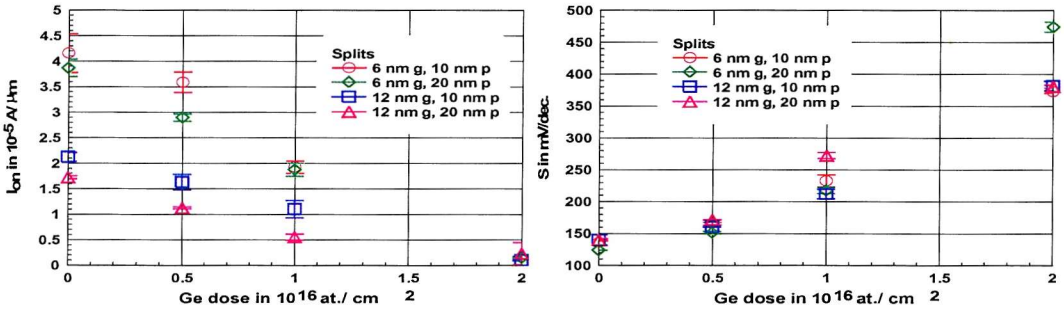
(a) (b)

Figure 8.11: Parameters as a function of the Ge implant (a) μ_o (b) R_{sd}



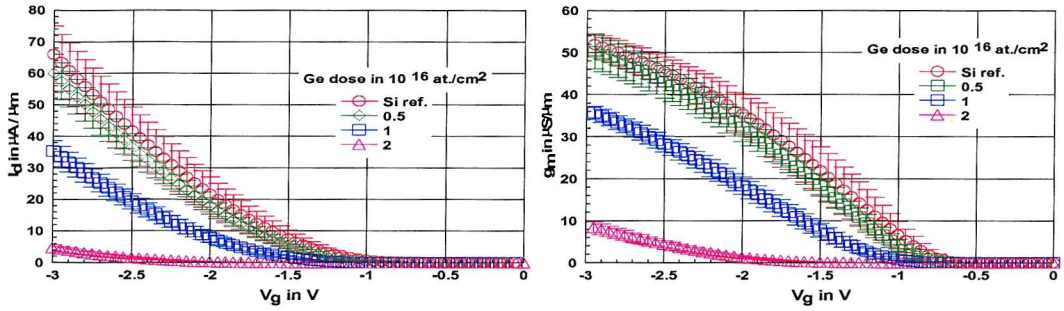
(a) (b)

Figure 8.12: Parameters as a function of the Ge implant (a) V_{th} ($L = 1\ \mu m$) (b) Θ_o



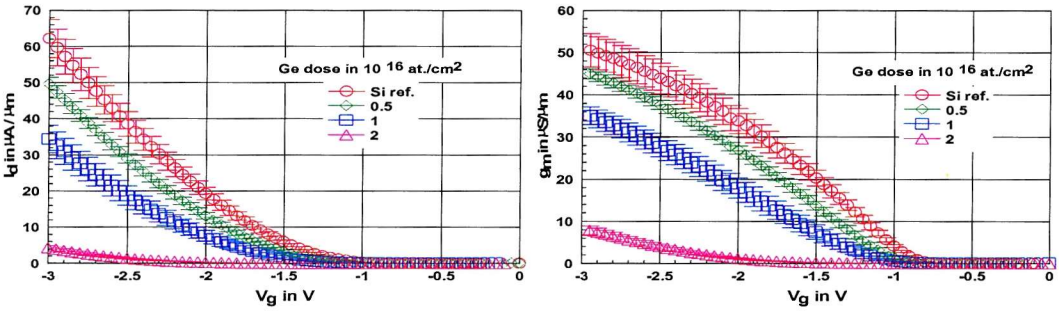
(a) (b)

Figure 8.13: Parameters as a function of the Ge implant for $L = 1\ \mu m$ (a) I_{on} (b) S



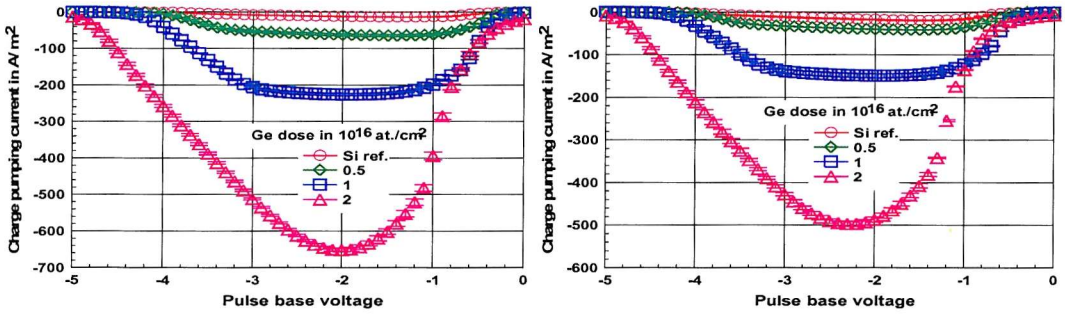
(a) (b)

Figure 8.14: P-MOSFETs with a 6 nm gate oxide and a 10 nm pad oxide ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m



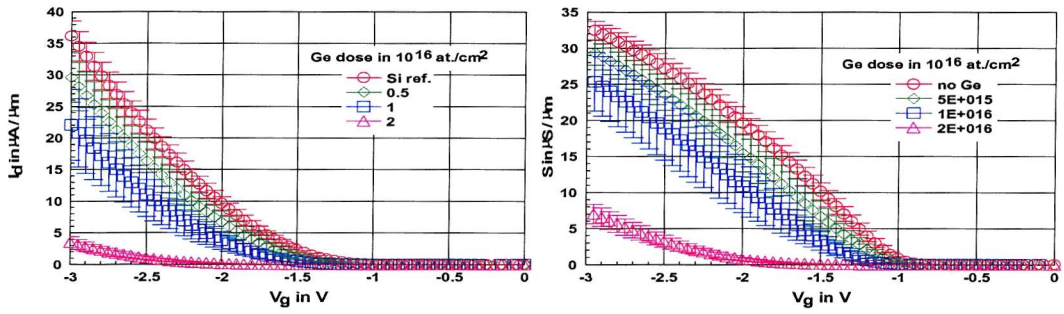
(a) (b)

Figure 8.15: P-MOSFETs with a 6 nm gate oxide and a 20 nm pad oxide ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m



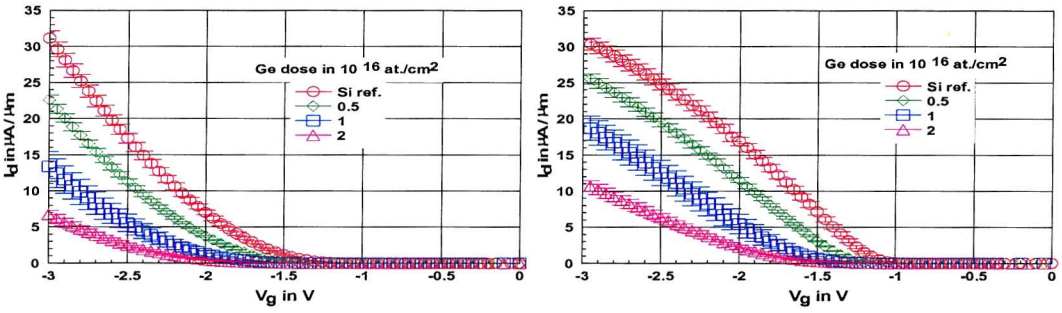
(a) (b)

Figure 8.16: I_{cp} for p-MOSFETs with a 6 nm gate oxide measured at 100 kHz and for $\Delta V_A = 4\ V$ (a) 10 nm pad oxide (b) 20 nm pad oxide



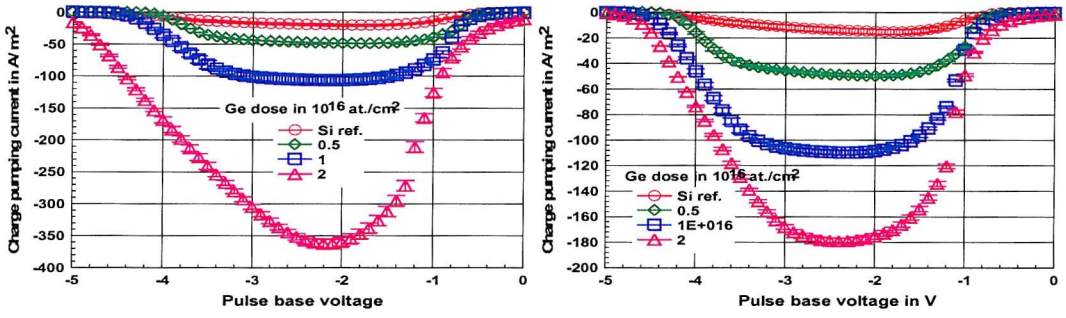
(a) (b)

Figure 8.17: P-MOSFETs with a 12 nm gate oxide and a 10 nm pad oxide ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m



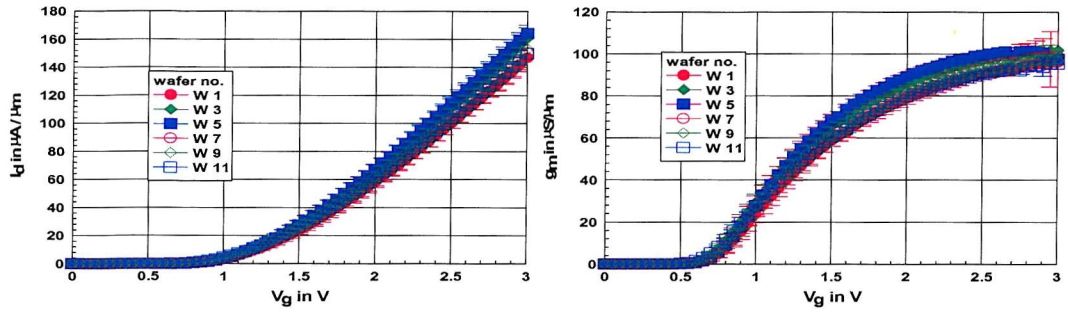
(a) (b)

Figure 8.18: P-MOSFETs with a 12 nm gate oxide and a 20 nm pad oxide ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m



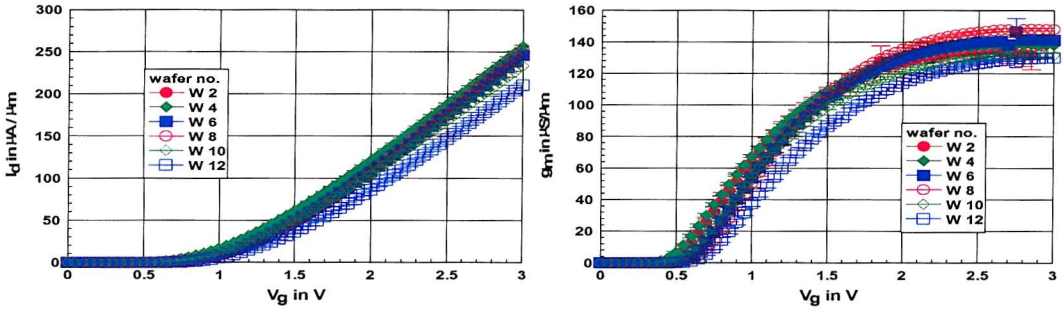
(a) (b)

Figure 8.19: I_{cp} for p-MOSFETs with a 12 nm gate oxide measured at 100 kHz and for $\Delta V_A = 4\ V$ (a) 10 nm pad oxide (b) 20 nm pad oxide



(a) (b)

Figure 8.20: N-MOSFETs with a 12 nm gate oxide ($L = 1, V_d = 2.5\ V$) (a) I_d (b) g_m



(a) (b)

Figure 8.21: N-MOSFETs with a 6 nm gate oxide ($L = 1, V_d = -2.5\ V$) (a) I_d (b) g_m

MOSFETs with a 12 nm gate oxide show the same dependence on the Ge^+ dose as the Ge^+ implanted MOSFET with a 6 nm gate oxide. Compared to the split with the 6 nm gate oxide, the magnitude of the threshold is higher and the on state current lower. The higher threshold voltage is a consequence of the thicker gate oxide. The on state current is lower due to the thicker gate oxide and the higher magnitude of the threshold voltage.

8.2.2 N-MOSFETs

The devices with the same gate oxide thickness are distinguished from each other only by the thickness of the pad oxide which was grown before well implantation (see tab. 8.1). As the pad oxide should not have a pronounced effect, devices with the same gate oxide thickness display very similar device characteristics. Slight variations are of a merely random nature. The reason for the low performance of wafer no. 12 is not obvious.

Devices with a 6 nm gate oxide For $L = 1\ \mu m$ (see tab. 8.8), I_{on} ($V_g = V_d = 2.5\ V$), I_{off} ($V_g = 0\ V; V_d = 2.5\ V$), S, and V_{th} range respectively from 147 to 190 $\mu A/\mu m$, 0.02 to 1.11 $nA/\mu m$, 116.9 to 123.2 $mV/dec.$, and 0.607 to 0.818 V. The drain currents and transconductances as a function of V_g for $V_d = -2.5\ V$ and $L = 1\ \mu m$ are given in fig. 8.21.

Wafer No.	R_{sd} in $k\Omega\mu m$	ΔL in μm	Θ_o in $1/V$	μ_o in cm^2/Vs
6 nm gate oxide				
2	0.52 ± 0.04	0.13 ± 0.03	0.241 ± 0.004	361.54 ± 2.07
4	0.51 ± 0.01	0.08 ± 0.01	0.249 ± 0.000	367.42 ± 0.85
6	0.64 ± 0.01	0.11 ± 0.01	0.243 ± 0.001	363.99 ± 1.06
8	0.59 ± 0.01	0.17 ± 0.01	0.240 ± 0.001	357.60 ± 0.79
10	0.50 ± 0.03	0.06 ± 0.01	0.252 ± 0.002	357.73 ± 1.18
12	0.71 ± 0.05	0.10 ± 0.01	0.198 ± 0.003	313.94 ± 1.66
12 nm gate oxide				
1	0.88 ± 0.04	0.14 ± 0.02	0.063 ± 0.001	379.38 ± 2.58
3	1.13 ± 0.17	0.30 ± 0.26	0.060 ± 0.001	370.67 ± 7.96
5	0.88 ± 0.04	0.31 ± 0.02	0.059 ± 0.001	371.10 ± 3.97
7	0.88 ± 0.04	0.19 ± 0.02	0.061 ± 0.001	374.92 ± 5.16
9	0.95 ± 0.06	0.28 ± 0.03	0.049 ± 0.001	329.21 ± 5.32
11	0.87 ± 0.01	0.17 ± 0.01	0.056 ± 0.000	360.04 ± 0.60

Table 8.7: R_{sd}, ΔL, Θ_o and μ_o for the n-MOSFETs

Wafer No.	I_{on} in $10^{-4}\ A/\mu m$	I_{off} in $10^{-11} A/\mu m$	S in mV/dec	V_{th} in V
6 nm gate oxide				
2	1.81 ± 0.04	2.08 ± 0.05	117.0 ± 4.0	0.652 ± 0.008
4	1.90 ± 0.01	111 ± 43	118.5 ± 1.4	0.607 ± 0.005
6	1.76 ± 0.02	4.91 ± 2.77	117.1 ± 9.1	0.748 ± 0.006
8	1.78 ± 0.01	61.9 ± 31.3	116.9 ± 0.8	0.797 ± 0.007
10	1.68 ± 0.01	61.4 ± 8.7	122.8 ± 0.9	0.678 ± 0.003
12	1.47 ± 0.01	12.2 ± 2.2	123.2 ± 1.1	0.818 ± 0.007
12 nm gate oxide				
1	1.00 ± 0.02	2.08 ± 0.51	123.2 ± 2.1	0.828 ± 0.003
3	1.12 ± 0.03	11.7 ± 6.1	121.4 ± 0.9	0.799 ± 0.016
5	1.14 ± 0.02	8.78 ± 3.30	124.1 ± 1.3	0.850 ± 0.025
7	1.03 ± 0.03	16.4 ± 7.1	125.7 ± 2.4	0.845 ± 0.021
9	1.06 ± 0.02	0.58 ± 0.23	128.5 ± 2.7	0.792 ± 0.009
11	1.04 ± 0.01	51.2 ± 13.6	133.2 ± 1.1	0.770 ± 0.004

Table 8.8: Parameters for the n-MOSFETs with $L = 1\ \mu m$

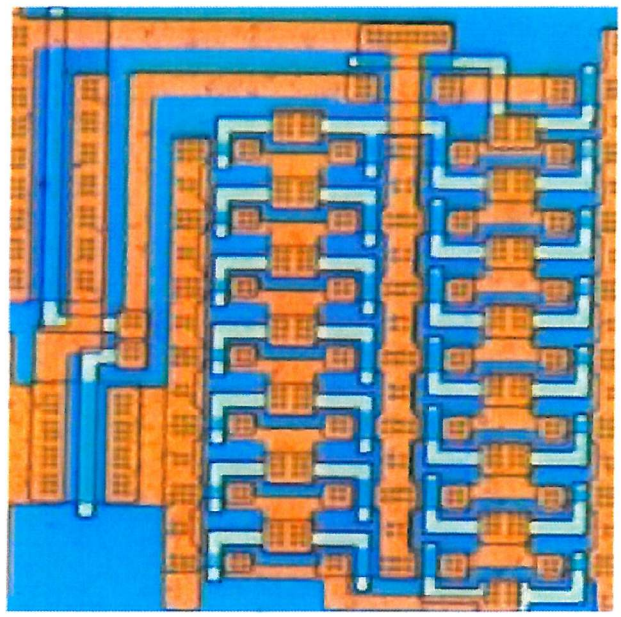


Figure 8.22: Ring oscillator

The n-MOSFETs are slightly superior to the ones of a similar process (batch k1543s, # 10) for which the following parameters have been found: $R_{sd} = 0.95 \pm 0.08\ k\Omega\mu m$, $\Delta L = 0.18 \pm 0.06\ \mu m$, $\Theta_o = 0.225 \pm 0.002\ V^{-1}$, $\mu_o = 307.91 \pm 4.06\ cm^2/Vs$, $I_{on} = 22.7 \pm 4.7\ \mu A/\mu m$, $I_{off} = 0.62 \pm 0.20\ pA/\mu m$, $V_{th} = 1.786 \pm 0.009\ V$, and $S = 236.0 \pm 3.6\ mV/dec$,

Devices with a 12 nm gate oxide The parameters extracted for various L are given in tab. 8.7, R_{sd}, ΔL, Θ_o, and μ_o, range respectively from 0.87 to 1.13 $k\Omega\mu m$, 0.14 to 0.31 μm, 0.049 to 0.63 V^{-1}, and 274.34 to 316.15 cm^2/Vs.

For $L = 1\ \mu m$ (see tab. 8.8), I_{on} ($V_g = V_d = 2.5\ V$), I_{off} ($V_g = 0\ V; V_d = 2.5\ V$), S, and V_{th} range respectively from 100 to 114 $\mu A/\mu m$, 5.80 to 512 $pA/\mu m$, 121.4 to 133.2 $mV/dec.$, and 0.770 to 0.850 V. The drain currents and transconductances as a function of V_g for $V_d = -2.5\ V$ and $L = 1\ \mu m$ are given in fig. 8.20.

Compared to the previous split, the source and drain series resistance, the threshold voltage, and the low field mobility are a bit higher. The mobility degradation factor typically decreases with increasing oxide thickness because mobility degradation depends on the variation of the electrical field. The increase of the threshold voltage can be attributed to the thicker gate oxide and the increase of the mobility to the lower dopant concentration in the substrate. The higher values for the source and drain series resistance will be further discussed in appendix C.

8.2.3 Ring Oscillators

Working ring oscillators could be found only on the chip in the upper left corner for the splits with the 6 nm gate oxide. The ring oscillators themselves are in the upper left corner of the chip.

These ring oscillators consist of 13 CMOS inverters ($L = 4\ \mu m$; $W = 4\ \mu m$ for the n-MOSFETs; $W = 10\ \mu m$ for the p-MOSFETs). For $V_{sup} = 4\ V$, the period was always within a range of a few nanoseconds around 45 ns.

The working of the very few ring oscillators is presumably due to their position. A possible explanation is that other ring oscillators fail because the contact windows are (1.5 μm by 1.5 μm) quite small and the oxide etch rate depends on the window size. The planarization oxide close to the edge of the wafer might be slightly thinner than in the centre. This explains as well why no working ring oscillators have been found for the split with a 12 nm gate oxide.

8.3 Conclusion

The integration of Ge implanted p-MOSFETs into a CMOS process [14] does not encounter any mayor obstacles. SiGe oxidizes at a slightly higher rate than Si. No negative effect of the Ge^+ implant on the quality of the gate oxide grown after Ge^+ implantation was found. A simultaneous performance degradation and increase of the interface state density has been observed as a function of the Ge^+ dose. The observed link between performance degradation and increase of the interface state density suggests that an anneal [14, 15; chap. 6.3] which decreases the interface state density might as well enhance the device performance of the Ge implanted MOSFETs.

