Interface Characterization of Metal-Gate MOS-Structures

and the Application to DRAM-Capacitors

Vom Promotionsausschuss der
Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation

von

Bernhard Sell

aus Hamburg

2002

Erster Gutachter:	Prof. Dr. Wolfgang Krautschneider
Zweiter Gutachter:	Prof. Dr. Wolfgang Albrecht
Tag der mündlichen Prüfung:	17.06.2002

Contents

List of Symbols

Symbol	Description	Unit
A_{G}	gate area	cm^2
C_{B}	bitline capacitance	F
C_{diel}	dielectric capacitance	F
C_{FB}	flatband capacitance	F
C_{HF}	high frequency capacitance	F
C_{inv}	minimum (strong inversion) capacitance	F
C_{LF}	low frequency capacitance	F
C_{ox}	oxide capacitance	F
C_{s}	storage capacitance	F
D	dissipation factor	
D_j	diffusion coefficient	$\mathrm{m}^2\ \mathrm{s}^{-1}$
D_{it}	interface trap charge density	$\mathrm{cm}^{-2}eV^{-1}$
ε_{d}	dielectric permittivity	
ε_{s}	silicon permittivity	
E_{c}	conduction band edge	eV
E_{g}	silicon band gap	eV
E_j	activation energy	J/mol
E_{ox}	oxide electric field	V/cm
E_{Si}	silicon electric field	V/cm
E_{v}	valence band edge	eV
ϕ_{B}	barrier height	V
ϕ_{F}	Fermi potential	V
ϕ_{M}	metal work function	V
ϕ_{MS}	metal semiconductor work function	V
ϕ_{S}	semiconductor work function	V
ϕ_{s}	surface potential	V
ϕ_{t}	dielectric trap potential	V
f	measurement frequency	Hz
F	minimum half pitch	
γ	charge distribution factor	
I_{Bulk}	bulk current of a MOSFET	A
I_{cp}	charge pumping current	A
I_{DS}	drain current of a MOSFET	A
I_{G}	gate current of a MOSFET	A
J_{FN}	Fowler-Nordheim tunnel current density	$\mathrm{A/cm}^2$

λ	mean free path	cm
L	channel length of a MOSFET	cm
L_{eff}	effective channel length of a MOSFET	cm
μ_{eff}	effective mobility	$\text{cm}^2/\text{V}\cdot\text{s}$
μ_{n}	electron mobility	$\text{cm}^2/\text{V}\cdot\text{s}$
m_{ox}	parabolic electron mass in SiO_2	
m_{Si}	electron mass in silicon	
M	molar mass	kg
n	electron density	cm^{-3}
n_{i}	intrinsic carrier density	cm^{-3}
N_{A}	acceptor doping density	cm^{-3}
N_{c}	conduction band density of states	cm^{-3}
N_{f}	fixed oxide charge density	cm^{-2}
N_{v}	valence band density of states	cm^{-3}
p	hole density	cm^{-3}
p_{s}	hole density at surface	cm^{-3}
Q_{cp}	charge pumping charge	C
Q_{d}	depletion charge density of a MOS capacitor	C/cm^2
Q_{f}	fixed interface charge density	C/cm^2
Q_{i}	inversion charge density of a MOS capacitor	C/cm^2
Q_{it}	interface state charge density	C/cm^2
Q_{m}	mobile oxide charge density	C/cm^2
Q_{ot}	oxide trapped charge density	C/cm^2
ρ	density	g cm^{-3}
R_j	deposition rate	cm/s
R_{s}	series resistance	Ω
R_{SD}	source drain resistance of a MOSFET	Ω
σ_{n}	electron capture cross section	cm^2
σ_{p}	hole capture cross section	cm^2
S	subthreshold slope of a MOSFET	V/decade
τ_{c}	capture time constant	s
$\tau_{\text{em,e}}$	electron emission time constant	s
$\tau_{\text{em,h}}$	hole emission time constant	s
t_{diel}	physical dielectric thickness	cm
t_{ox}	physical oxide thickness	cm
T	temperature	K
T_{s}	surface temperature	K
v_{th}	thermal velocity	cm/s
V_{DS}	drain source voltage of a MOSFET	V
V_{FB}	flatband voltage	V
V_{G}	gate voltage	V
V_{GS}	gate source voltage of a MOSFET	V
V_{SB}	source bulk voltage of a MOSFET	V
V_{T}	threshold voltage of a MOSFET	V
ω	radial frequency	s^{-1}
W_{eff}	effective channel width of a MOSFET	cm