Bibliography

[1] H. Jiang and R. G. Elliman, "Electrical Properties of GeSi Surface- and Buried-Channel p-MOSFET's Fabricated by Ge Implantation", IEEE Transactions on Electron Devices, Vol. 43, No. 1, pp. 97-103, 1996

[2] R. G. Elliman and W. C. Wong, "The fabrication of epitaxial Ge_xSi_{1-x} layers by ion implantation", Nuc. Instr. and Meth. in Phys. Res., Vol. B80/81, pp. 768-772, 1993

[3] Z. Xia, "Solid Phase epitaxial regrowth of $Si_{1-x}Ge_x$ layers formed by single-energy Ge^+, double-energy Si^+ and Ge^+, and Ge^+ and Ge^{2+} ion implantation", Mat. Sci. Eng , Vol. B 34, pp. 42-46, 1995

[4] R. G. Elliman, W.-C. Wong and P. Kringhoj, "Solid-Phase Epitaxial Crystallisation of Ge_xSi_{1-x} Alloy Layers", Mat. Res. Soc. Symp. Proc., Vol. 321, pp. 375-385, 1994

[5] K. C. Liu, S. K. Ray, S. K. Oswal and S. K. Banerjee, "A deep submicron $Si_{1-x}Ge_x/Si$ Vertical PMOSFET Fabricated by Ge Ion Implantation", IEEE Electron Dev. Lett., Vol. 19, No. 1, pp. 13-15, 1998

[6] S. Kal, N and B. Chakrabarti, "Strained Silicon-SiGe Devices using Germanium Implantation", IETE J. Res., Vol. 43, No. 2 & 3, pp. 185-192, 1997

[7] Z. Xia, E. O. Ristolainen and P. H. Holloway, "$SiGe/Si$ heterostructures produced by double-energy Si^+ and Ge^+, and Ge^+ and Ge^{2+} ion implantations", J. Vac. Sci. Technol. B, Vol. 16, No. 5, pp. 2629-2632, 1998

[8] F. Christiano, "Extended defects in SiGe device structures formed by ion implantation", Ph.D. Thesis, University of Surrey, 1998

[9] Y. Suprun-Belevich, F. Christiano, A. Nejim, P. L. F. Hemment, and B. J. Sealy, "Evolution of mechanical strain and extended defects in annealed (1 0 0) silicon samples implanted with Ge^+ ions ", Nucl. Instr. and Meth. in Phys. Res. B 140, pp. 91-98, 1998

[10] A. Nejim, N. P. Barradas, C. Jeynes, F. Cristiano, E. Wendler, K. Gaertner, and B. J. Sealy, "Residual post anneal damage of Ge and C co-implantation of Si determined by quantitative RBS-channeling", Nucl. Instr. and Meth. in Phys. Res. B 139, pp. 244-248, 1998

[11] F. Christiano, A. Nejim, B. de Maudit, A. Claverie, P. L. F. Hemment, "Characterization of extended defects in SiGe alloys formed by high dose Ge^+ implantation into Si", Nucl. Instr. and Meth. in Phys. Res. B 120, pp. 156-160, 1996

[12] F. Christiano, A. Nejim, Y. Suprun-Belevich, A. Claverie, and P. L. F. Hemment, "Formation of extended defects and strain relaxation in ion beam synthesised SiGe alloys", Nucl. Instr. and Meth. in Phys. Res. B 147, pp. 35-42, 1999

[13] F. Christiano, C. Bonafos, A. Nejim, S. Lombardo, M. Omri, D. Alquier, A. Martinez, S. U. Campisano, P. L. F. Hemment, and A. Claverie, "Interstitial trapping efficiency of C^+ implanted into preamorphised silicon - control of EOR defects", Nucl. Instr. and Meth. in Phys. Res. B 127 / 128, pp. 22-26, 1997

[14] U. N. Straube, A. M. Waite, A.G.R Evans, A. Nejim, and P. L.F. Hemment, "Adverse effect of Ge^+ implantation for fabrication of SiGe PMOS", Electronics Lett., Vol. 37, No. 25, pp. 1549-1550, 2001

[15] D. Tchikatilov, Y. F. Yang and E. S. Yang, "Improvement of SiGe oxide grown by electron cyclotron resonance using H_2O vapor annealing", Appl. Phys. Lett., Vol. 69, No. 17, pp. 2578-2580, 1996

[16] C.-G. Ahn, H.-S. Kang, Y.-K. Kwon, and B. Kang, "Effects of Segregated Ge on Electrical Properties of $SiO_2/SiGe$ Interface", Jpn. J. Appl. Phys., Vol. 37, pp. 1316-1319, 1998

Chapter 9

SiGe Virtual Substrate/Limited Area PMOS Process

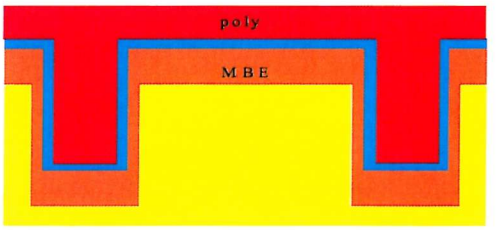
A high Ge content in the strained SiGe channel resulting in a considerable mobility enhancement can be achieved only if grown on top of a virtual substrate [1-6]. As the virtual substrate has to be relaxed, a substantial thickness of the MBE layer is in general required. However, the relaxation of the virtual substrate can be eased up by the growth on microscopic pillars [6-10] . The edges of the pillars are suitable sites for the nucleation of dislocations. The lateral dimensions and orientations are crucial for the relaxation of the virtual substrate (see section 5.2.1).

This chapter deals with a SiGe virtual substrate/limited area PMOS process. Whereas typical virtual substrates are $2-5\ \mu m$ thick, the thickness of the presented virtual substrate batch is only $1.25\ \mu m$. The required process deviates considerably from a standard process because of the trenches which surround the tiny pillars on which the devices are fabricated. The fabrication of MOSFETs on tiny pillars requires an extremely tight mask control. A Si development batch was used in order to investigate the process and masks. The SiGe p-MOSFETs are compared to Si reference p-MOSFETs and p-HMOSFETs which are not fabricated on pillars.

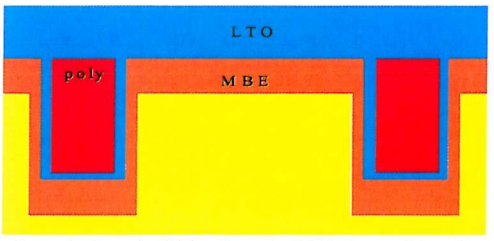
9.1 Fabrication Process

The process for which 4-inch wafers (p-type, CZ, $17-33\ \Omega cm$ resistivity, orientation 100) are used is depicted in fig. 9.1. The complete process listing is given in appendix B. The process sequence starts with the growth of a 1000 nm thick furnace oxide at 1000^oC in which trenches are dry etched while the rest of the wafer is protected by resist. The oxide is used as a mask in the following Si etch (step 1). The trenches cover a few percent of the total chip area. The etch controllability is not good if the exposed area is too small. Large trenches are used in order to increase the size of the area exposed during the etch. Unfortunately, the large trenches fill up with resist in further processing and lead to a nonuniformity of the resist thickness. The

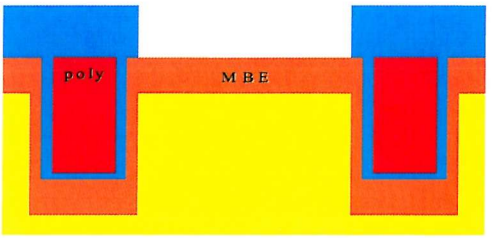
Step 1: Trench Etch



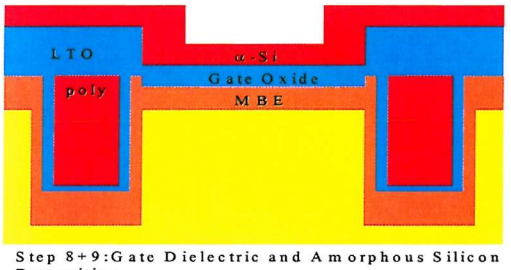
Step 2+3+4: Growth of MBE layer, LTO Deposition I and poly-Si Deposition



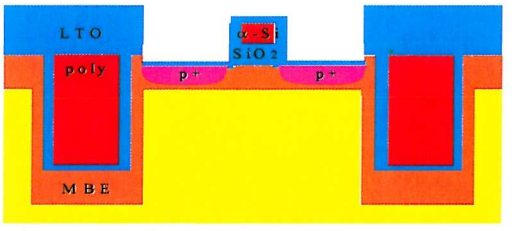
Step 5+6: poly-Si back etch, LTO Deposition II



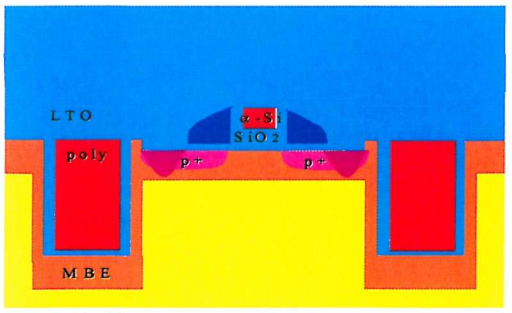
Step 7: Etch of Active Area



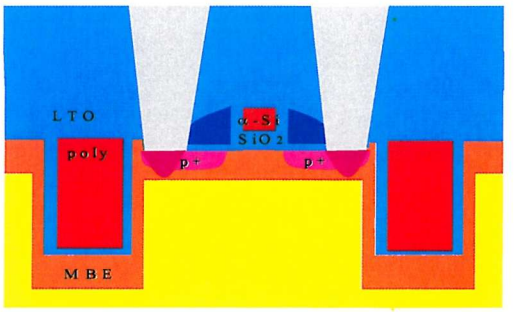
Step 8+9:Gate Dielectric and Amorphous Silicon Deposition



Step 10+11: Gate Stack Creation and LDD Implant



Step 12+13+14: Spacer Creation, HDD, and LTO Deposition



Step 15+16+17: Anneal, Contact Windows, and Metallisation

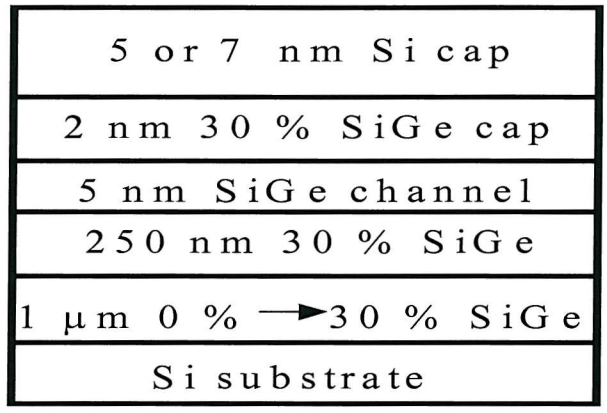


Figure 9.1: Grown MBE layer structure

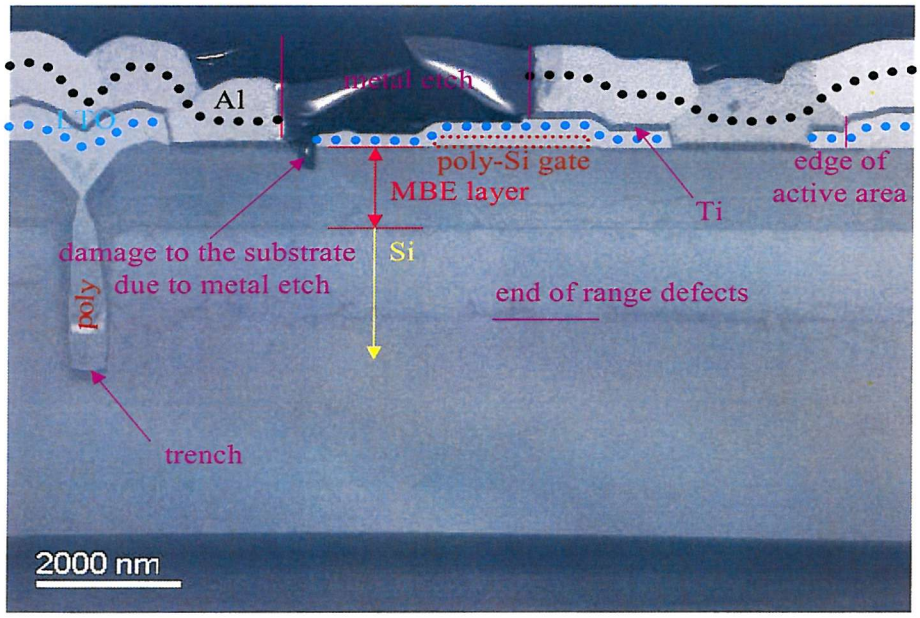
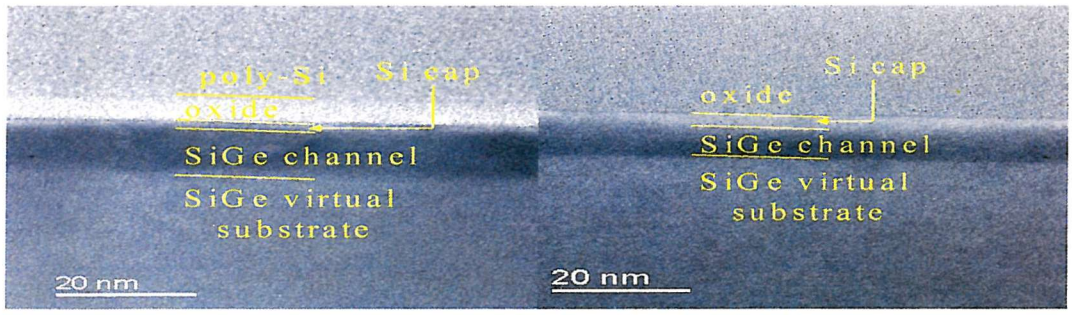
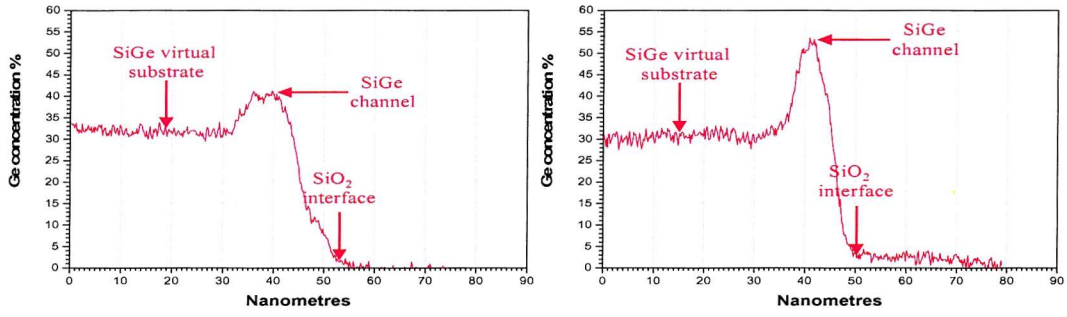


Figure 9.2: TEM photograph of a MOSFETs fabricated on a square pillar



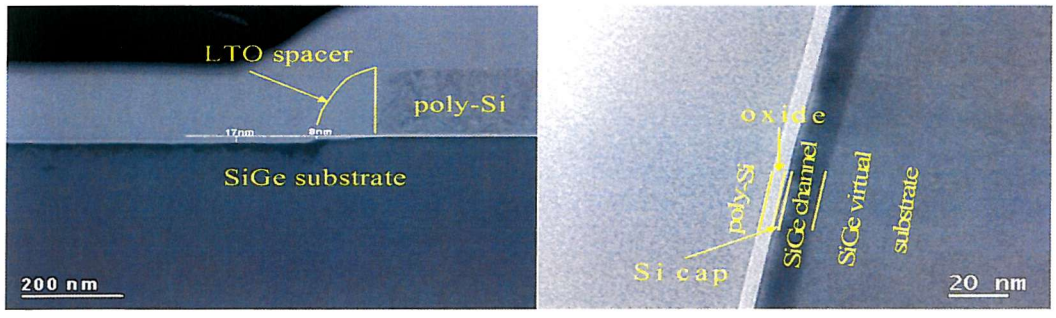
(a) (b)

Figure 9.3: SiGe underneath the gate dielectric (a) Thermal Oxide (b) LTO



(a) (b)

Figure 9.4: Ge profile underneath the gate dielectric (a) Thermal Oxide (b) LTO



(a) (b)

Figure 9.5: (a) Loss of substrate in source and drain of the processed devices with a thermal gate oxide (b) Fluctuation of the Ge content in the SiGe channel

wafer number	Si reference wafers		gate oxide in nm
5			5
9			50
11			5
wafer number	Ge content in the channel	Si cap in nm	gate oxide in nm
3	60 %	5	50
4	70 %	7	5
8	60 %	5	5
12	70 %	5	50
13	60 %	7	5

Table 9.1: Splits for the SiGe virtual substrate/limited area PMOS process

oxide is stripped off and a thin thermal oxide is grown (20 nm @ 900^oC) in order to protect the substrate against contamination during the n-well implant ($2 \cdot 10^{13}$ P^+/cm^2 @ 160 keV, $4 \cdot 10^{12}$ P^+/cm^2 @ 70 keV). The dopants are activated during an RTA (10 s at 1100^oC).

Now, the MBE layer is grown (step 2). The MBE layer structure for the SiGe wafer is given in fig. 9.1. The splits are given in tab. 9.1. A 250 nm Si layer is grown for the Si reference wafers. The MBE layer is not intentionally doped and contains according to the growers about $1 \cdot 10^{15} at./cm^3$ of phosphorus which comes from the stainless steel of the MBE growth reactor. From here on, the cleaning procedure deviates from the standard in order to prevent damage to the SiGe substrate. The SC-1 is omitted for all cleans which come in contact with the substrate.

The deposition of either 100 nm of LTO (step 3) is followed by the deposition of 900 nm of poly-Si for the refill of the trenches (step 4). The LTO protects the substrate in the following poly back etch leaving the trenches filled with poly-Si (step 5) and serves as part of the field oxide. Then, 500 nm LTO is deposited (step 6) as a field oxide. The active area is etched into the field oxide (step 7). A 5 nm oxide is grown at 800^oC (step 8) or a 50 nm thick gate oxide is deposited as gate dielectric (see tab. 9.1). 200 nm of insitu p-type doped α-Si is deposited as a gate material (step 9).

Fig.'s 9.3 (a) and (b) shows that a thin Si cap layer is left after gate oxidation/LTO deposition. The Si cap was initially 5 nm thick. The SiGe channel is substantially thicker for the thermally oxidized wafer than for the wafer with an LTO oxide indicating Ge diffusion during gate oxidation. Variations in color show that the Ge content in the Si cap and the SiGe channel underneath the thermal gate oxide can vary substantially over a small distance (see fig. 9.5 (b)).

In fig.'s 9.4(a) and (b), the Ge profile of the channel for the same samples is displayed. Whereas the virtual substrate layer and the SiGe channel can be recognized due to its Ge content, the localization of the oxide interface is difficult as a high Ge content can be found at the surface of the substrate much further away from the Ge peak in the SiGe channel than expected. This suggests that the profile may be broadened by the analysis technique. According to the growth

specifications, the Ge concentration in the channel should be 70 %. For the thermally oxidized sample the Ge peak concentration is reduced due to diffusion during gate oxidation, whereas no significant Ge diffusion can occur during LTO deposition. The Ge peak measured by EDS is far lower than intended. The specified error for the measurement technique is 1-2 %. However, a broadening of the Ge profile due to the analysis technique could also explain the discrepancy. Further processing will reduce the peak Ge concentration even more.

The gates are defined by e-beam lithography (step 10). Many devices were lost because their gates fell off. The problems probably were persevered by the trenches leading to a nonuniformity in the resist. The longest possible overetch is applied during dry etching in order to ensure that no fillets are left in the edges of the active areas. These are typical of devices which have the active area etched into the substrate instead of LOCOS because the edges are very steep.

For the wafers with a thermal gate oxide, a 5 nm oxide is grown at 800^oC) in order to protect the wafer during subsequent processing and anneal the damage due to the etch. Then, a LDD implant that consists of four implantation steps at the same energy and doping concentration is performed (step 11). Between each implantation step, the wafers are rotated by 90^oC ($4 \times 2.5 \cdot 10^{13}\ BF_2^+/cm^2$ @ $33\ keV$). The LTO spacers are created (step 12). The wafers with a thermal gate oxide are reoxidized ($5nm$ @ 800^oC) in order to protect the Si from contamination during implantation. After the HDD implant into source, drain, and gate ($5.0 \cdot 10^{15}\ BF_2^+/cm^2$ @ $45\ keV$ (step 13), a fluorine removal anneal is performed for 20 min at 650^oC.

400 nm LTO (step 14) are deposited for planarization of the wafer surface. The dopants are activated in an anneal (30 min. furnace anneal at 750^oC for the wafers with a deposited gate oxide, 30 min. at 800^oC for the wafers with a thermal gate oxide, step 15). Then, the contact windows are created (step 16) by the LTO oxide with $CHF_3 + Ar$. A dip etch in BHF is performed before the metallization stage in order to make good source, drain and gate contacts.

25 nm of Ti, 75 nm of TiN, 1 μm of Al, and 30 nm TiN are deposited and defined (step 17). The Ti layer reduces the contact resistance between metal and semiconductor. The first TIN layer prevents the spiking of Si and the second reflections during optical lithography. The metallic wires are created by dry etching the aluminium and titanium layer. Finally, the wafers are annealed at 420^oC for 30 min.. Several cleaning steps are implemented when ever necessary.

A cross section of a fabricated device is given in fig. 9.2. In this particular case, the metal etch seems to have penetrated into the substrate because metal mask overlapped with the contact window mask. Therefore, the metal etch was not stopped by the LTO oxide. This process problems may be responsible for the high source and drain series resistance found for many devices. A substantial amount of the substrate is lost in the source and drain due to cleaning and thermal oxidation for the investigated MOSFET with a thermal gate oxide (see fig. 9.5 (a)). Close to the poly-Si gates, less substrate is lost because it is protected by the LTO spacers.

9.2 Measurements and Parameter Extraction

P-MOSFETs on square pillars ($10\ \mu m \times 10\ \mu m$) with $W = 8\ \mu m$ and $L = 2, 1, 0.5, 0.2\ \mu m$, their reference devices (not fabricated on square pillars), p-MOSFETs on long, $7.07\ \mu m$ or $14.85\ \mu m$ wide pillars with $W = 5.66\ \mu m$ and $L = 1.41, 9.19, 19.80\ \mu m$ along the [010] direction, and their reference devices (not fabricated on pillars) were measured.

Out of twelve measured devices of a certain type, the best six representative devices were chosen for parameter extraction with the two routines - routine 1 described in section 4.4 and routine 2 described in section 4.5. The drain current was fixed at $-50\ mV$, the gate voltage was varied between $1\ V$ and $-3\ V$ for the devices with a thermal gate oxide and between $2\ V$ and $5\ V$ for the devices with an LTO gate oxide. Additionally, the MOSFETs on square pillars with $W = 8\ \mu m$ and $L = 0.5\ \mu m$ were measured for $V_d = -2.5\ V$.

In this chapter, only the parameters extracted with routine 1 are given because many devices displayed bad turn-off characteristics rendering the application of routine 2 inaccurate. For wafer # 4, # 8, # 12, # 13, the leakage current, i. e. the drain current measured for the highest applied gate voltage was subtracted before application of routine 2 ($1\ V$ or $2\ V$). However, this modification was not very successful because the leakage changed as a function of the gate voltage in some cases. The mobilities extracted as a function of the gate voltage with routine 2 suggest that routine 1 can be applied. The complete results are given in appendix D.

The mobilities extracted with routine 1 were displayed as a function of the effective vertical field E_{eff}:

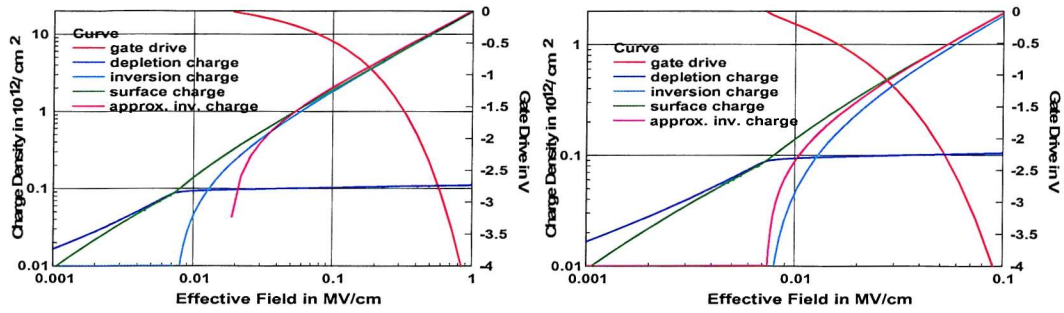
$$E_{eff} = \frac{1}{\epsilon_{ch}}[Q_b(V_g) + \eta Q_i(V_g)] \tag{9.1}$$

Where ϵ_{ch} is the permittivity of the channel, $\eta = 1/3$ for holes, Q_b the bulk charge, and Q_i the inversion charge. It was assumed that the substrate was n-type ($N_d = 1 \cdot 10^{15}/cm^2$) and always $\epsilon_{ch} = \epsilon_{si}$. For the calculation of Q_b and $Q_i = Q_s - Q_b$, eq.'s (3.33) and (3.22) were used. $V_g(\phi_s)$ was calculated with Eq. (3.8) and $V_{fb} = 0\ V$. The approximate inversion charge $Q_{iapp} = C_{ox} \cdot (V_g - V_{th})$ was calculated from a constant capacitance approximation and the threshold voltage was chosen so that the approximate inversion charge equals the surface charge $Q_{iapp} = Q_s$ for a hole density of $1 \cdot 10^{12}/cm^2$. The depletion charge, inversion, surface and approximate inversion charge are given as a function of the effective field and the gate drive in fig.'s 9.6 and 9.7 for $t_{ox} = 5\ nm$ and $t_{ox} = 50\ nm$.

9.3 Results

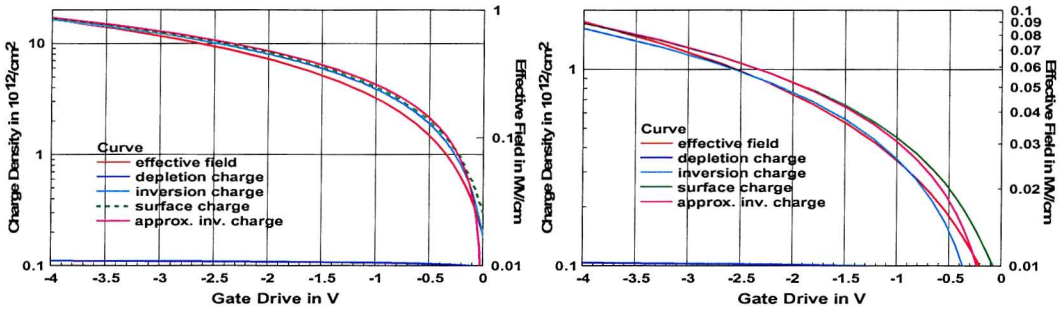
9.3.1 MOSFETs with an LTO Gate Oxide

Devices on square pillars V_{th}, ΔL, and R_{sd} for the LTO devices on square pillars and their reference devices are given in tab. 9.2. The magnitude of the threshold voltage decreases as a



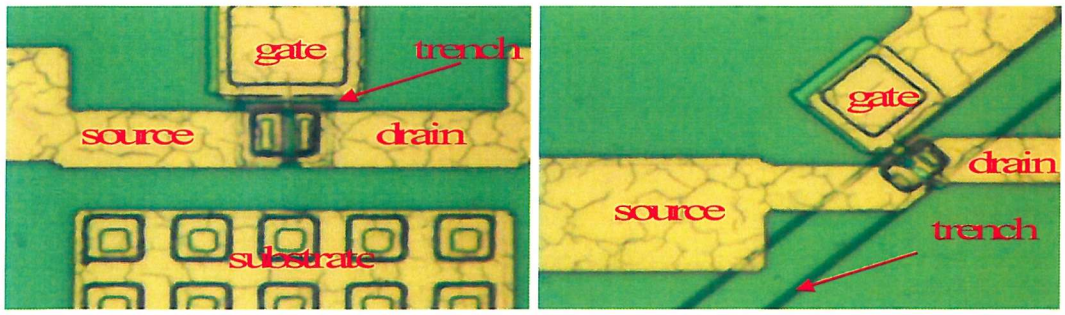
(a) (b)

Figure 9.6: Charge densities and gate drive as a function of the effective field for $N_d = 1{\cdot}10^{15}/cm^2$ (a) 5 nm oxide (b) 50 nm oxide



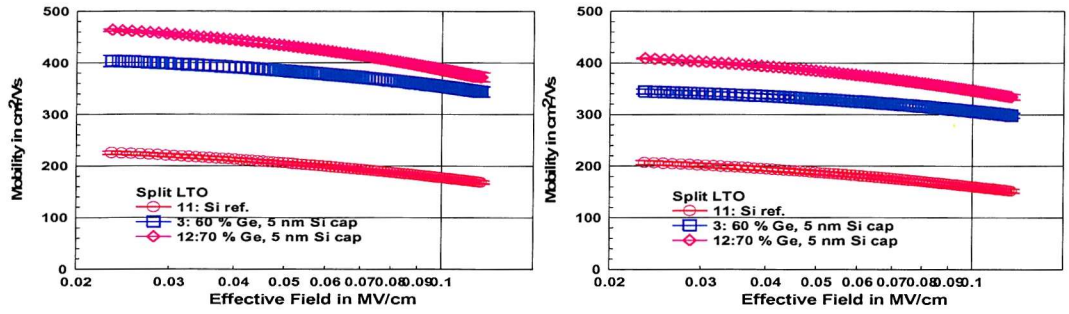
(a) (b)

Figure 9.7: Charge densities and effective field as a function of the gate drive for $N_d = 1{\cdot}10^{15}/cm^2$ (a) 5 nm oxide (b) 50 nm oxide



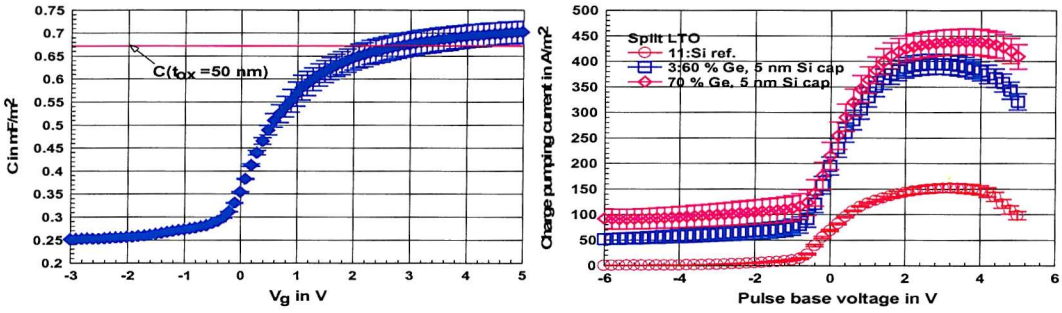
(a) (b)

Figure 9.8: MOSFETs fabricated on (a) a square pillar (b) a long pillar oriented 45° to the mayor crystal orientation



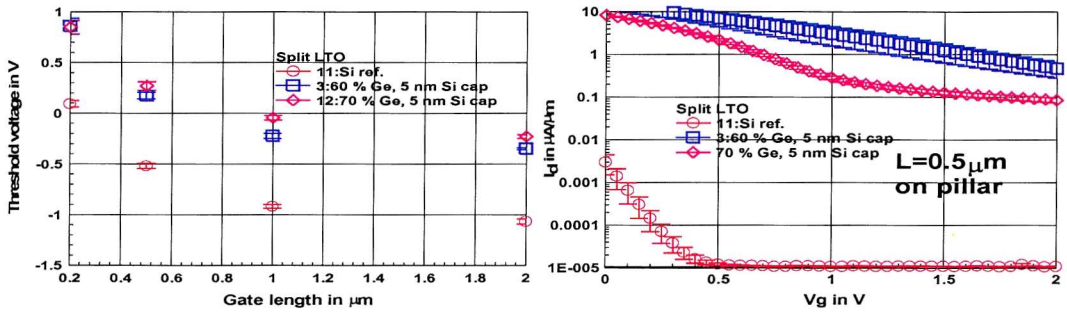
(a) (b)

Figure 9.9: Mobility for the MOSFETs with an LTO gate oxide as a function of the effective electric field (a) on square pillars (b) not on square pillars



(a) (b)

Figure 9.10: (a) HF capacitance for the Si reference (b) I_{cp} for p-MOSFETs with $L = 2\ \mu m$ and $W = 8\ \mu m$ measured at 500 kHz and for $\Delta V_A = 10V$



(a) (b)

Figure 9.11: (a) Threshold voltage as a function of the gate length (b) Subthreshold characteristics for $V_d = -2.5\ V$ for $L = 0.5\ \mu m$

Wafer no. Ge cont.	$V_{th} in V$ ($L = 1\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$	Int. states in $10^{11}\ /cm^2$
On pillars				
11; Si ref.	-0.917 ± 0.017	0.08 ± 0.01	2.37 ± 0.20	1.90 ± 0.11
3; 60 % Ge	-0.224 ± 0.025	0.23 ± 0.01	2.44 ± 0.88	4.92 ± 0.25
12; 70 % Ge	-0.041 ± 0.023	0.32 ± 0.02	1.82 ± 0.20	5.48 ± 0.31
Not on pillars				
11; Si ref.	-0.996 ± 0.016	0.09 ± 0.01	2.71 ± 0.34	not meas.
3; 60 % Ge	-0.169 ± 0.015	0.16 ± 0.00	3.07 ± 0.12	not meas.
12; 70 % Ge	0.013 ± 0.023	0.29 ± 0.01	2.33 ± 0.20	not meas.

Table 9.2: V_{th}, ΔL, and R_{sd} for the LTO devices on square pillars and their reference devices

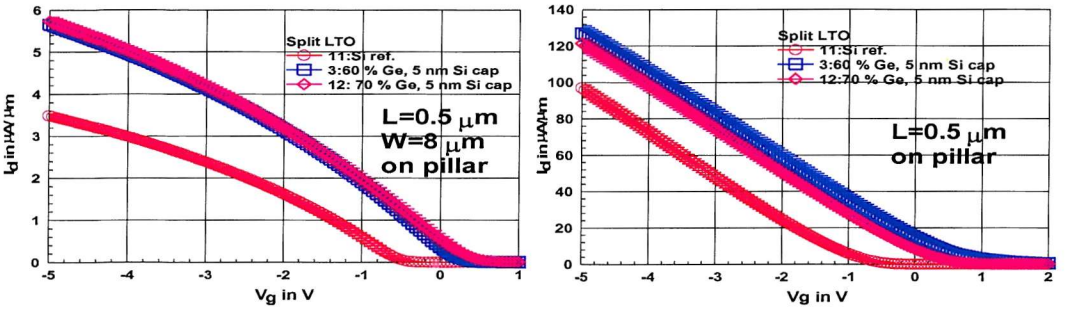
function of the Ge content due to the band offset in the SiGe channel. The devices display a considerable threshold voltage roll-off because the substrate is not heavily doped. Surprisingly, ΔL is higher for the Si reference devices than for the SiGe devices. The values for the SiGe devices are unreasonable as devices with $L = 0.2\ \mu m$ were measured which is smaller than ΔL for the SiGe devices.

The SiGe wafers and Si devices display a comparable source and drain series resistance. The fact that the source and drain series resistances were not extracted very accurately should not surprise although devices with a small channel length were used. The influence of the source and drain series resistance is especially important when the channel is highly populated. However, this spilt has a thick gate oxide and even a high gate drive does not populate the channel considerably.

For the calculation of the mobility, the constant capacitance approximation was used. The effect of the Si cap was neglected. HF CV-measurements show that the gate oxide is about 50 nm thick as assumed in the calculations (see fig. 9.10 (a)). The mobility as a function of the effective field is given in fig's 9.9 (a) and (b). The SiGe devices beat the Si references by more than a factor 2 for the devices fabricated on pillars. The SiGe devices not fabricated on pillars display much worse transport properties than the SiGe devices fabricated on pillars. It is worth noting that the Si devices on pillars seems as well to have a slightly better performance than the Si devices not fabricated on pillars.

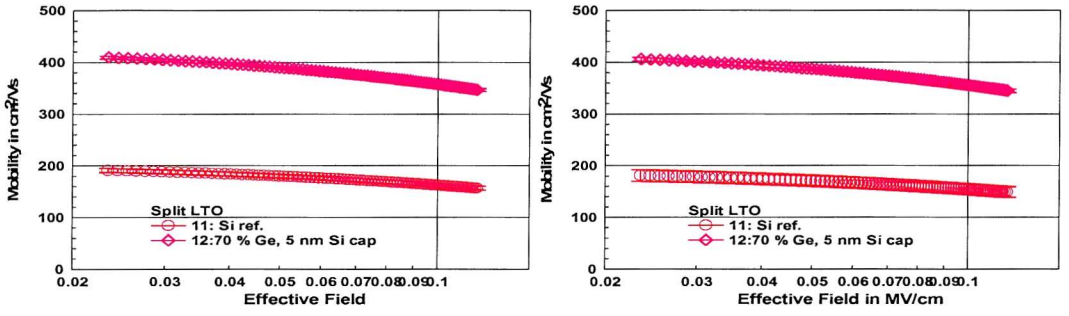
The interface state density was extracted from charge pumping measurement for $\Delta V_A = 10\ V$ and $f = 500\ kHz$ (see fig. 9.10 (b)). Despite the fact that the gate dielectric is deposited, a higher density of interface states was found for the SiGe devices than for the Si devices.

Measurements at high gate drive In fig. 9.12 (a), the drain currents of the Si reference devices and the SiGe devices with $L = 0.5\ \mu m$ are compared for $V_d = 50\ mV$. The SiGe devices have clearly a steeper slope than the Si reference. In contrast (see fig. 9.12(b)), the slope for the Si and SiGe devices is pretty similar when the devices are measured at $V_d = -2.5\ V$. Obviously, the velocity of the SiGe devices saturates at high drain voltages.



(a) (b)

Figure 9.12: Drain current as a function of the gate voltage for for $L = 0.5\ \mu m$ and (a) $V_d = -50\ mV$ (b) $V_d = -2.5\ V$



(a) (b)

Figure 9.13: Mobility for the MOSFETs with an LTO gate oxide fabricated as a function of the effective electric field (a) fabricated on long, 14.85 μm wide pillars (b) not fabricated on long pillars

Fig. 9.11 (b) shows the turn-off characteristics of the same devices. The SiGe devices are far more leaky than the Si reference. The leakage current passes from the drain to the substrate and the source. It can be attributed to the low doping concentration of the substrate. Instead of being unintentionally doped (n-type: $1 \cdot 10^{15}/cm^2$), some parts of the substrate may be p-type. The leakage current may be higher for the SiGe devices because they have a thicker MBE layer or because the p-type dopants are introduced by the Ge source in MBE.

Devices on long pillars Fig.'s 9.13 (a) and (b) compare devices which are fabricated on long, 14.85 μm wide pillars along the [010] direction to devices which are not fabricated on such structures. Long pillars along the [010] direction could relieve both strain components ([011] and [0$\bar{1}$1] directions, see section 5.2.1). However, the performance is obviously not enhanced by the fabrication on long pillars [1]. Devices on long, 7.07 μm wide pillars can as well not beat their

[1]Three different gate lengths ($L = 1.41, 9.19, 19.80\ \mu m$) were used for the extraction of the parameters of the devices fabricated on long, 14.85 μm wide pillars, whereas two devices were used for their references and devices

Wafer no. ; Ge cont.	$V_{th} in V$ ($L = 1\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$	Int. states in $10^{11}/cm^2$
On pillars				
5; Si ref.	0.040 ± 0.003	0.10 ± 0.01	2.54 ± 0.13	1.11 ± 0.06
9; Si ref.	0.041 ± 0.003	0.26 ± 0.00	2.71 ± 0.19	0.98 ± 0.04
13; 60 % Ge	0.396 ± 0.003	0.16 ± 0.07	5.92 ± 1.08	2.47 ± 0.17
8; 60 % Ge	0.465 ± 0.006	0.36 ± 0.05	2.91 ± 0.52	5.49 ± 0.58
4; 70 % Ge	0.372 ± 0.006	0.34 ± 0.02	1.55 ± 0.13	2.66 ± 0.13
Not on pillars				
5; Si ref.	0.021 ± 0.004	0.09 ± 0.01	2.66 ± 0.16	not meas.
9; Si ref.	0.033 ± 0.005	0.23 ± 0.01	2.15 ± 0.20	not meas.
13; 60 % Ge	0.442 ± 0.004	0.47 ± 0.11	10.49 ± 2.02	not meas.
8; 60 % Ge	0.541 ± 0.012	1.00 ± 0.15	11.47 ± 2.07	not meas.
4; 70 % Ge	0.423 ± 0.008	0.28 ± 0.02	1.88 ± 0.18	not meas.

Table 9.3: V_{th}, ΔL, and R_{sd} for the devices with a thermal gate oxide on square pillars and their reference devices

references (see appendix D). The extracted mobilities are very similar to those in fig. 9.9 (b). On wafer # 3, no working devices of these types were found due to the problem with gate definition.

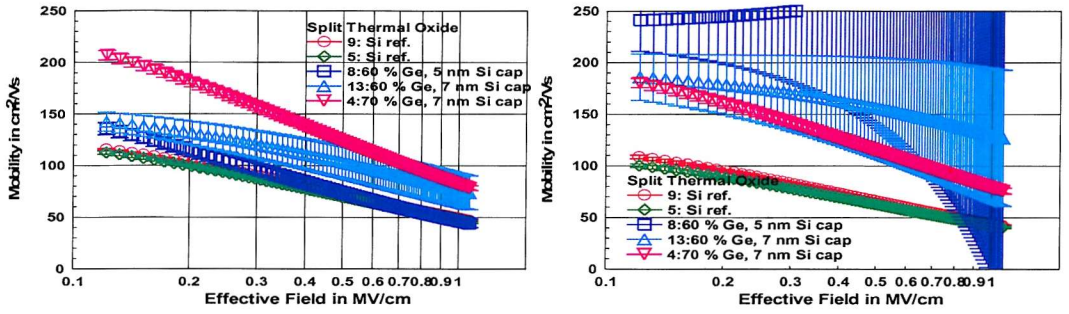
9.3.2 MOSFETs with a Thermal Gate Oxide

Devices on square pillars V_{th}, ΔL, and R_{sd} for the devices with a thermal gate oxide on square pillars and their reference devices are given in tab. 9.3. The magnitude of the threshold voltage decreases as a function of the Ge content due to the band offset in the SiGe channel. The devices with a thermal gate oxide display less threshold voltage roll-off than the devices with an LTO gate oxide because the gate dielectric is thinner (see fig. 9.16 (b)).

The devices on wafer # 8 and # 13 suffer from a high source and drain series resistance. As the source and drain series resistances are very dispersed , all parameter cannot be accurately extracted for these wafers. Their dispersion and the fact that they can be comparable to the channel resistance of the longest devices ($L = 2\ \mu m$) results for the devices of wafer # 8 on pillars in the extraction of a low value for the source and drain series resistance. The low source and drain series resistance for the SiGe devices with a 70 % Ge channel channel indicates that the virtual substrate devices can have a lower source and drain series resistance than the Si references due to better dopant activation and higher mobility as long as the process does not go wrong. High source and drain series resistances can be attributed to the misalignment due to metal definition as the current path between source and drain can be interrupted by damage due to the penetration of the metal etch in to the substrate (see fig. 9.2).

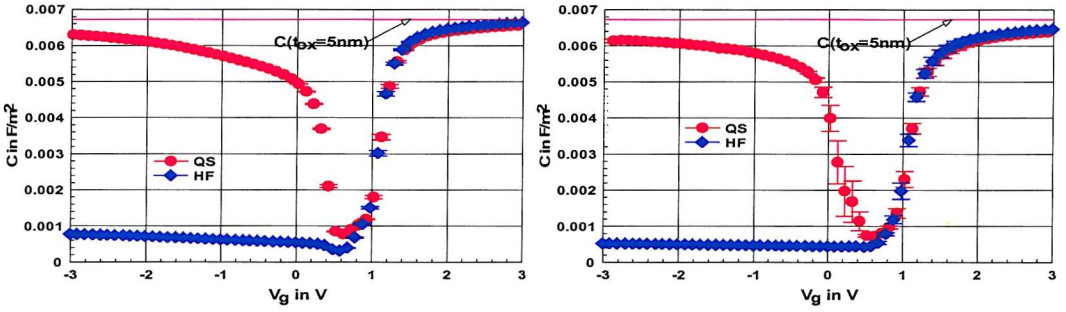
For the calculation of the mobility, the constant capacitance approximation was used. The effect of the Si cap was neglected. HF CV-measurements show that the gate oxide is about

fabricated on long, 7.07 μm wide pillars ($L = 1.41, 9.19\ \mu m$).



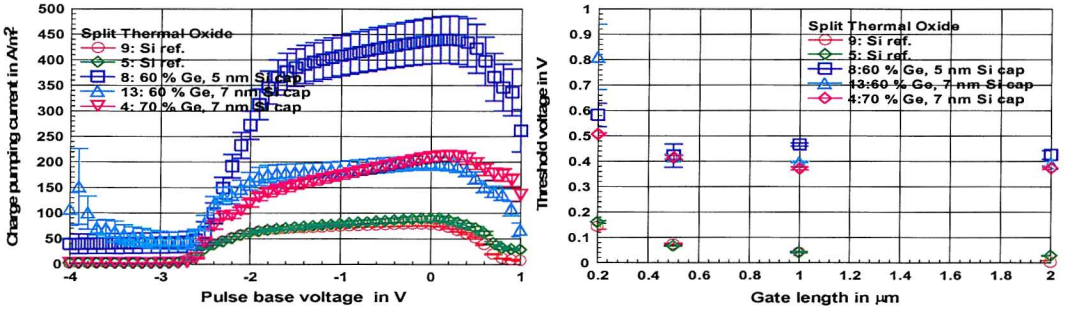
(a) (b)

Figure 9.14: Mobility for the MOSFETs with a thermal gate oxide as a function of the effective electric field (a) on square pillars (b) not on square pillars



(a) (b)

Figure 9.15: (a) HF and quasistatic capacitance (a) 70 % Ge (b) Si reference



(a) (b)

Figure 9.16: (a) I_{cp} for p-MOSFETs with $L = 2\ \mu m$ and $W = 8\ \mu m$ measured at 500 kHz and for $\Delta V_A = 10V$ (b) Threshold voltage as a function of the gate length

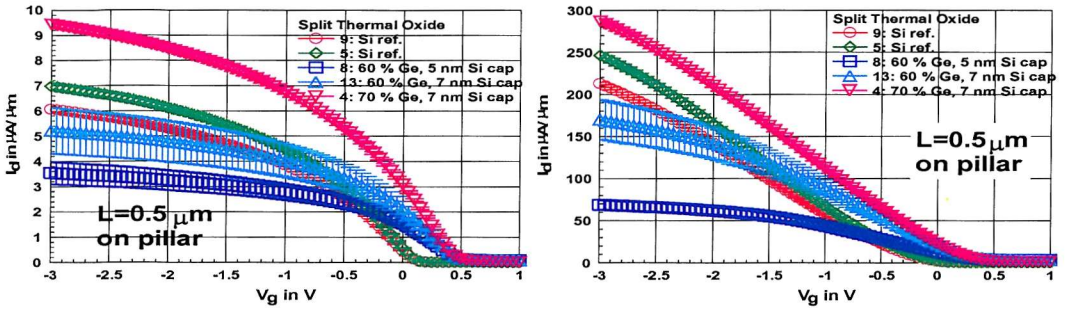


Figure 9.17: Drain current as a function of the gate voltage for for $L = 0.5\ \mu m$ and (a) $V_d = -50\ mV$ (b) $V_d = -2.5\ V$

5 nm thick as assumed in the calculations (see fig.'s 9.15 (a) and (b)). The quasi-static CV-measurements suggest that the application of the constant capacitance approximation may reduce the extracted mobility by 30 %. As the quasi-static CV-curves are comparable for the Si and SiGe wafers, the neglect of the Si cap on the gate capacitance is justified.

The mobility as a function of the effective field is given in fig's 9.14 (a) and (b). The mobilities were extracted at considerably higher effective vertical fields than for the LTO split and are therefore considerably lower. The SiGe devices beat the Si references by nearly a factor 2 for the devices fabricated on pillars. The reduced performance of the SiGe devices compared to the LTO split can be attributed to the higher thermal budget and enhanced Ge diffusion. The SiGe devices with a 70 % SiGe not fabricated on pillars channel displays slightly worse transport properties than those on pillars. The Si devices on pillars seem again to have a slightly better performance than the Si devices not on pillars. Precise mobilities cannot be extracted for wafer # 8 and # 13 because of the high and dispersed values for source and drain series resistance which has an important effect on the device characteristics for relatively short gates. The error in the extracted mobility increases with the gate drive or the effective field because the contribution of the source and drain series resistance becomes more important.

The interface state density was extracted from charge pumping measurements for $\Delta V_A = 10\ V$ and $f = 500\ kHz$ (see fig. 9.16 (a)). As expected, a higher density of interface states was found for the SiGe devices and for wafer # 8 which has a thinner Si cap layer than the two other SiGe wafers (see tab. 9.3).

Measurements at high gate drive In fig. 9.17 (a) the drain currents of the Si reference devices and the SiGe device with a 70 % Ge channel with $L = 0.5\ \mu m$ are compared for $V_d = -50\ mV$. The SiGe device has clearly a steeper slope than the Si reference. In contrast (see fig. 9.17 (b)), the slopes for the Si and SiGe devices with a 70 % Ge channel are pretty similar when the devices are measured at $V_d = -2.5\ V$. This suggests that carrier velocity of the SiGe devices

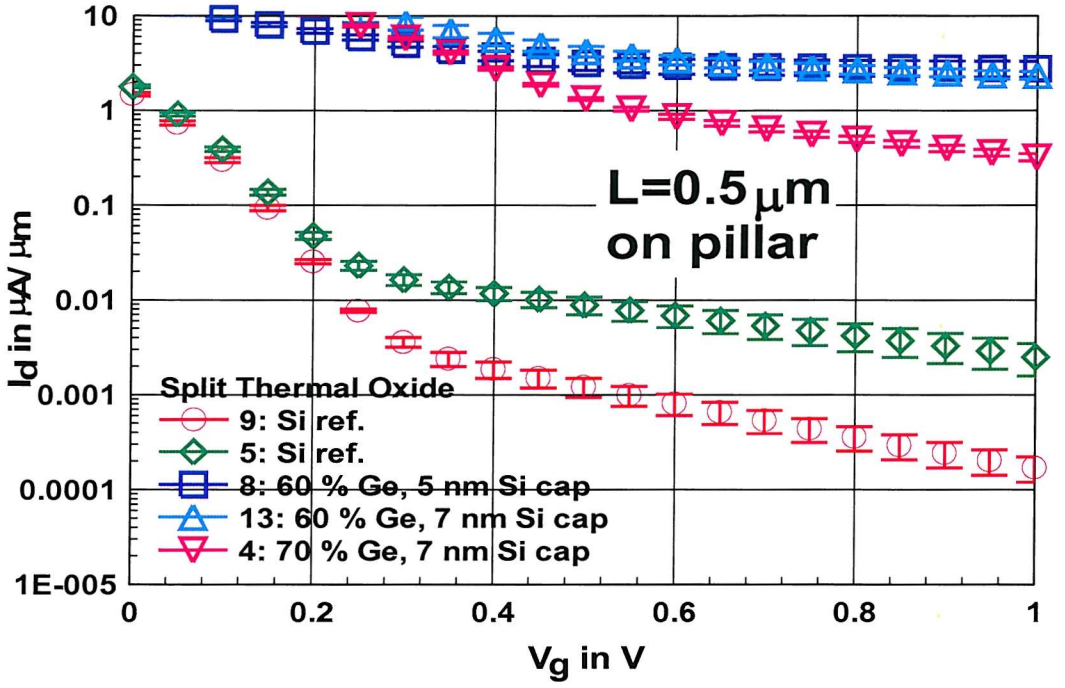


Figure 9.18: Subthreshold characteristics for $V_d = -2.5\ V$ and $L = 0.5\ \mu m$

saturates as it is the case for the LTO devices. The SiGe devices display as well for this split worse turn-off characteristics than the Si devices (see fig. 9.18).

Devices on long pillars Fig.'s 9.19 (a) and (b) compare devices which are fabricated on long, 14.85 μm wide pillars along the [010] direction to devices which are not fabricated on such structures. As for the LTO split, the performance is not enhanced by the fabrication on pillars[2]. Devices on long, 7.07 μm wide pillars can as well not beat their references (see appendix D). On wafer # 13, no working devices of these types were found due to the problem with gate definition. The mobilities for wafer # 8 could be determined accurately due to the long devices lengths. The thicker cap layer may be to blame for the higher mobility degradation on wafer # 8 than wafer # 4.

9.4 Conclusion

The mobility enhancement of more than 100 % for MOSFETs with an LTO gate oxide is far lower than values cited in literature for virtual substrate MOSFETs [1-6]. According to reference [1], a substantial mobility enhancement is only found for Ge concentrations above 60 %. However, the fabricated MOSFETs seem to have a Ge content of less than 55 % in the channel. The Ge peak concentration is substantially reduced by the interdiffusion of Si and Ge enhanced by strain [11]

[2]Three different gate lengths ($L = 1.41, 9.19, 19.80\ \mu m$) were used for the extraction of the parameters of the devices fabricated on long, 14.85 μm wide pillars, whereas two devices were used for their references and devices fabricated on long, 7.07 μm wide pillars($L = 1.41, 9.19\ \mu m$).

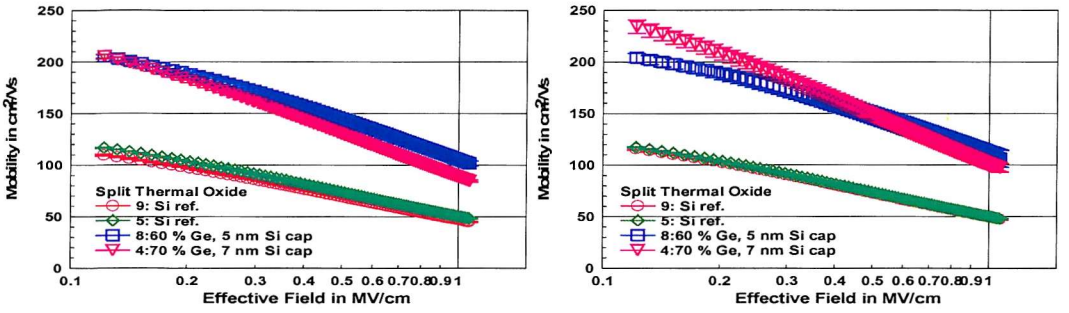


Figure 9.19: Mobility for the MOSFETs with a thermal gate oxide fabricated as a function of the effective electric field (a) fabricated on long, 14.85 μm wide pillars (b) not fabricated on long pillars

during thermal oxidation.

Although the mobility of 350 cm^2/Vs cited in ref [5] for a process with a thermal oxide is considerably higher than the mobilities found for the split with a thermal oxide, the mobility enhancement itself is comparable, because the same authors quote in ref. [4] a mobility in the range of 200 cm^2/Vs for Si devices. This corresponds to a mobility enhancement of less than a factor 2 (= 1.75) and is slightly lower than the mobility enhancement of a factor 1.87 obtained for the spilt with a thermal gate oxide. The fact that the measurements for this batch were performed at higher vertical fields and the use of the constant capacitance approximation can explain the extraction of low mobilities. For the same split, a lower source and drain series resistance has been observed for SiGe MOSFETs than for the Si reference MOSFETs.

A better performance could be demonstrated for SiGe MOSFETs fabricated on microscopic pillars than those fabricated on a plain substrate due to a reduction in crosshatch and and improved relaxation of the virtual substrate. However, SiGe p-MOSFETs do not beat Si references for short channel lengths and high gate drives ($L = 0.5\ \mu m$ and $V_d = -2.5\ V$) supporting the idea that Si and SiGe have similar saturation velocities [12]. As the SiGe p-MOSFETs display bad turn-off characteristics, it is therefore recommended to use a doped virtual substrate for further batches.

Bibliography

[1] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, E. A. Fitzgerald, "Hole mobility enhancement in strained $Si/Si_{1-y}Ge_y$ p-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-x}G_x$ $(x < y)$ virtual substrates", Appl. Phys. Lett., Vol. 79, No. 25, pp. 4246-4248, 2001

[2] M. L. Lee, C. W. Leitz, Z. Cheng, A. J. Pitera, T. Langado, M. T. Currie, G. Taraschi, and E. A. Fitzgerald, "Strained Ge channel p-type metal-oxide-semiconductor field-effect transistors grown on $Si_{1-x}Ge_x/Si$ virtual substrates", Appl. Phys. Lett. Vol. 79, No. 20, pp. 3344-3446, 2001

[3] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Tarasci, and E. A. Fitzgerald, "Carrier mobilities and process stability of strained Si n- and p-MOSFETs on SiGe virtual substrates", J. Vac. Sci. Technol. B, Vol. 19, No. 6, pp. 2268-2279, 2001

[4] G. Hoeck, E. Kohn, C. Rosenblad, H. v. Kaenel, H.-J. Herzog, and U. Koenig, "High hole mobility in $Si_{0.17}Ge_{0.83}$ channel metal-oxide-semiconductor field-effect transistors grown by plasma-enhanced chemical vapor deposition", Appl. Phys. Lett., Vol. 76, No. 26, pp. 3920-3922, 2000

[5] U. Koenig, M. Zeuner, G. Hoeck, T. Hackbarth, M. Glueck, T. Ostermann, and M. Saxarra, "n- and p-Type SiGe HFETs and circuits", Solid-State Electronics, Vol. 43, pp. 1383-1388, 1999

[6] S. J. Koester, R. Hammond, and J. O. Chu, "Extremely High Transconductance $Ge/Si_{0.4}Ge_{0.6}$ p-MODFET's Grown by UHV-CVD", IEEE Electron Device Lett., Vol. 21, No. 3, pp. 110-112, 2000

[7] E. Muller, R. Hartmann, C. David, and D. Grutzmacher, "Structural characterization of SiGe step graded buffer layers grown on prestructured Si [100] substrates by molecular beam epitaxy", Thin Solid Films, Vol. 336, pp. 92-95, 1998

[8] T. Rupp, F. Kaesen, W. Hansch, E. Hemmerl, D. J. Gravesteijn. R. Schorer, E. Silveira, G. Abstreiter, and I. Eisele, "Defect-free strain relaxation in locally MBE-grown SiGe heterostructures", Thin Solid Films, Vol. 294, pp. 27-32, 1997

[9] E. A. Fitzgerald, P. D. Kirchner, R. Proano, G. D. Pettit, J. M. Woodall, and D. G. Ast, "Elimination of interface defects in mismatched epilayers by a reduction in growth area", Appl. Phys. Lett., Vol. 52, No. 18, pp. 1496-1498, 1988

[10] R. Hammond, P. J. Phillips, T. E. Whall, E. H. C. Parker, T. Graf, H. von Kanel, and A. J. Shields, "The elimination of surface cross-hatch from relaxed, limited-area $Si_{1-x}Ge_x$ buffer layers", Appl. Phys. Lett., Vol. 71, No. 17, pp. 2517-2519, 1997

[11] N. E. B. Cowern, P. C. Zalm, P. van der Sluis, D. J. Gravesteijn, and W. B. de Boer, "Diffusion in Strained Si(Ge)", Phys. Rev. Lett., Vol. 72, No. 16, pp. 2585-2588, 1994

[12] R. J. P. Lander, Y. V. Ponomarev, J. G. M. van Berkum, and W. B. de Boer, ""High Hole Mobilities in Fully-Strained $Si_{1-x}Ge_x$ Layers ($0.3 < x < 0.4$) and their Significance for SiGe pMOSFET Performance", IEEE Trans. Electron. Dev., Vol. 48, No. 8, pp. 1826-1832, 2001

Chapter 10

Conclusions and Suggestions for Further Work

SiGe is a novel compound material that can be beneficially used in the channel, gate, source and drain of MOSFETs. In source and drain, SiGe can improve the short channel characteristics, reduce the source and drain series resistance and parasitic bipolar transistor action. Poly-SiGe gates can be used for the adjustment of the threshold voltage and reduction of the gate resistance.

Strained SiGe heterostructures possess transport properties superior to Si. Their integration in the channel of MOSFETs can lead to an enhanced speed performance and is therefore an alternative to the scaling of the channel length. This thesis has focused on the integration of strained SiGe layers in the channel of p-MOSFETs. The SiGe layers have either been fabricated by implantation or MBE.

10.1 Integration of Heterostructures in the Channel of MOSFETs

As both design and material properties determine the performance of MOSFETs, SiGe MOSFETs will beat conventional MOSFETs only if their design and fabrication process are competitive. The implementation of heterostructures into a competitive design struggles in particular with the following problems.

10.1.1 Substrate Doping

MOSFETs are continuously scaled because their speed is inversely proportional to the square of the channel length. However, submicron MOSFETs require a highly doped substrate in order to maintain their turn-off capability. Whereas the incorporation of dopants in SiGe during MBE is not well investigated, the implantation of the substrate after epilayer growth can damage the

heterolayer and introduce dopants in the channel region reducing the mobility. A very thin MBE layer would not have to be doped in order to guarantee good device characteristics.

10.1.2 Thermal Budget

Annealing temperatures in excess of $1000^{o}C$ are used for the dopant activation of conventional MOSFETs and the diffusion of the dopants in the poly gates to the gate oxide. Such elevated temperatures are in general not acceptable for SiGe MOSFETs as they can lead to the relaxation of the heterolayer and Ge diffusion. The diffusion of dopants into the SiGe channel is a possible negative side effect of a high temperature anneal as the superior mobilities found for SiGe MOSFETs are partly attributed to the low dopant concentration in the conducting channel.

10.1.3 Layer Architecture and Process Flow

The layer architecture is closely linked to the the thermal budget and the doping of the substrate. The supportable thermal budget is determined by the layer structure and usually not exactly known. In thick MBE layers with high Ge content, particularly high intrinsic mobilities can be achieved. However, these layers may be especially susceptible to strain relaxation. Lowering the thermal budget reduces the dopant activation as well as the damage to the heterostructure by strain relaxation.

In order to incorporate thick MBE layers into CMOS requiring n- and p-type dopants in the wells, the process flow has to be modified. An insitu doped MBE layer would have to be removed over some parts of the wafer. If dopants are introduced by implantation, the strained SiGe channel might be damaged. Alternatively, the MBE growth could be interrupted for well implantation before the growth of the SiGe channel. Further complications arise from the fact that the optimum performance for n- and p-MOSFETs may be achieved for different layer architectures. All effects have to be balanced in order to achieve a cost-effective process with optimum performance.

10.1.4 Wafer Cleaning

The high yield obtained for conventional MOSFETs is based on the well established cleaning procedures. The amount of Si removed during cleaning is negligible for conventional MOSFETs. In contrast, the etch of the substrate during cleaning can be critical for HMOSFETs requiring an exact control of the vertical layer architecture. This control is made more difficult because the etch rate of the standard cleaning procedures increases considerably with the Ge content.

Wafer cleaning will be a crucial issue for industrial HMOSFET production when a high yield has to be achieved and should be studied before the industrial fabrication of HMOSFETs starts. Currently, it is far more important to ensure that the vertical layer structure is well controlled

during fabrication so that the potential benefits of SiGe MOSFETs can be proven. Instead of maximizing the yield, the preliminary aim is to minimize the loss of substrate during cleaning.

10.2 Fabricated Batches

Three batches addressing the issues raised in the above section have been fabricated. Actually, there is only a limited number of ways to integrate heterolayers into MOSFETs. Due to the considerable amount of time it takes to complete a full process, it is impossible to choose and optimize a particular way in the course of a Ph.D. project. Instead, three different SiGe process were carried out in order to investigate the fabrication of SiGe MOSFETs from a different angle.

10.2.1 Pseudomorphic SiGe CMOS Process

The pseudomorphic SiGe CMOS process was designed as a gentle approach towards a competitive industrial process and advancement of the process given in [1]. The performance of HMOSFETs with an insitu doped MBE layer having experienced an elevated thermal budget was to be analyzed. The dopant concentration in the substrate satisfied threshold voltage and short-channel requirements.

A 2 nm Si cap on top of the heterochannel underneath the gate oxide was expected according to experiments on Si samples. However, extensive SIMS, TEM, and x-ray analysis revealed no heterochannel at all and only negligible Ge concentrations attributed to shutter leakage. Experimental evidence suggests that the SiGe channel was not grown.

However, the control of the Si cap layer thickness is far more critical than previously assumed. Recent research [2] shows that the interfaces of SiGe heterostructures grown by MBE are not perfectly defined. Residual Ge in the Si cap enhances the etch rate of the HF [3] and SC-1 [4] cleans considerably. For future batches, it is advised to use only SC-2 in order to clean the MBE layer.

Even so no SiGe devices were fabricated, the SiGe CMOS batch highlighted the following processing problems:

1. Poly-Si fillets: Many devices displayed a considerable gate leakage which may be attributed to poly-Si fillets left in the edges of the active areas after gate definition. Poly-Si fillets are typical for SiGe batches using LTO as field oxide. The etch of active area windows into the field oxide results in steep edges. A deposited layer is much thicker in the edges of the active area windows than on the plane wafer surface. In these edges, the gate material can only be removed if a sufficient overetch is applied. However the possible overetch is limited by the thickness of the gate oxide. Modern devices with a thin gate oxide do not support a considerable over etch. Therefore, thinner poly-Si layers should be employed for future batches.

2. Poly depletion: The reduced thermal budget for this batch did not allow the diffusion of a sufficient number of dopants to the oxide interface in the gate in order to avoid poly depletion. The increase of the implantation energy is no solution because implantation profiles have in general a long tail and dopants would penetrate into the substrate during the implantation of the gate. The use of insitu doped poly-Si is a suitable alternative for PMOS, whereas a short fix may not be found for CMOS using n-type n-MOSFET gates and p-type p-MOSFET gates satisfying threshold voltage requirements.

3. Integration of the MBE layer: A thick insitu doped MBE buffer layer was used in order to satisfy threshold voltage requirements and ensure a good crystal quality of the SiGe channel and good short channel characteristics. Introducing dopants by implantation can damage the heterostructure. However, the insitu doping is especially problematic for CMOS requiring n- and p-wells. N-MOSFETs with a counterdoped MBE have poor device characteristics. Further alternatives are far more complicated (e.g., the etch of the MBE layer for the n-MOSFETs, selective epitaxial growth). The growth of a thin, undoped MBE layer would simplify CMOS processing considerably.

10.2.2 SiGe CMOS/Ge^+ Implantation Process

In order to explore Ge^+ implantation as an alternative to MBE for CMOS fabrication, the SiGe CMOS/Ge^+ implantation process [5] was carried out. Although only low Ge concentrations and no abrupt interfaces can be achieved by implantation, the use of Ge^+ implantation may be so attractive because the integration of Ge^+ implanted MOSFETs into a CMOS process is very simple. Besides, implantation is less expensive and time-consuming than MBE.

The process is based on a conventional process and the EPIFAB process [6] developed at the University of Surrey. Ge^+ can be implanted before or after gate oxidation. For the SiGe CMOS/Ge^+ implantation batch, Ge^+ implantation before gate oxidation was chosen as it enables a relatively high peak Ge concentration in the SiGe channel and prevents the contamination of the gate oxide during Ge^+ implantation.

In contrast to previous work [7], a decrease in the MOSFET performance was found with increasing Ge^+ dose. The performance deterioration can be attributed to the following three factors:

1. A slight increase of the oxidation rate for Ge^+ implanted samples reduces the gate drive of Ge^+ implanted MOSFETs in comparison with Si references.

2. Carriers trapped in interface states do not contribute to conduction.

3. The mobility is lowered due to carrier scattering at the degraded silicon interface.

This suggest that the performance could be improved by changing the the annealing conditions

[8, 9]. Furthermore, these findings question whether the Ge^+ implantation of the channel before gate oxidation is really the better alternative as stated in ref. [7].

10.2.3 SiGe Virtual Substrate/Limited Area Process

Unlike the two above mentioned batches, the main emphasis of the SiGe virtual substrate/ limited area process is not put on the realization of an industrial process, but the achievement of high hole mobilities. Such high mobilities can only be achieved if the SiGe channel has a high Ge content and is grown on a virtual substrate [10-13].

The SiGe MOSFETs were subject to an intermediate thermal budget ($750^oC - 800^oC$) in order to obtain a low source and drain series resistance. The wafers have a $< 100 >$ substrate. MBE growth took place on the plain substrate and on microscopic pillars. The growth on pillars can reduce crosshatch and renders the relaxation of the virtual substrate more easy [14, 15].

SiGe MOSFETs fabricated on square microscopic pillars show a considerable performance improvement compared to Si reference MOSFETs and SiGe MOSFETs which were not fabricated on pillars whereas the performance of SiGe MOSFETs fabricated on long, narrow pillars aligned along the [010] direction is only enhanced with respect to Si devices.

The SiGe devices with an LTO gate oxide display a mobility enhancement of more than a factor 2 compared to Si devices. However, their performance is relatively low in comparison with previous work because the Ge content in the channel is far lower than intended. A considerable performance improvements starts for Ge concentrations above 60 % [10], whereas the devices of this batch contain less than 55 % Ge. In contrast, the mobility enhancement of the SiGe devices with a thermal gate oxide is about the same as the one found for comparable devices [12, 13]. For the same split, a lower source and drain series resistance has been observed for SiGe MOSFETs than for the Si reference MOSFETs.

The bad turn-off characteristics of the investigated SiGe devices are probably caused by the presence of p-type dopants in the MBE layer. Unfortunately, the fabricated submicron SiGe MOSFETs are not better than Si devices for a high drain bias. This result can only be explained by velocity saturation.

10.3 Further Work

In general, the SC-1 clean should be omitted for the cleaning of the MBE layer. Future batches should profit from the experience gained during the fabrication of the presented three batches and try to answer the following questions:

- Pseudomorphic SiGe CMOS Process:
  The missing of the Ge channel was realized very late. Should the monitoring of MBE growth and fabrication be improved? For example, spare wafers could be examined after a

mayor processing step. Poly-Si gate deposition is a suitable step. The next step is e-beam lithography and often takes quite long. The SiGe channel of the processed wafers is altered only by diffusion during further processing. The spare wafer can be used for TEM analysis. These results can be obtained long before the batch is completed.

The integration of a thick insitu doped MBE layer into CMOS poses extreme difficulties. What would be the performance of devices with a thin undoped MBE layer?

- SiGe CMOS/Ge^+ Implantation Process:
  Ge^+ implanted MOSFETs are easy to integrate into a CMOS process. However, they can suffer from the degradation of the oxide interface. How much could annealing improve the performance of Ge^+ implanted MOSFETs? Is the implantation of the channel after gate oxidation the better alternative?

- SiGe Virtual Substrate/Limited Area Process:
  The processed MOSFETs did not possess the expected high Ge content in the SiGe channel. What would be the performance of SiGe MOSFETs with a higher Ge content in the channel or even a pure Ge channel?

  A considerable Ge diffusion has been observed. But, the MOSFETs still display a considerable performance improvement. Could a performance improvement as well be obtained for SiGe MOSFETs subject to the thermal budget of a conventional fabrication process?

  The reduced source and drain resistance found for a SiGe MOSFETs with a thermal gate oxide opens interesting perspectives for the application of SiGe virtual substrate devices. Would it be possible to reduce the source and drain series resistance further by modifying the dose of the source and drain implant or for different annealing conditions?

  Submicron SiGe devices measured at high drain voltages could not beat the Si reference devices. How would the SiGe devices perform if the gate length was further scaled down?

  The SiGe MOSFETs did not display good turn-off characteristics. Thus, it is recommended to use an insitu doped MBE layer for future batches.

The characterization of HMOSFETs requires suitable techniques. Split-CV measurements are essential for the mobility extraction of HMOSFETs, charge pumping measurements for the characterization of their oxide interface which may be altered by Ge segregation. S-parameter measurements may be used for the assessment of the speed performance and possibly mobility extraction. Devices suitable for these measurements should be part of any batch.

In general, I recommend to investigate as many problems as possible separately. Although the time gained by using a PMOS process instead of a CMOS process should not be overestimated, presently, it may often be advisable to produce PMOS and NMOS devices separately and to minimize the fabrication time. CMOS requirements interfere to a certain extent with

the requirements for SiGe MOSFETs. Especially, the trial to fulfill threshold requirements for CMOS renders the investigation of SiGe MOSFETs difficult. MBE growth often delays processing considerably.

Batches based on a reliable process can investigate the transport properties of layer structures suitable for integration into CMOS dependent on the thermal budget and optimize the balance between the layer architecture and the thermal budget. Si trial batches can be carried out in order to investigate problems arising from the integration of a particular layer architecture within a given thermal budget into CMOS such as threshold voltage requirements, n- and p-channel behavior. Si CVD can be used in order to simulate the MBE layer in the process flow.

Laser annealing enables the local activation of dopants close to the wafer surface. It is a prospective technique for the SiGe technology for which the thermal budget is critical. Its capability to activate dopants in source, drain and gate without damaging the strained SiGe channel should be assessed. This technique may reduce the thermal budget considerably in conjunction with the uses of novel deposited oxides.

Bibliography

[1] R. M. Sidek, U. N. Straube, A. M. Waite, A. G. R. Evans, C. Parry, P. Phillips, T. E. Whall, and E. H. C. Parker, "SiGe CMOS fabrication using SiGe MBE and anodic/LTO gate oxide", Semicond. Sci. Technol., Vol. 15, pp. 135-138, 2000

[2] D. J. Norris, A. G. Cullis, T. J. Grasby, and E. H. C. Parker, "Investigation of nanoscale Ge segregation in p-channel SiGe/Si field effect transistor structures by electron energy loss imaging", J. Appl. Phys., Vol. 82, No. 12, pp. 7183-7185, 1999

[3] D. J. Godbey, A. H. Krist, K. D. Hobart, and M. E. Twigg, "Selective Removal of $Si_{1-x}Ge_x$ from (100) Si using HNO_3 and HF", J. Electrochem. Soc., Vol. 139, No. 10, pp. 2943-2945, 1992

[4] F. S. Scott, D. S. Miles, D. T. Grider, and J. J. Wortman, "Selective Chemical Etching of Polycrystalline SiGe Alloys with Respect to Si and SiO_2", J. Electronic Mat., Vol. 21, No. 8, pp. 805-810, 1992

[5] U. N. Straube, A. M. Waite, A.G.R Evans, A. Nejim, and P. L.F. Hemment, "Adverse effect of Ge^+ implantation for fabrication of SiGe PMOS", Electronics Lett., Vol. 37, No. 25, pp. 1549-1550, 2001

[6] F. Christiano, A. Nejim, Y. Suprun-Belevich, A. Claverie, and P. L. F. Hemment, "Formation of extended defects and strain relaxation in ion beam synthesised SiGe alloys", Nucl. Instr. and Meth. in Phys. Res. B 147, pp. 35-42, 1999

[7] H. Jiang and R. Elliman, "Electrical Properties of GeSi Surface- and Buried-Channel p-MOSFET's Fabricated by Ge Implantation", IEEE Transactions on Electron Devices, Vol. 43, No. 1, pp. 97-103, 1996

[8] D. Tchikatilov, Y. F. Yang, and E. S. Yang, "Improvement of SiGe oxide grown by electron cyclotron resonance using H_2O vapor annealing", Appl. Phys. Lett., Vol. 69, No. 17, pp. 2578-2580, 1996

[9] C.-G. Ahn, H.-S. KANG, Y.-K. Kwon, and B. Kang, "Effects of Segregated Ge on Electrical Properties of $SiO_2/SiGe$ Interface", Jpn. J. Appl. Phys., Vol. 37, Part 1, No. 3B, pp. 1316-1319, 1998

[10] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, E. A. Fitzgerald, "Hole mobility enhancement in strained $Si/Si_{1-y}Ge_y$ p-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-x}G_x$ $(x < y)$ virtual substrates", Appl. Phys. Lett., Vol. 79, No. 25, pp. 4246-4248, 2001

[11] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Tarasci, and E. A. Fitzgerald, "Carrier mobilities and process stability of strained Si n- and p-MOSFETs on SiGe virtual substrates", J. Vac. Sci. Technol. B, Vol. 19, No. 6, pp. 2268-2279, 2001

[12] G. Hoeck, E. Kohn, C. Rosenblad, H. v. Kaenel, H.-J. Herzog, and U. Koenig, "High hole mobility in $Si_{0.17}Ge_{0.83}$ channel metal-oxide-semiconductor field-effect transistors grown by plasma-enhanced chemical vapor deposition", Appl. Phys. Lett., Vol. 76, No. 26, pp. 3920-3922, 2000

[13] U. Koenig, M. Zeuner, G. Hoeck, T. Hackbarth, M. Glueck, T. Ostermann, and M. Saxarra, "n- and p-Type SiGe HFETs and circuits", Solid-State Electronics, Vol. 43, pp. 1383-1388, 1999

[14] E. Muller, R. Hartmann, C. David, and D. Grutzmacher, "Structural characterization of SiGe step graded buffer layers grown on prestructured Si [100] substrates by molecular beam epitaxy", Thin Solid Films, Vol. 336, pp. 92-95, 1998

[15] R. Hammond, P. J. Phillips, T. E. Whall, E. H. C. Parker, T. Graf, H. von Kanel, and A. J. Shields, "The elimination of surface cross-hatch from relaxed, limited-area $Si_{1-x}Ge_x$ buffer layers", Appl. Phys. Lett., Vol. 71, No. 17, pp. 2517-2519, 1997

Appendix A

Calculation of the Error for Linear Curve Fitting

Both parameter extraction routines (see sections 4.4 and 4.5) are based on the parameter fit for a linear curve. When the error in x is assumed to be negligible [1, chap. 8], the probability of obtaining the observed value y_i for a linear problem ($y_i = A + Bx_i$) is proportional to:

$$Prob_{A,B}(y_i) \propto \frac{1}{\sigma_i} e^{-(y_i - A - Bx_i)^2/2\sigma_i^2} \tag{A.1}$$

Where σ_i corresponds to the standard deviation in y $(\sigma_y)_i$ at a particular measurement point (x_i, y_i) The probability of obtaining the complete set of N measurements is proportional to the product:

$$Prob_{A,B}(y_1, \ldots, y_N) = Prob_{A,B}(y_1) \ldots Prob_{A,B}(y_N) \propto \frac{1}{\sigma_1 \ldots \sigma_N} e^{-\chi^2/2} \tag{A.2}$$

Where the exponent can be expressed:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - A - Bx_i)^2}{\sigma_i^2} \tag{A.3}$$

Applying the method of least-squares fitting, the derivatives of χ^2 with respect to A and B have to be equal to zero:

$$\frac{\partial \chi^2}{\partial A} = \sum_{i=1}^{N} \frac{-2(y_i - A - Bx_i)}{\sigma_i^2} = 0 \tag{A.4}$$

$$\frac{\partial \chi^2}{\partial B} = \sum_{i=1}^{N} \frac{-2x_i(y_i - A - Bx_i)}{\sigma_i^2} = 0 \tag{A.5}$$

Thus, we obtain the two following equations:

$$A \sum_{i=1}^{N} \frac{1}{\sigma_i^2} + B \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} y_i \tag{A.6}$$

$$A \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i + B \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i^2 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i y_i \tag{A.7}$$

Thus, A and B are given by:

$$A = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i^2 \sum_{i=1}^{N} \frac{1}{\sigma_i^2} y_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i y_i}{\Delta} \tag{A.8}$$

$$B = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i y_i - \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i \sum_{i=1}^{N} \frac{1}{\sigma_i^2} y_i}{\Delta} \tag{A.9}$$

$$\Delta = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i^2 - (\sum_{i=1}^{N} \frac{1}{\sigma_i^2} x_i)^2 \tag{A.10}$$

The standard deviations σ_A and σ_B are given by:

$$\sigma_A = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} x^2}{\Delta}} \tag{A.11}$$

$$\sigma_B = \sqrt{\frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}{\Delta}} \tag{A.12}$$

An error Δx in x produces exactly the same effect as an error in y expressed as:

$$\Delta y = \frac{dy}{dx} \Delta x \tag{A.13}$$

As the errors in x and y are independent, they must be combined in quadrature. For a linear problem $y_i = A + Bx_i$, the equivalent standard deviation is therefore given by:

$$\sigma_i = \sqrt{(\sigma_y)_i^2 + (B(\sigma_x)_i^2)} \tag{A.14}$$

The equivalent standard deviation substitutes the standard deviation in x. B can be calculated neglecting the error in x. A more exact result might be found by iteration.
For a particular measurement point, y_i and x_i are given by:

$$y_i = \frac{1}{n} \sum_{j=1}^{n} y_{j,i} \tag{A.15}$$

$$(\sigma_y)_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^{n} (y_{j,i} - y_i)^2} \tag{A.16}$$

$$x_i = \frac{1}{n} \sum_{j=1}^{n} x_{j,i} \tag{A.17}$$

$$(\sigma_x)_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^{n} (x_{j,i} - x_i)^2} \tag{A.18}$$

$(\sigma_y)_i$ and $(\sigma_x)_i$ are the standard deviations for a particular measurement point. n is the number of measurements at a particular measurement point (e. g. channel length), In contrast to the index i that refers to a measurement point, the index j refers to a single measurement.

Bibliography

[1] John R. Taylor, "An introduction to error analysis", University Science Books, second edition, 1997

Appendix B

Process Listings

B.1 Pseudomorphic SiGe CMOS Process

1. Substrate material: 11 wafers, $17 - 33\ \Omega cm$, CZ, $< 100 >$, p-type
2. RCA clean
3. Pad oxidation: 20 nm at 900^oC
4. Photolith mask AL DF
5. Hardbake
6. Dry etch SiO_2: 20 nm
7. Alignment marks: dry etch Si 1 μm deep
8. Resist strip
9. Fuming nitric acid clean
10. Photolith mask NW DF
11. Hardbake
12. N-well implant:
    (a) $2.0 \cdot 10^{13}\ P^+/cm^2$ @ 160 keV
    (b) $4.0 \cdot 10^{12}\ P^+/cm^2$ @ 70 keV
13. Resist strip
14. Photolith mask PW LF
15. Hardbake

16. P-well implant:
    (a) $2.0 \cdot 10^{13}\ B^+/cm^2$ @ 160 keV
    (b) $4.0 \cdot 10^{12}\ B^+/cm^2$ @ 25 keV
17. Resist strip
18. Front spin resist
19. Hardbake
20. Dry etch SiO_2
21. Resist strip
22. RCA
23. RTA: 10 s at 1100^oC
24. MBE growth at the University of Warwick
25. RCA
26. LTO deposition: 20 nm
27. Photolith mask PW LF
28. Hardbake
29. Counterdoping implant for the n-MOSFETs: $2.5 \cdot 10^{13}\ B^+/cm^2$ @ 20 keV
30. Resist strip
31. RCA
32. LTO deposition: 500 nm
33. Photolith mask AA DF
34. Hardbake
35. Active area etch: wet etch SiO_2 until hydrophobic: 20:1 BHF
36. Resist strip
37. RCA
38. Gate oxidation: 6 nm at 800^oC
39. α-Si deposition: 200 nm
40. E-beam lithography mask PF/PC LF

41. Hardbake
42. Gate definition: dry etch α-Si
43. Resist strip
44. RCA
45. Sidewall oxidation: 6 nm at 800^oC
46. Photolith mask NA DF
47. HDD Implant (NMOS): $5.0 \cdot 10^{15}\ As^+/cm^2$ @ 40 keV
48. Resist strip
49. Fuming nitric acid clean
50. Photolith mask PA DF
51. Hardbake
52. HDD Implant (PMOS): $5.0 \cdot 10^{15}\ BF_2^+/cm^2$ @ 45 keV
53. Resist strip
54. RCA clean
55. LTO deposition: 400 nm
56. Frontspin resist
57. Hardbake
58. Wet etch SiO_2 until hydrophobic: 7:1 BHF
59. Resist strip
60. RCA clean
61. RTA
    (a) wafer # 1, 2, 5, 9: 1 min. at 850^oC
    (b) wafer # 3, 6, 10: 3 min. at 850^oC
    (c) wafer # 4, 7, 8, 12: 1 min. at 900^oC
62. Photolith CW DF
63. Hardbake
64. Contact window creation: dry etch SiO_2

65. Resist strip
66. RCA
67. Pre-metal dip etch: 20:1 BHF for 30 s
68. Sputter Ti (100 nm) and Al (1000 nm, 1 % Si)
69. Photolith M1 LF
70. Hardbake
71. Metal wires: dry etch Al/Si + Ti
72. Resist strip
73. Fuming nitric acid clean
74. Alloy anneal: 30 min. at 420^oC

B.2 SiGe CMOS/Ge^+ Implantation Process

1. Substrate material: 12 wafers, $17 - 33\ \Omega cm$, CZ, $< 100 >$, p-type
2. RCA clean
3. Pad oxidation: 20 nm at 900^oC
4. Deposit nitride: 130 nm
5. Photolith mask AA LF
6. Hardbake
7. Dry etch Si_3N_4
8. Resist strip
9. Full clean
10. LOCOS: 400 nm at 1000^oC
11. Dip etch: 20:1 BHF for 30 s
12. Wet etch Si_3N_4: orthophosphoric acid at 160^oC
13. Full clean
14. White ribbon oxidation: 20 nm at 900^o
15. Dip etch: 20:1 BHF

16. RCA clean
17. Pad oxidation:
    (a) wafers # 1-6: 10 nm at 900^oC
    (b) wafers # 7-12: 20 nm at 900^oC
18. Photolith mask PW LF
19. Hardbake
20. P-well implant:
    (a) wafers # 1, 3, 5, 7, 9, 11:
        i. $2.0 \cdot 10^{12}\ B^+/cm^2$ @ 200 keV
        ii. $1.5 \cdot 10^{13}\ B^+/cm^2$ @ 120 keV
        iii. $2.0 \cdot 10^{12}\ B^+/cm^2$ @ 25 keV
    (b) wafers # 2, 4, 6, 8, 10, 12:
        i. $4.0 \cdot 10^{12}\ B^+/cm^2$ @ 200 keV
        ii. $3.0 \cdot 10^{13}\ B^+/cm^2$ @ 120 keV
        iii. $4.0 \cdot 10^{12}\ B^+/cm^2$ @ 25 keV
21. Resist strip
22. Photolith mask NW DF
23. Hardbake
24. N-well implant:
    (a) wafers # 1, 3, 5, 7, 9, 11:
        i. $2.0 \cdot 10^{12}\ P^+/cm^2$ @ 400 keV
        ii. $1.0 \cdot 10^{13}\ P^+/cm^2$ @ 280 keV
        iii. $4.0 \cdot 10^{12}\ P^+/cm^2$ @ 70 keV
    (b) wafers # 2, 4, 6, 8, 10, 12:
        i. $4.0 \cdot 10^{12}\ P^+/cm^2$ @ 400 keV
        ii. $2.0 \cdot 10^{13}\ P^+/cm^2$ @ 280 keV
        iii. $8.0 \cdot 10^{12}\ B^+/cm^2$ @ 70 keV
25. Resist strip
26. Photolith mask NW DF
27. Hardbake

28. Ge implant into the n-well
    (a) wafer # 1, 2, 7, 8: $5 \cdot 10^{15}\ Ge^{+}/cm^{2}$ @ 30 keV
    (b) wafer # 3, 4, 9, 10: $1 \cdot 10^{16}\ Ge^{+}/cm^{2}$ @ 30 keV
    (c) wafer # 5, 6, 11, 12: $2 \cdot 10^{16}\ Ge^{+}/cm^{2}$ @ 30 keV
29. Resist strip
30. Si implant: $6 \cdot 10^{15}\ Si^{+}/cm^{2}$ @ 500 keV
31. Sulphuric/peroxide clean
32. Dip etch until hydrophobic: 20:1 BHF
33. RCA clean
34. Gate oxidation and recrystallization:
    (a) wafer # 1, 3, 5, 7, 9, 11: 12 nm at $900^{o}C$
    (b) wafer # 2, 4, 6, 8, 10, 12: 6 nm at $800^{o}C$
35. α-Si deposition: 250 nm
36. Implant P: $2 \cdot 10^{14}\ P^{+}/cm^{2}$ @ 20 keV
37. RCA clean
38. Dopant activation: 30 min. at $850^{o}C$
39. Fuming nitride acid clean
40. E-beam lithography mask PF/PC LF
41. Hardbake
42. Gate definition: Dry etch α-Si
43. Resist strip
44. Full clean
45. Sidewall oxidation: 6 nm at $800^{o}C$
46. Photolith mask PP DF
47. Hardbake
48. LDD Implant (PMOS): $1.0 \cdot 10^{14}\ BF_{2}^{+}/cm^{2}$ @ 33 keV
49. Resist strip

50. Fuming nitric acid clean
51. Photolith mask NN DF
52. LDD Implant (NMOS): $1.6 \cdot 10^{14}\ P^+/cm^2$ @ 50 keV
53. Resist strip
54. RCA clean
55. LTO deposition for spacers: 180 nm
56. Dry etch SiO_2 to leave fillets
57. Polymer removal
58. Full clean
59. HDD implant oxide: 6 nm at 800^oC
60. Photolith mask NN DF
61. Hardbake
62. S & D implant (NMOS): $5 \cdot 10^{15}\ As^+/cm^2$ @ 60 keV
63. Resist strip
64. Fuming nitric acid
65. Photolith mask PP DF
66. Hardbake
67. S & D implant (PMOS): $5 \cdot 10^{15}\ BF_2^+/cm^2$ @ 45 keV
68. Resist strip
69. RCA clean
70. Post-implant F removal anneal: 20 min. at 650^oC
71. RCA clean
72. Deposit 100 nm undoped Silox & 500 nm BSPG
73. RCA clean
74. RTA for 20 s at 1050^oC
75. Photolith CW DF
76. Hardbake

77. Dry etch SiO_2
78. Resist strip
79. Full clean
80. Pre-metal dip etch: 20:1 BHF for 20 s
81. Sputter 1000 nm Ti-TiN-Al-Si
82. Photolith M1 LF
83. Hardbake
84. Dry Etch Al/Si + Ti
85. Resist strip
86. Alloy anneal: 30 min. at 420^oC

B.3 SiGe virtual substrate/limited area PMOS Process

1. Substrate material: 15 wafers, $17 - 33\ \Omega cm$, CZ, $< 100 >$, p-type
2. RCA clean
3. Wet oxidation: 1000 nm at 1000^oC
4. Photolith mask XTR DF
5. Hardbake
6. Dry etch SiO_2
7. Resist strip
8. Fuming nitric acid clean
9. Trench etch: dry etch SiO_2
10. Strip all oxide: 7:1 BHF
11. RCA clean
12. Pad oxidation: 20 nm at 900^oC
13. N-well implant:
    (a) $2.0 \cdot 10^{13}\ P^+/cm^2$ @ 160 keV
    (b) $4.0 \cdot 10^{12}\ P^+/cm^2$ @ 70 keV

14. RTA for 10 s at 1100^oC
15. MBE growth in Warwick
16. SC-2 clean only
17. LTO deposition: 100 nm
18. poly-Si deposition: 900 nm
19. Dry etch poly-Si to leave only trenches filled
20. LTO deposition: 500 nm
21. Photolith mask AA DF
22. Hardbake
23. Dry etch SiO_2, leave a few nm of oxide in all AA
24. Resist strip
25. Fuming nitric acid clean
26. Dip etch until AA hydrophobic: 20:1 BHF
27. SC-2 clean only
28. Gate Oxide
    (a) wafer # 4, 5, 8, 9, 13: 5 nm at 800^oC
    (b) wafer # 3, 11, 12: 50 nm LTO
29. poly-Si deposition: p-type, insitu, 200 nm
30. RCA clean
31. E-beam lithography mask PF/PC LF
32. Hardbake
33. Gate definition: Dry etch poly-Si
34. Resist strip
35. SC-2 clean only; # 4, 5, 8, 9, 13
36. Sidewall oxidation: 5 nm at 800^oC; # 4, 5, 8, 9, 13
37. Photolith mask PA DF
38. Hardbake

39. LDD Implant: $1.0 \cdot 10^{14}\ BF_2^+/cm^2$ @ 33 keV
40. Resist strip
41. SC-2 clean only
42. LTO deposition for spacers: 180 nm
43. Dry etch SiO_2 to leave fillets
44. Polymer removal
45. Full clean
46. HDD implant oxide: 5 nm at 800^oC; # 4, 5, 8, 9, 13
47. Photolith mask NA DF
48. Hardbake
49. S & D implant: $5 \cdot 10^{15}\ As^+/cm^2$ @ 60 keV
50. Resist strip
51. Fuming nitric acid clean
52. Photolith mask PA DF
53. Hardbake
54. S & D implant: $5 \cdot 10^{15}\ BF_2^+/cm^2$ @ 45 keV
55. Resist strip
56. SC-2 clean only
57. Post-implant F removal anneal: 20 min. at 650^oC
58. LTO deposition: 400 nm
59. RCA clean
60. S & D activation:
    (a) wafer # 4, 5, 8, 9, 13; Furnace anneal: 30 min. at 800^oC
    (b) wafer # 3, 11, 12; Furnace anneal: 30 min. at 750^oC
61. Photolith CW DF
62. Hardbake
63. Dry etch SiO_2

64. Resist strip
65. SC-2 clean only
66. Pre-metal dip etch: 20:1 BHF for 20 s
67. Sputter 1000 nm Ti-TiN-Al-Si
68. Photolith M1 LF
69. Hardbake
70. Dry Etch Al/Si + Ti
71. Resist strip
72. Fuming nitric acid clean
73. Alloy anneal: 30 min. at 420^oC

Appendix C

Comparison between two parameter extraction routines

C.1 Measurements

The parameter extraction routine 1 for conventional MOSFETs (see section 4.4) and the parameter extraction routine 2 for HMOSFETs (see section 4.5) were employed in order to assess the SiGe CMOS/ Ge^+ implantation process (see chapter 8). The measurements based on the routine 1 are discussed in chapter 8. The same data was used for routine 2. The threshold voltage for routine 2 was defined as the voltage for which $I_{ds}(in\ A) \geq 1 \cdot 10^{-7} W/L$. R_{to} was determined at $V_g - V_{th} = -0.5$, -1.0, -1.25, and $-1.5\ V$ for the p-MOSFETs and $V_g - V_{th} = 0.5$, 1.0, 1.5, and $2.0\ V$ for the n-MOSFETs.

C.2 Comparison of Results

R_{sd}, ΔL, and V_{th} extracted with both routines are compared for n- and p-MOSFETs without Ge^+ implant, and Ge^+ implanted p-MOSFETs (see tab. C.1 and C.2):

- R_{sd}:

  Routine 1 uses the following equation for the determination of R_{sd} (see section 4.4):

  $$\Theta = \beta R_{sd} + \Theta_o \tag{C.1}$$

  It leads in general to consistent results for the source and drain series resistance. However, it is based on a mobility model and a constant gate capacitance. Hence, the extracted results depend on the applicability of both assumptions. The constant capacitance approximation is not valid for MOSFETs with a thin gate oxide. As a consequence, β is not an appropriate parameter for their description and can explain the dependence of R_{sd} on the

routine	1	2	1	2	1	2
Wafer	R_{sd} in $k\Omega\mu m$		ΔL in μm		V_{th} in V	
p-MOSFETs						
	6 nm gate oxide					
2	0.89 ± 0.04	0.33 ± 0.90	0.30 ± 0.01	0.22 ± 0.04	-0.91 ± 0.02	-0.80 ± 0.02
4	0.94 ± 0.10	-0.96 ± 0.88	0.20 ± 0.00	0.08 ± 0.04	-0.86 ± 0.04	-0.75 ± 0.03
6	0.79 ± 0.07	-1.01 ± 0.82	0.19 ± 0.01	0.08 ± 0.03	-1.02 ± 0.02	-0.90 ± 0.02
8	0.94 ± 0.07	-2.66 ± 1.35	0.40 ± 0.03	0.16 ± 0.05	-0.93 ± 0.00	-0.85 ± 0.00
10	1.00 ± 0.03	0.90 ± 0.45	0.23 ± 0.01	0.23 ± 0.02	-0.80 ± 0.00	-0.75 ± 0.00
12	1.42 ± 0.06	0.30 ± 0.55	0.21 ± 0.01	0.15 ± 0.02	-1.01 ± 0.01	-0.95 ± 0.02
	12 nm gate oxide					
1	1.37 ± 0.10	1.65 ± 0.76	0.23 ± 0.01	0.23 ± 0.02	-1.00 ± 0.01	-0.95 ± 0.02
3	1.80 ± 0.26	0.56 ± 1.49	0.29 ± 0.02	0.24 ± 0.04	-1.20 ± 0.05	-1.30 ± 0.07
5	1.46 ± 0.17	0.37 ± 1.34	0.39 ± 0.02	0.31 ± 0.04	-1.17 ± 0.01	-1.30 ± 0.09
7	1.17 ± 0.06	1.03 ± 0.57	0.22 ± 0.01	0.22 ± 0.01	-1.19 ± 0.01	-1.20 ± 0.02
9	1.24 ± 0.14	0.29 ± 1.75	0.28 ± 0.02	0.24 ± 0.05	-1.24 ± 0.01	-1.20 ± 0.01
11	1.26 ± 0.01	0.40 ± 0.54	0.23 ± 0.01	0.19 ± 0.01	-1.07 ± 0.00	-1.05 ± 0.00
n-MOSFETs						
6 nm gate oxide						
2	0.52 ± 0.04	0.81 ± 0.22	0.13 ± 0.03	0.21 ± 0.04	0.65 ± 0.01	0.50 ± 0.00
4	0.51 ± 0.01	0.86 ± 0.05	0.08 ± 0.01	0.19 ± 0.01	0.61 ± 0.01	0.45 ± 0.00
6	0.64 ± 0.01	0.93 ± 0.09	0.11 ± 0.01	0.20 ± 0.02	0.75 ± 0.01	0.60 ± 0.01
8	0.59 ± 0.01	0.79 ± 0.07	0.17 ± 0.01	0.23 ± 0.01	0.80 ± 0.01	0.60 ± 0.01
10	0.50 ± 0.03	0.80 ± 0.09	0.06 ± 0.01	0.14 ± 0.02	0.68 ± 0.00	0.50 ± 0.00
12	0.71 ± 0.05	0.09 ± 0.54	0.10 ± 0.01	0.15 ± 0.02	0.82 ± 0.01	0.65 ± 0.01
12 nm gate oxide						
1	0.88 ± 0.04	0.76 ± 0.26	0.14 ± 0.02	0.13 ± 0.03	0.83 ± 0.00	0.70 ± 0.00
3	1.13 ± 0.17	1.00 ± 0.25	0.30 ± 0.26	0.26 ± 0.03	0.80 ± 0.02	0.70 ± 0.02
5	0.88 ± 0.04	0.92 ± 0.18	0.31 ± 0.02	0.32 ± 0.02	0.85 ± 0.03	0.80 ± 0.04
7	0.88 ± 0.04	1.11 ± 0.20	0.19 ± 0.02	0.23 ± 0.02	0.85 ± 0.02	0.80 ± 0.04
9	0.95 ± 0.06	1.31 ± 0.35	0.28 ± 0.03	0.34 ± 0.04	0.79 ± 0.01	0.70 ± 0.02
11	0.87 ± 0.01	1.20 ± 0.10	0.17 ± 0.01	0.23 ± 0.01	0.77 ± 0.00	0.65 ± 0.00

Table C.1: Comparison between R_{sd}, ΔL, and V_{th} extracted for n- and p-MOSFETs without Ge^+ implant with both routines

oxide thickness (see tab. C.1). Besides, unreasonable values can be found for MOSFETs with a Ge^+ implant of $2 \cdot 10^{16}$ $at./cm^2$ because the mobility model does not apply to them.

Routine 2 uses the following equation for the simultaneous determination of ΔL and R_{sd} (see section 4.5):

$$R_{indL} = R_{sd} + \Delta L \cdot R_{depL} \tag{C.2}$$

Exact values for R_{sd} can only be obtained if the difference $R_{indL} - \Delta L \cdot R_{depL}$ is relative large in comparison to the respective error. As this is not the case, even physically unacceptable negative values for R_{sd} are obtained.

- ΔL:

  Similar values for ΔL are found by both routines. The fact that routine 1 is based on a mobility model and a constant gate capacitance does not influence the extraction of ΔL considerably.

- V_{th}:

  The magnitude of V_{th} extracted by routine 1 is usually larger than the one extracted by routine 2 for a 6 nm thick gate oxide. Similar values are extracted for a 12 nm gate oxide (see tab. C.1). This disparity can again be attributed to the mobility model and the constant capacitance approximation used by routine 1.

  In contrast with the extraction of R_{sd} and ΔL, the extraction of V_{th} by routine 2 is also vague. The threshold voltages extracted by routine 2 depend on the drain current and therefore the transport parameters. Consequently, the magnitude of the threshold voltage extracted by routine 1 is even higher than the one extracted by routine 2 for MOSFET with a Ge^+ implant of $2 \cdot 10^{16}$ $at./cm^2$ (see tab. C.2).

C.3 Conclusion

Routine 1 can provide consistent results for R_{sd}, ΔL, and V_{th}. Whereas the impact on the extraction of ΔL is negligible, the constant capacitance approximation and the mobility model seem to influence the extraction of R_{sd} and V_{th}.

Routine 2 leads to reasonable values for ΔL, but cannot extract R_{sd} with the required precision. The extracted threshold voltage depends systematically on the transport properties.

routine	1	2	1	2	1	2
Ge^+dose in $at./cm^2$	R_{sd} in $k\Omega\mu m$		ΔL in μm		V_{th} in V	
6 nm gate oxide						
	10 nm pad oxide					
$5 \cdot 10^{15}$	0.90 ± 0.08	-0.46 ± 0.90	0.26 ± 0.02	0.18 ± 0.03	-1.00 ± 0.03	-0.90 ± 0.03
$1 \cdot 10^{16}$	0.87 ± 0.14	-0.17 ± 0.81	0.10 ± 0.01	0.06 ± 0.03	-1.25 ± 0.03	-1.10 ± 0.03
$2 \cdot 10^{16}$	15.61 ± 1.71	-4.75 ± 4.10	0.11 ± 0.06	0.01 ± 0.03	-1.76 ± 0.01	-1.90 ± 0.04
	20 nm pad oxide					
$5 \cdot 10^{15}$	0.80 ± 0.05	-0.09 ± 7.06	0.24 ± 0.01	0.19 ± 0.03	-1.06 ± 0.03	-0.95 ± 0.03
$1 \cdot 10^{16}$	0.72 ± 0.14	-0.42 ± 1.03	0.15 ± 0.02	0.09 ± 0.04	-1.16 ± 0.02	-1.05 ± 0.01
$2 \cdot 10^{16}$	18.96 ± 0.99	0.05 ± 6.18	0.29 ± 0.03	-0.03 ± 0.05	-1.86 ± 0.01	-1.90 ± 0.03
12 nm gate oxide						
	10 nm pad oxide					
$5 \cdot 10^{15}$	1.05 ± 0.02	2.22 ± 0.41	0.20 ± 0.01	0.24 ± 0.01	-1.13 ± 0.00	-1.10 ± 0.00
$1 \cdot 10^{16}$	1.34 ± 0.09	0.69 ± 0.60	0.23 ± 0.01	0.20 ± 0.02	-1.38 ± 0.03	-1.30 ± 0.03
$2 \cdot 10^{16}$	1.89 ± 0.57	-4.30 ± 1.75	0.26 ± 0.04	0.11 ± 0.02	-1.95 ± 0.01	-2.15 ± 0.04
	20 nm pad oxide					
$5 \cdot 10^{15}$	2.51 ± 0.36	-0.15 ± 1.23	-0.08 ± 0.05	0.15 ± 0.03	-1.38 ± 0.00	-1.35 ± 0.01
$1 \cdot 10^{16}$	1.04 ± 0.06	-0.54 ± 0.83	0.20 ± 0.01	0.15 ± 0.02	-1.65 ± 0.01	-1.55 ± 0.01
$2 \cdot 10^{16}$	1.23 ± 0.25	-0.34 ± 1.82	0.20 ± 0.01	0.06 ± 0.03	-1.75 ± 0.01	-1.85 ± 0.02

Table C.2: Comparison between R_{sd}, ΔL, and V_{th} extracted for Ge^+ implanted p-MOSFETs with both routines

Appendix D

Virtual Substrate/Limited Area PMOS Process - Complete Results

D.1 Measurements

The parameter extraction routine 1 for conventional MOSFETs (see section 4.4) and the parameter extraction routine 2 for HMOSFETs (see section 4.5) were employed in order to assess the virtual substrate/limited area PMOS process (see chapter 9). In chapter 9, only the parameters extracted with routine 1 are given because many devices displayed bad turn-off characteristics rendering the application of routine 2 inaccurate. The same data was used for routine 2.

For wafer # 4, # 8, # 12, # 13, the leakage current, ie the drain current measured for the highest applied gate voltage was subtracted before application of routine 2 (1 V or 2 V). However, this modification was not very successful because the leakage changed as a function of the gate voltage in some cases. Then, the threshold voltage for routine 2 was defined as the voltage for which $I_{ds}(in\ A) \geq 1 \cdot 10^{-7} W/L$. R_{to} was determined at $V_g - V_{th} = -1.0$, -2, -2.5, and -3.0 V for the split with a thermal gate oxide and $V_g - V_{th} = -1$, -2, -3, and -3.5 V for the split with an LTO gate oxide.

This appendix lists all extracted parameters and drain current measurements. The provided information shows the considerable impact of the bad turn-off characteristics on the extracted parameters.

D.2 Extracted Parameters

Wafer no. ; Ge content	V_{th} in V ($L = 1\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$
Wafers with a thermal gate oxide			
5; Si ref.	0.1 ± 0.00	0.07 ± 0.03	2.65 ± 0.28
9; Si ref.	0.1 ± 0.01	0.20 ± 0.03	2.76 ± 0.35
13; 60 % Ge	0.45 ± 0.00	0.14 ± 0.22	6.84 ± 1.70
8; 60 % Ge	0.50 ± 0.02	0.21 ± 0.07	3.29 ± 0.67
4; 70 % Ge	0.50 ± 0.01	0.19 ± 0.03	2.19 ± 0.19
Wafers with an LTO gate oxide			
11; Si ref.	-1.05 ± 0.02	0.07 ± 0.01	2.60 ± 0.32
3; 60 % Ge	-0.20 ± 0.02	0.15 ± 0.02	2.96 ± 0.28
12; 70 % Ge	0.20 ± 0.07	0.17 ± 0.02	3.32 ± 0.28

Table D.1: routine 2, square pillars

Wafer no. ; Ge content	V_{th} in V ($L = 1\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$
Wafers with a thermal gate oxide			
5; Si ref.	0.10 ± 0.00	0.04 ± 0.02	3.24 ± 0.28
9; Si ref.	0.05 ± 0.01	0.19 ± 0.02	2.69 ± 0.27
13; 60 % Ge	0.40 ± 0.02	0.33 ± 0.13	8.72 ± 1.02
8; 60 % Ge	0.45 ± 0.00	0.42 ± 0.20	8.66 ± 1.12
4; 70 % Ge	0.65 ± 0.02	0.4 ± 0.02	2.95 ± 0.12
Wafers with an LTO gate oxide			
11; Si ref.	-1.10 ± 0.02	0.05 ± 0.01	3.44 ± 0.50
3; 60 % Ge	-0.15 ± 0.02	0.11 ± 0.01	3.08 ± 0.22
12; 70 % Ge	0.10 ± 0.02	0.13 ± 0.01	3.45 ± 0.14

Table D.2: routine 2, no square pillars

Wafer no. ; Ge cont.	V_{th} in V ($L = 1\ \mu m$)	Θ_o in $1/V$	ΔL in μm	R_{sd} in $k\Omega\mu m$	μ_o in cm^2/Vs
Wafers with a thermal gate oxide					
5; Si ref.	0.040 ± 0.003	0.419 ± 0.008	0.10 ± 0.01	2.54 ± 0.13	138.52 ± 1.72
9; Si ref.	0.041 ± 0.003	0.399 ± 0.011	0.26 ± 0.00	2.71 ± 0.19	141.45 ± 0.39
13; 60 % Ge	0.396 ± 0.003	0.237 ± 0.095	0.16 ± 0.07	5.92 ± 1.08	161.16 ± 7.44
8; 60 % Ge	0.465 ± 0.006	0.627 ± 0.072	0.36 ± 0.05	2.91 ± 0.52	182.79 ± 5.73
4; 70 % Ge	0.372 ± 0.006	0.462 ± 0.014	0.34 ± 0.02	1.55 ± 0.13	259.68 ± 5.53
Wafers with an LTO gate oxide					
11; Si ref.	-0.917 ± 0.017	0.080 ± 0.003	0.08 ± 0.01	2.37 ± 0.20	235.86 ± 3.36
3; 60 % Ge	-0.224 ± 0.025	0.040 ± 0.002	0.23 ± 0.01	2.44 ± 0.88	412.98 ± 11.23
12; 70 % Ge	-0.041 ± 0.023	0.057 ± 0.006	0.32 ± 0.02	1.82 ± 0.20	478.29 ± 1.65

Table D.3: routine 1, square pillars

Wafer no. ; Ge cont.	V_{th}in$V$ ($L = 1\ \mu m$)	Θ_o in $1/V$	ΔL in μm	R_{sd} in $k\Omega\mu m$	μ_o in cm^2/Vs
Wafers with a thermal gate oxide					
5; Si ref.	0.021 ± 0.004	0.414 ± 0.011	0.09 ± 0.01	2.66 ± 0.16	123.06 ± 2.00
9; Si ref.	0.033 ± 0.005	0.442 ± 0.008	0.23 ± 0.01	2.15 ± 0.20	135.19 ± 1.90
13; 60 % Ge	0.442 ± 0.004	0.112 ± 0.160	0.47 ± 0.11	10.49 ± 2.02	197.39 ± 17.42
8; 60 % Ge	0.541 ± 0.012	-0.039 ± 0.177	1.00 ± 0.15	11.47 ± 2.07	235.79 ± 17.67
4; 70 % Ge	0.423 ± 0.008	0.375 ± 0.017	0.28 ± 0.02	1.88 ± 0.18	217.31 ± 5.08
Wafers with an LTO gate oxide					
11; Si ref.	-0.996 ± 0.016	0.086 ± 0.004	0.09 ± 0.01	2.71 ± 0.34	216.70 ± 4.03
3; 60 % Ge	-0.169 ± 0.015	0.036 ± 0.002	0.16 ± 0.00	3.07 ± 0.12	351.51 ± 4.77
12; 70 % Ge	0.013 ± 0.023	0.052 ± 0.004	0.29 ± 0.01	2.33 ± 0.20	420.29 ± 0.80

Table D.4: routine 1, no square pillars

Wafer no. ; Ge cont.	V_{th} in V ($L = 1.41\ \mu m$)	Θ_o in $1/V$	ΔL in μm	R_{sd} in $k\Omega\mu m$	μ_o in cm^2/Vs
Wafers with a thermal gate oxide					
5; Si ref.	0.035 ± 0.005	0.391 ± 0.005	0.08 ± 0.01	1.90 ± 0.08	141.97 ± 0.38
9; Si ref.	0.039 ± 0.002	0.394 ± 0.001	0.14 ± 0.03	2.23 ± 0.10	133.39 ± 0.98
8; 60 % Ge	0.440 ± 0.006	0.263 ± 0.011	0.36 ± 0.03	7.08 ± 1.01	235.06 ± 1.33
4; 70 % Ge	0.415 ± 0.006	0.390 ± 0.002	0.21 ± 0.02	1.34 ± 0.06	249.32 ± 2.74
Wafers with an LTO gate oxide					
11; Si ref.	-1.081 ± 0.036	0.051 ± 0.003	0.63 ± 0.08	-0.81 ± 0.80	196.87 ± 3.95
12; 70 % Ge	-0.174 ± 0.014	0.042 ± 0.001	0.10 ± 0.02	1.76 ± 0.10	419.35 ± 2.72

Table D.5: routine 1, long, 14.85 μm wide pillars, three devices

Wafer no. ; Ge cont.	V_{th} in V ($L = 1.41\ \mu m$)	Θ_o in $1/V$	ΔL in μm	R_{sd} in $k\Omega\mu m$	μ_o in cm^2/Vs
Wafers with a thermal gate oxide					
5; Si ref.	0.035 ± 0.005	0.392 ± 0.005	0.08 ± 0.01	1.88 ± 0.84	142.02 ± 0.38
9; Si ref.	0.039 ± 0.002	0.396 ± 0.001	0.08 ± 0.03	2.20 ± 0.11	128.32 ± 2.51
8; 60 % Ge	0.440 ± 0.006	0.255 ± 0.031	0.37 ± 0.03	7.15 ± 1.54	235.84 ± 1.43
4; 70 % Ge	0.415 ± 0.006	0.384 ± 0.003	0.17 ± 0.03	1.43 ± 0.06	243.07 ± 4.83
Wafers with an LTO gate oxide					
11; Si ref.	-1.081 ± 0.036	0.045 ± 0.004	0.45 ± 0.09	0.19 ± 0.94	182.14 ± 5.85
12; 70 % Ge	-0.174 $\pm$0.014	0.043 ± 0.001	0.09 ± 0.02	1.72 ± 0.09	416.68 ± 3.41

Table D.6: routine 1, long, 14.85 μm wide pillars, two devices

Wafer no. ; Ge cont.	V_{th} in V ($L = 1.41\ \mu m$)	Θ_o in $1/V$	ΔL in μm	R_{sd} in $k\Omega\mu m$	μ_o in cm^2/Vs
Wafers with a thermal gate oxide					
5; Si ref.	0.016 ± 0.005	0.407 ± 0.001	0.05 ± 0.01	1.72 ± 0.04	140.49 ± 0.94
9; Si ref.	0.040 ± 0.002	0.399 ± 0.004	0.10 ± 0.03	2.31 ± 0.14	133.23 ± 2.55
8; 60 % Ge	0.398 ± 0.008	0.250 ± 0.023	0.21 ± 0.10	6.21 ± 1.58	193.37 ± 2.89
4; 70 % Ge	0.347 ± 0.006	0.413 ± 0.002	0.17 ± 0.02	1.35 ± 0.06	235.29 ± 1.19
Wafers with an LTO gate oxide					
11; Si ref.	-1.189 ± 0.040	0.050 ± 0.006	0.44 ± 0.09	1.34 ± 1.23	177.00 ± 8.63
12; 70 % Ge	-0.210 ± 0.012	0.046 ± 0.001	0.07 ± 0.03	1.62 ± 0.28	389.27 ± 4.20

Table D.7: routine 1, long, 7.07 μm wide pillars

Wafer no. ; Ge cont.	V_{th} in V ($L = 1.41\ \mu m$)	Θ_o in $1/V$	ΔL in μm	R_{sd} in $k\Omega\mu m$	μ_o in cm^2/Vs
Wafers with a thermal gate oxide					
5; Si ref.	0.028 ± 0.006	0.403 ± 0.001	0.05 ± 0.01	1.80 ± 0.06	143.16 ± 0.82
9; Si ref.	0.038 ± 0.002	0.402 ± 0.001	0.11 ± 0.03	2.27 ± 0.09	141.39 ± 1.32
8; 60 % Ge	0.460 ± 0.006	0.236 ± 0.034	0.34 ± 0.05	8.89 ± 1.72	230.51 ± 3.09
4; 70 % Ge	0.435 ± 0.003	0.384 ± 0.020	0.37 ± 0.04	1.51 ± 0.20	283.65 ± 7.49
Wafers with an LTO gate oxide					
11; Si ref.	-1.055 ± 0.021	0.050 ± 0.007	0.37 ± 0.11	-0.35 ± 1.36	186.06 ± 11.61
12; 70 % Ge	-0.161 ± 0.022	0.041 ± 0.001	0.05 ± 0.02	1.99 ± 0.14	415.20 ± 3.64

Table D.8: routine 1, no pillars

Wafer no. ; Ge content	V_{th} in V ($L = 1.41\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$
Wafers with a thermal gate oxide			
5; Si ref.	0.05 ± 0.02	0.02 ± 0.03	2.51 ± 0.26
9; Si ref.	0.1 ± 0.00	0.13 ± 0.06	2.34 ± 0.65
8; 60 % Ge	0.50 ± 0.01	1.47 ± 0.36	4.06 ± 1.92
4; 70 % Ge	0.70 ± 0.01	-0.38 ± 0.05	3.86 ± 0.28
Wafers with an LTO gate oxide			
11; Si ref.	-1.25 ± 0.02	0.31 ± 0.06	5.83 ± 1.93
12; 70 % Ge	-0.05 ± 0.02	0.02 ± 0.02	2.55 ± 0.43

Table D.9: routine 2, long, 14.85 μm wide pillars

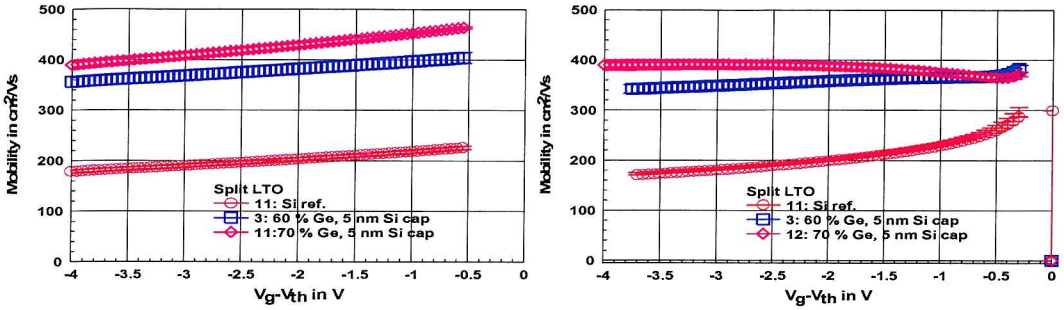
Wafer no. ; Ge content	V_{th} in V ($L = 1.41\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$
Wafers with a thermal gate oxide			
5; Si ref.	0.1 ± 0.01	-0.01 ± 0.04	2.16 ± 0.37
9; Si ref.	0.1 ± 0.00	0.10 ± 0.07	2.31 ± 0.71
8; 60 % Ge	0.45 ± 0.00	0.17 ± 0.52	5.38 ± 3.39
4; 70 % Ge	0.45 ± 0.01	0.12 ± 0.03	1.55 ± 0.13
Wafers with an LTO gate oxide			
11; Si ref.	-1.25 ± 0.03	0.31 ± 0.11	1.21 ± 3.58
12; 70 % Ge	-0.05 ± 0.01	-0.02 ± 0.03	2.44 ± 0.52

Table D.10: routine 2, long, 7.07 μm wide pillars

Wafer no. ; Ge content	V_{th} in V ($L = 1.41\ \mu m$)	ΔL in μm	R_{sd} in $k\Omega\mu m$
Wafers with a thermal gate oxide			
5; Si ref.	0.05 ± 0.02	0.01 ± 0.03	2.21 ± 0.24
9; Si ref.	0.10 ± 0.00	0.06 ± 0.05	2.64 ± 0.49
8; 60 % Ge	0.45 ± 0.01	0.24 ± 0.51	8.37 ± 3.34
4; 70 % Ge	0.85 ± 0.04	-0.04 ± 0.09	3.05 ± 0.43
Wafers with an LTO gate oxide			
11; Si ref.	-1.20 ± 0.01	0.16 ± 0.12	4.22 ± 4.11
12; 70 % Ge	0.05 ± 0.04	-0.11 ± 0.04	3.44 ± 0.69

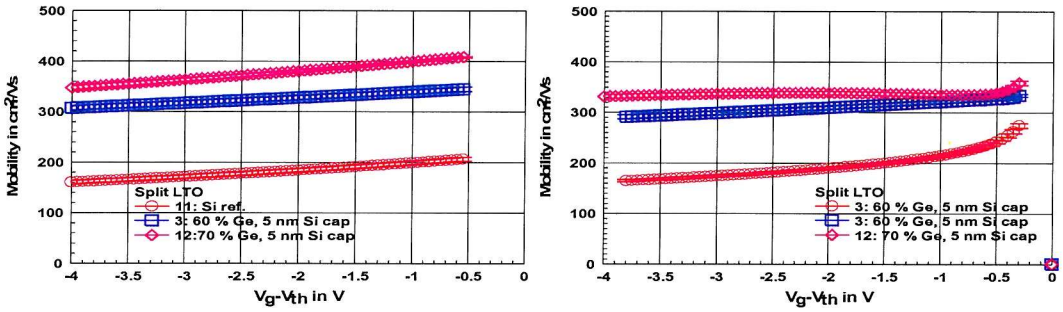
Table D.11: routine 2, no long pillars

D.3 Extracted Mobilities as a Function of the Gate Drive



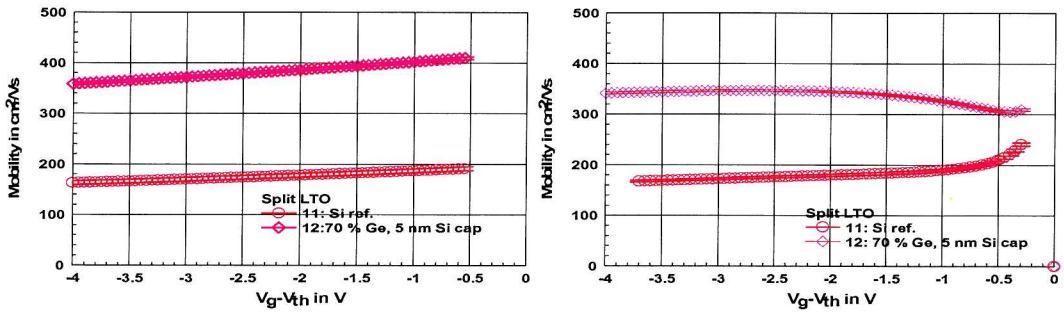
(a) (b)

Figure D.1: Mobility for the MOSFETs with a LTO gate oxide fabricated on square pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



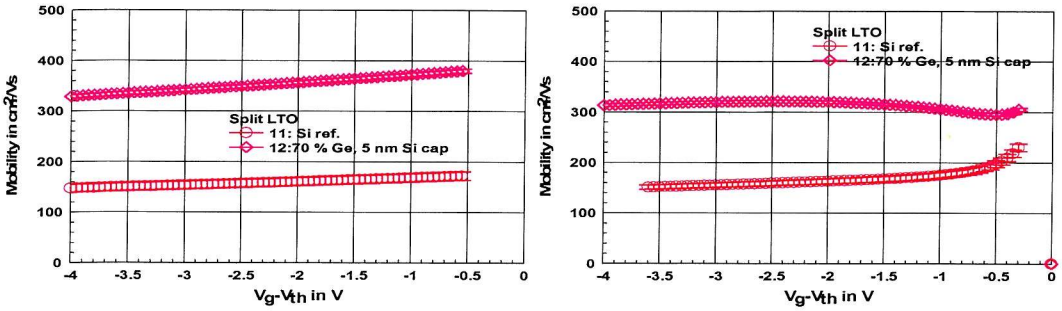
(a) (b)

Figure D.2: Mobility for the MOSFETs with an LTO gate oxide not fabricated on square pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



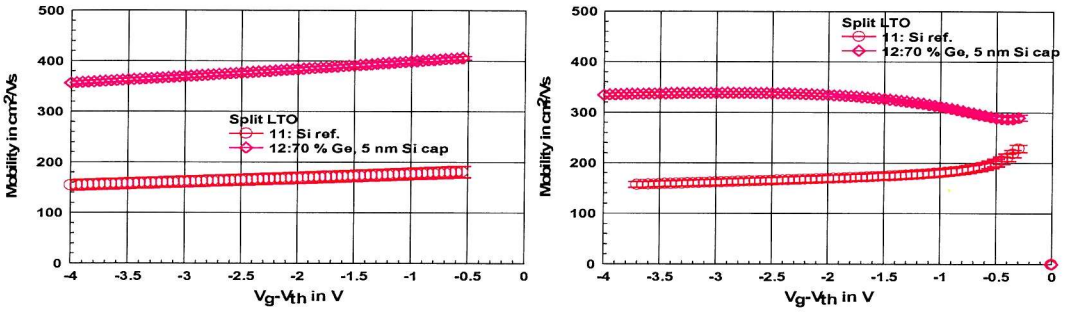
(a) (b)

Figure D.3: Mobility for the MOSFETs with an LTO gate oxide fabricated on long, 14.85 μm wide pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



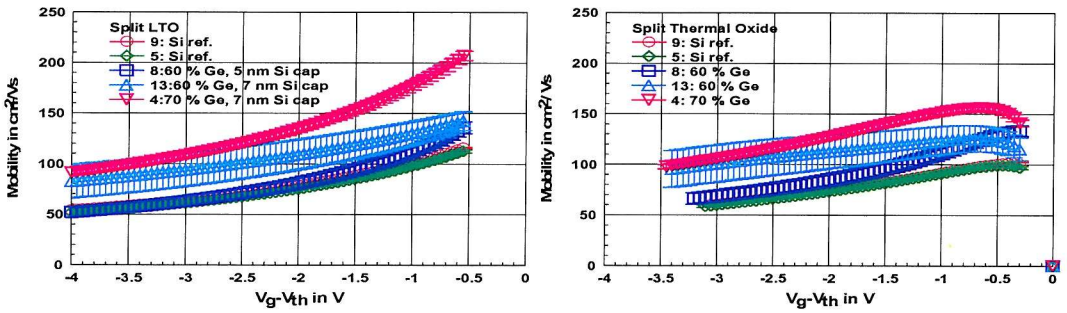
(a) (b)

Figure D.4: Mobility for the MOSFETs with an LTO gate oxide fabricated on long, 7.07 μm wide pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



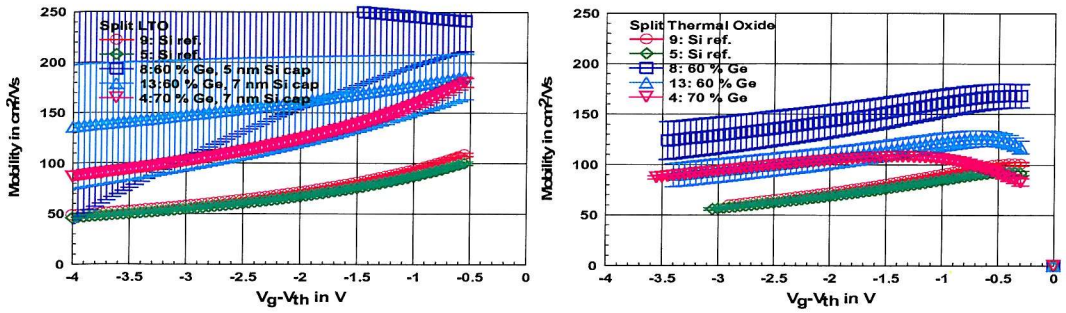
(a) (b)

Figure D.5: Mobility for the MOSFETs with an LTO gate oxide not fabricated on long pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



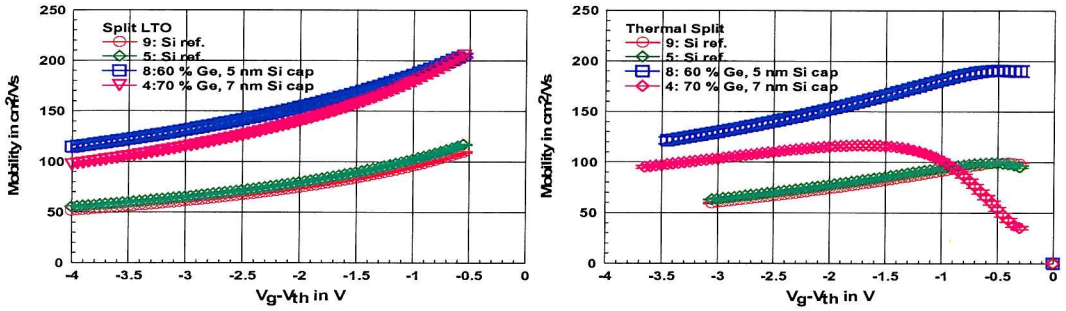
(a) (b)

Figure D.6: Mobility for the MOSFETs with a thermal gate oxide fabricated on square pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



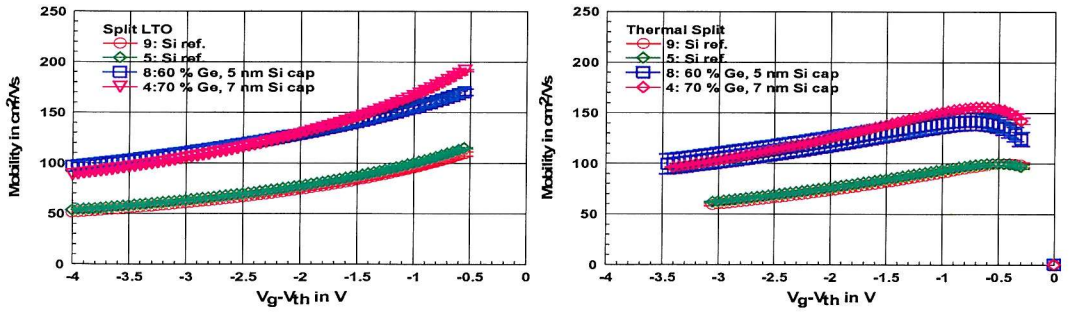
(a) (b)

Figure D.7: Mobility for the MOSFETs with a thermal gate oxide not fabricated on square pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



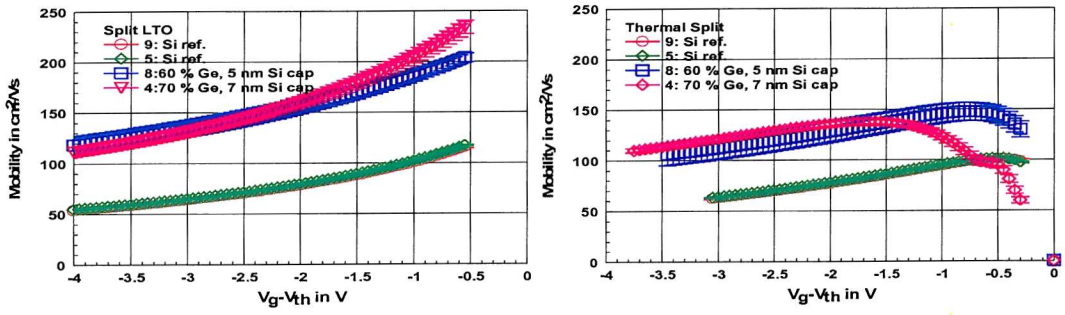
(a) (b)

Figure D.8: Mobility for the MOSFETs with a thermal gate oxide fabricated on long, 14.85 μm wide pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



(a) (b)

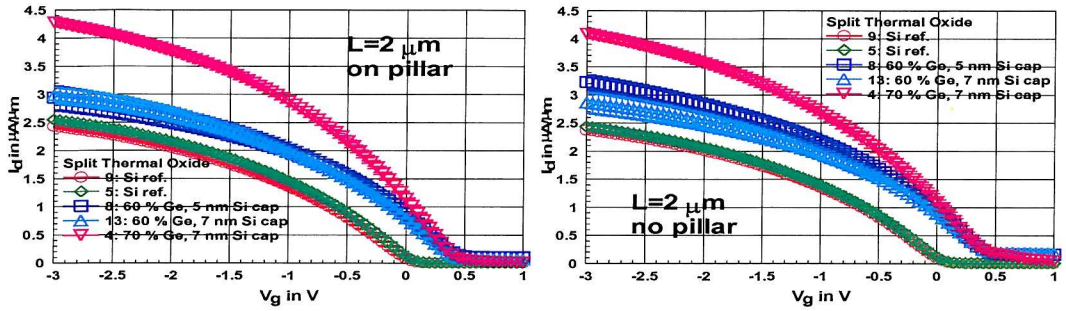
Figure D.9: Mobility for the MOSFETs with a thermal gate oxide fabricated on long, 7.07 μm wide pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2



(a) (b)

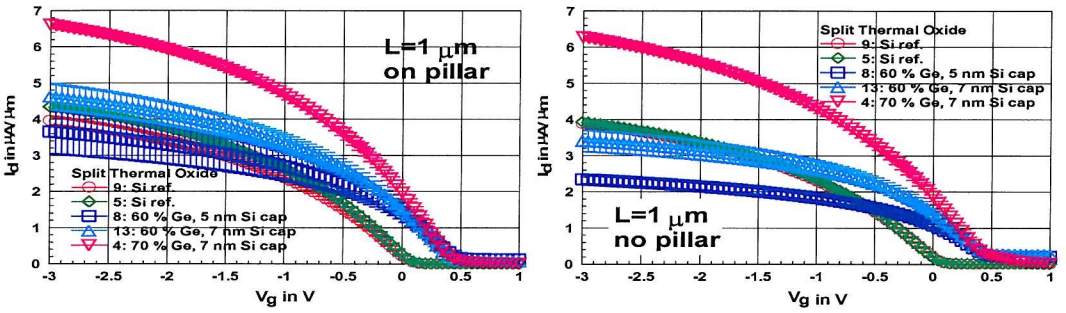
Figure D.10: Mobility for the MOSFETs with a thermal gate oxide not fabricated on long pillars as a function of the gate drive extracted by (a) routine 1 (b) routine 2

D.4 Drain Current Measurements for $V_d = -50\ mV$



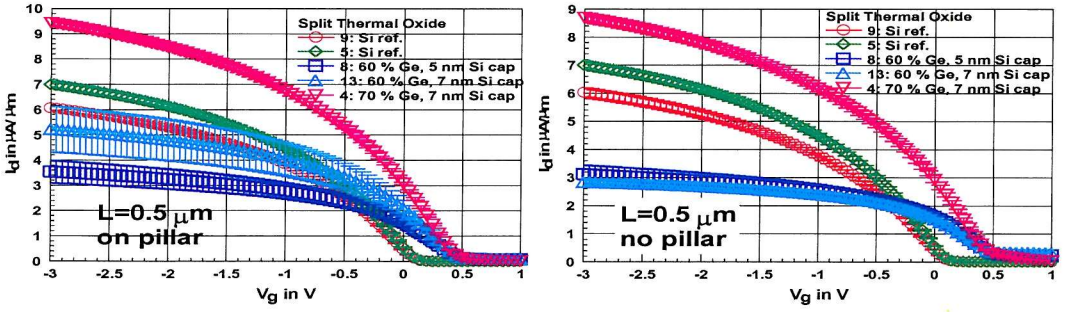
(a) (b)

Figure D.11: p-MOSFETs ($L = 2\ \mu m$) with thermal gate oxide (a) on pillar (b) no pillar



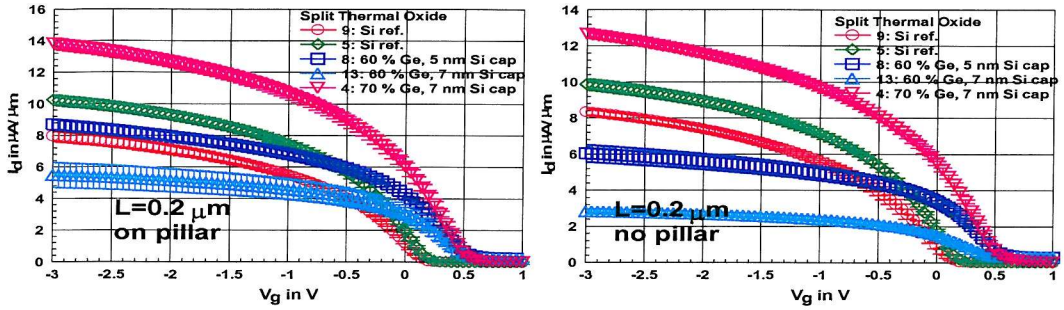
(a) (b)

Figure D.12: p-MOSFETs ($L = 1\ \mu m$) with thermal gate oxide (a) on pillar (b) no pillar



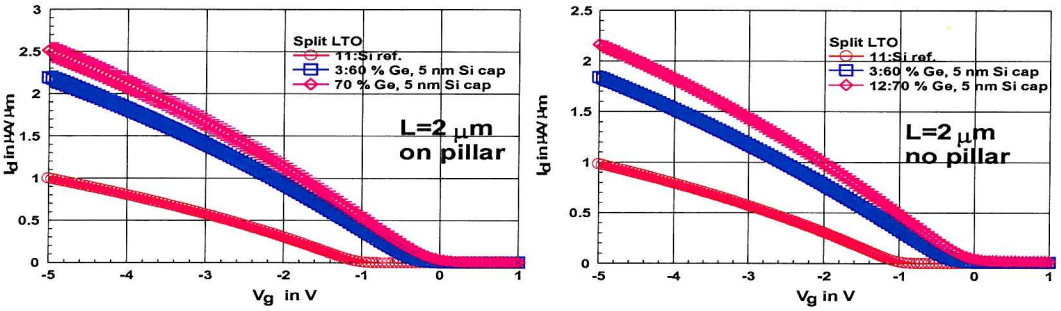
(a) (b)

Figure D.13: p-MOSFETs ($L = 0.5\ \mu m$) with thermal gate oxide (a) on pillar (b) no pillar



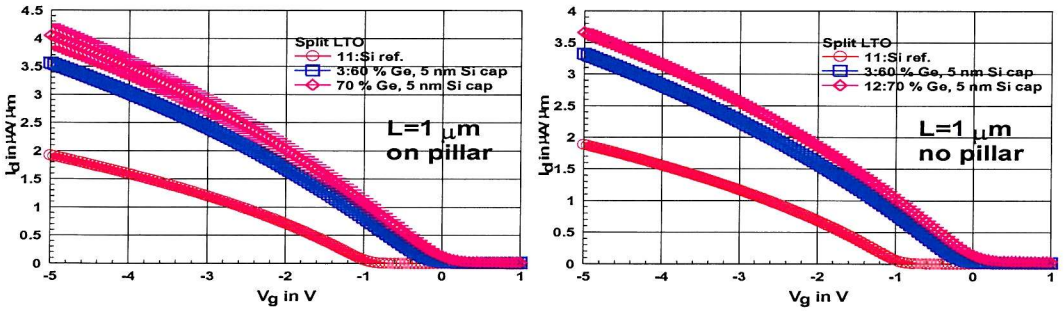
(a) (b)

Figure D.14: p-MOSFETs ($L = 0.2\ \mu m$) with thermal gate oxide (a) on pillar (b) no pillar



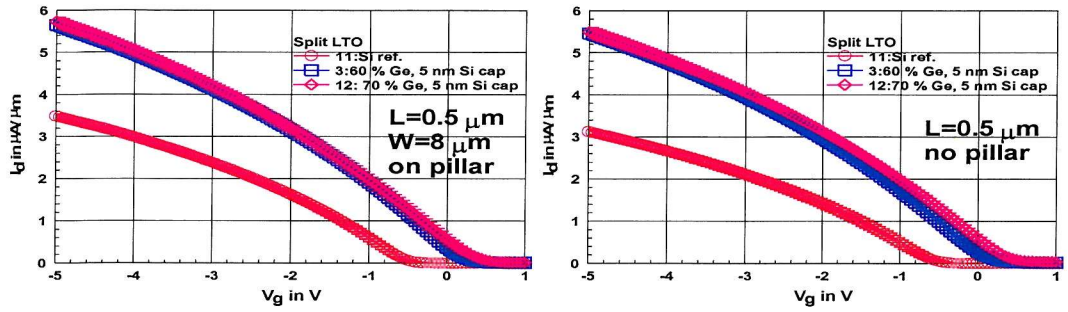
(a) (b)

Figure D.15: p-MOSFETs ($L = 2\ \mu m$) with LTO gate oxide (a) on pillar (b) no pillar



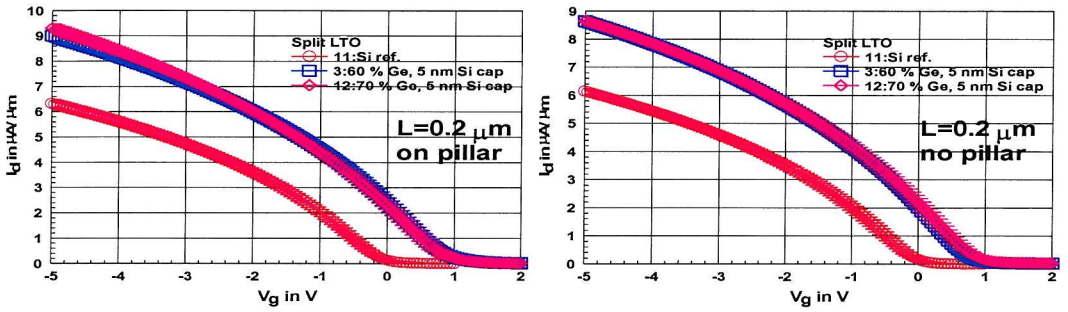
(a) (b)

Figure D.16: p-MOSFETs ($L = 1\ \mu m$) with LTO gate oxide (a) on pillar (b) no pillar



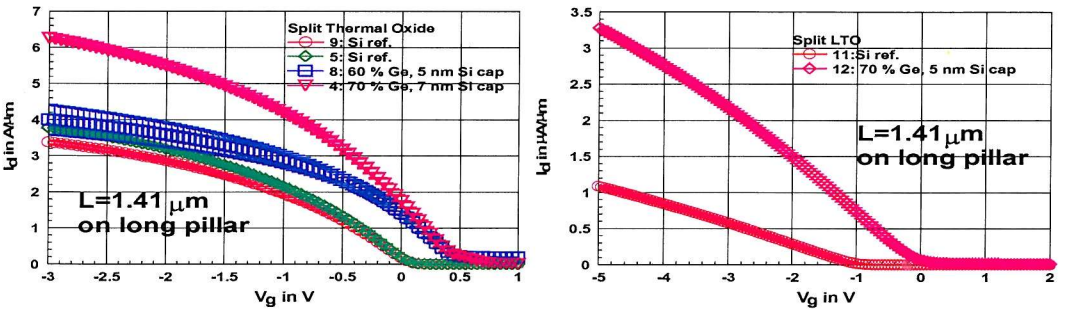
(a) (b)

Figure D.17: p-MOSFETs ($L = 0.5\ \mu m$) with LTO gate oxide (a) on pillar (b) no pillar



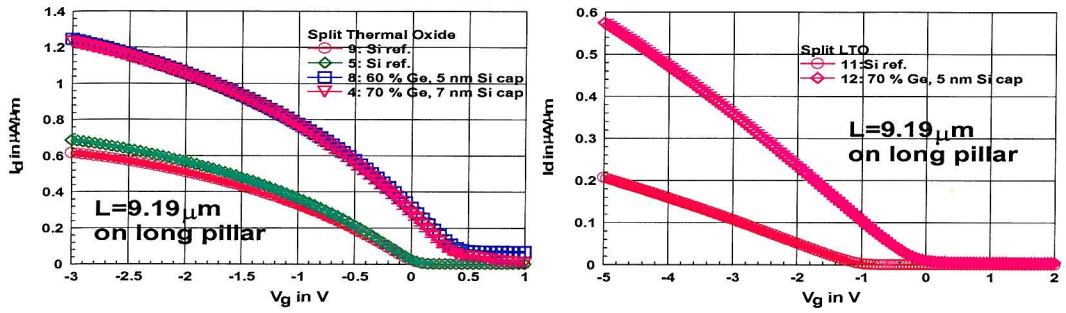
(a) (b)

Figure D.18: p-MOSFETs ($L = 0.2\ \mu m$) with LTO gate oxide (a) on pillar (b) no pillar



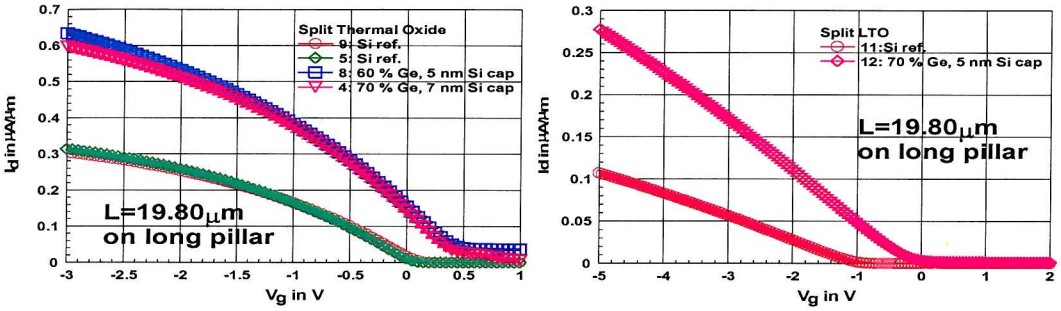
(a) (b)

Figure D.19: p-MOSFETs ($L = 1.41\ \mu m$) on a long, 14.85 μm wide pillar (a) Thermal Oxide (b) LTO



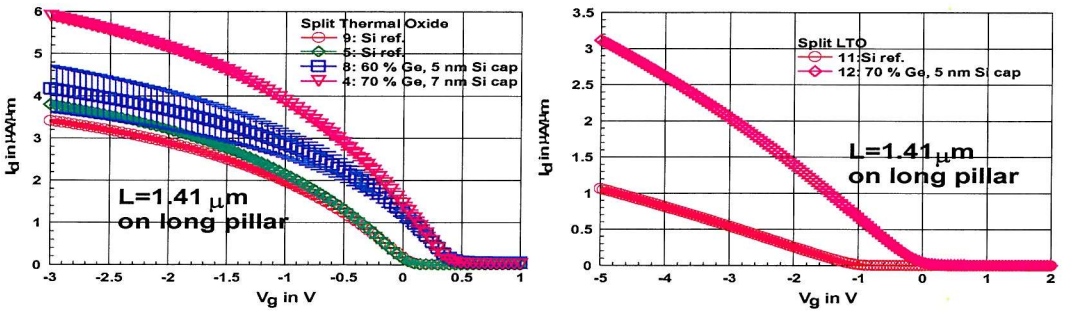
(a) (b)

Figure D.20: p-MOSFETs ($L = 9.19\ \mu m$) on a long, 14.85 μm wide pillar (a) Thermal Oxide (b) LTO



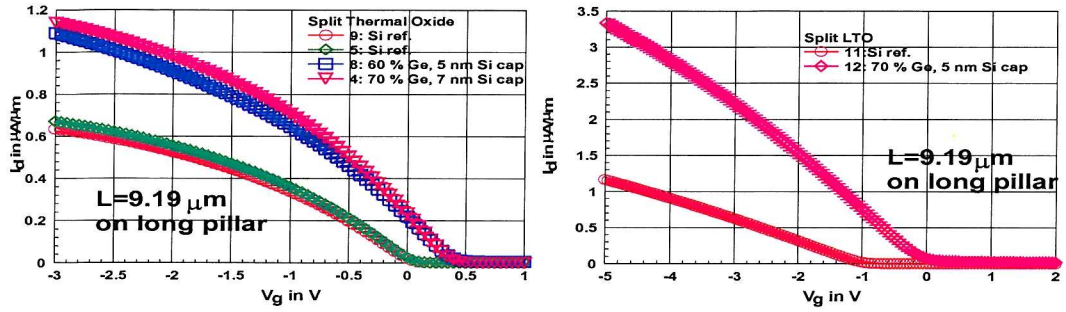
(a) (b)

Figure D.21: p-MOSFETs ($L = 19.80\ \mu m$) on a long, 14.85 μm wide pillar (a) Thermal Oxide (b) LTO



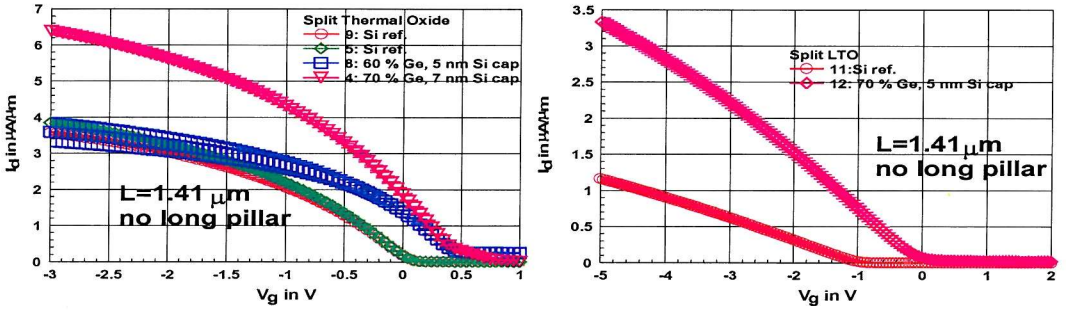
(a) (b)

Figure D.22: p-MOSFETs ($L = 1.41\ \mu m$) on a long, 7.07 μm wide pillar (a) Thermal Oxide (b) LTO



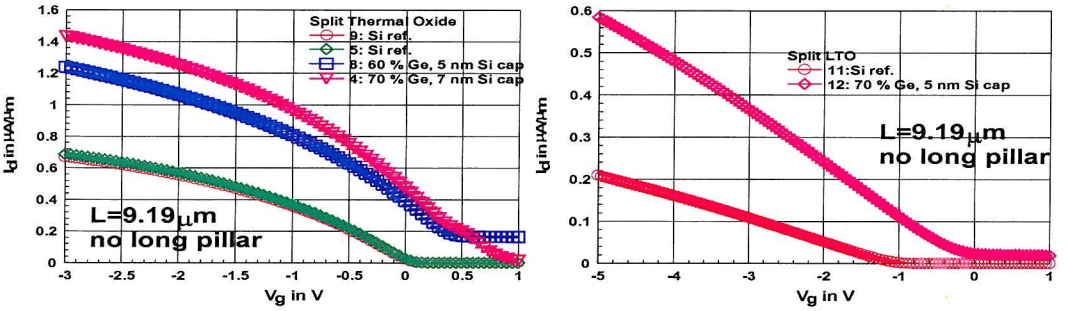
(a) (b)

Figure D.23: p-MOSFETs ($L = 9.19\ \mu m$) on a long, 7.07 μm wide pillar (a) Thermal Oxide (b) LTO



(a) (b)

Figure D.24: p-MOSFETs ($L = 1.41\ \mu m$) on no long pillar (a) Thermal Oxide (b) LTO



(a) (b)

Figure D.25: p-MOSFETs ($L = 9.19\ \mu m$) on no long pillar (a) Thermal Oxide (b) LTO

Appendix E

List of Publications

1. R. M. Sidek, U. N. Straube, A. M. Waite, A. G. R. Evans, C. Parry, P. Phillips, T. E. Whall, and E. H. C. Parker, "SiGe CMOS fabrication using SiGe MBE and anodic/LTO gate oxide", Semicond. Sci. Technol., Vol. 15, pp. 135-138, 2000

2. U. N. Straube, A. G. R. Evans, G. Braithwaite, S. Kaya, J. Watling, and A. Asenov, "On the mobility extraction for HMOSFETs", Sol.-St. Electron., Vol. 45, pp. 527-529, 2000

3. U. N. Straube, A. M. Waite, A.G.R Evans, A. Nejim, and P. L.F. Hemment, "Adverse effect of Ge^+ implantation for fabrication of SiGe PMOS", Electronics Lett., Vol. 37, No. 25, pp. 1549-1550, 2001

4. J. Bonar, G. D. M. Dilliway, U. N. Straube, B. Ault, A. Brunschweiler, and A. G. R. Evans, "n-MOSFETs fabricated on SiGe virtual substrates grown by LPCVD", 4th International Conference on Materials for Microelectronics and Nanoengineering, 10-12 June 2002, Finland

5. Further conference paper about the SiGe Virtual Substrate/Limited Area PMOS Process to be submitted at ESSDERC 2002

6. Further journal paper about the SiGe Virtual Substrate/Limited Area PMOS Process to be submitted