Abstract

Many devices of today's information technology like personal computers and personal digital assistants require a memory with fast read- and write-times and a low cost per bit. State of the art DRAMs are the only mature circuits fulfilling these requirements. Continuous shrinking of device dimensions is necessary to improve the productivity and hence to reduce the price per bit even further. At the same time, the storage capacitance of a DRAM cell remains the same for every generation which leads to three dimensional capacitors with very high aspect ratios and therefore to high series resistances. According to the ITRS roadmap, the commonly used poly-crystalline silicon will have to replaced by a higher conductive material from the 70 nm generation onwards.

In this work, a stack of poly-crystalline silicon and thin metal is proposed as a suitable replacement for pure silicon. A full analysis procedure for metal-gate MOS-structures is developed that allows to identify problems during processing and to extract physical parameters of metal electrodes. Quantities gained from device simulations of MOS-capacitors enable an automatic extraction of flatband potential and physical oxide thickness within an accuracy of 1-2 Å as compared to ellipsometric measurements and IV-analysis. An improved model for the gate-leakage current is presented that describes the measurement data accurately for all voltages and for all gate oxide thicknesses under study. To apply the analysis procedure, a new test chip and pertinent process technology is developed which enables the fabrication of metal-gate MOS-structures on short loops as well as on fully-integrated wafers. TiN-gate structures are successfully used to characterize a TiN-atomic layer deposition process, to identify process problems and to eliminate them.

The thermal stability of pure metal gate-electrodes is found to be insufficient for deep-trench DRAM applications. A stack of 25 nm poly-crystalline silicon and 20 nm TiN is identified as a suitable low-resistance deep trench fill that is thermally stable up to 1050 °C. To integrate this stack into DRAMs, the metal has to be deposited into deep trenches with a good step coverage. A tungsten silicide chemical vapor deposition process is investigated with the result that the major challenge is identified as the simultaneous reaching of good step-coverage and a thermally stable composition. A regime for such a deposition process is proposed which, however, lies beyond the parameter space available with the tool employed in this study. On the other hand, a TiN atomic layer deposition process is successfully modified to give 70% step coverage in trenches with an aspect ratio of 40:1. Using this process, deep trench short loops are fabricated that withstand a 60 s anneal at 1050 °C while reaching the target values for the maximum allowed leakage current and the minimum capacitance. In summary, results presented in this thesis enable the fabrication of low-resistance deep trench fills for deep trench DRAM-capacitors.

Zusammenfassung

Viele Applikationen der modernen Informationstechnologie benötigen einen Arbeitsspeicher mit kurzen Schreib- und Lesezeiten sowie geringen Kosten pro Bit. Moderne DRAMs sind zur Zeit die einzigen Schaltkreise, die diese Anforderungen erfüllen. Eine kontinuierliche Verkleinerung der Strukturabmessungen ist für die Erhöhung der Produktivität und somit für eine weitere Verringerung der Kosten pro Bit erforderlich. Gleichzeitig bleibt jedoch die Kapazität der DRAM Speicherzelle mit jeder Generation konstant, was zu dreidimensionalen Kondensatoren mit hohen Aspektverhältnissen und hohen Reihenwiderständen führt. Nach der ITRS Roadmap muß daher ab der 70 nm Generation das üblicherweise verwendete polykristalline Silizium durch ein Material mit höherem Leitwert ersetzt werden.
In dieser Arbeit wird ein Stapel aus polykristallinem Silizium und einer dünnen Metallschicht als geeigneter Ersatz für das reine Silizium vorgeschlagen. Es wird eine vollständige Analyse-Prozedur für Metall-Gate MOS-Strukturen entwicklet, die die Identifizierung von Problemen während der Prozessierung und die Extraktion physikalischer Parameter ermöglicht. Aus Simulationen von MOS-Kondensatoren werden Größen gewonnen, die eine automatische Extraktion der Flachbandspannung und der physikalischen Oxyddicke zulassen und dabei eine Genauigkeit von 1-2 Å verglichen mit ellipsometrischen Messungen und IV-Analysen zeigen. Weiterhin wird ein verbessertes Modell zur Simulation des Leckstroms präsentiert welches die Meßdaten für alle untersuchten Spannungen und Oxyddicken sehr gut beschreibt. Zur Anwendung der Analyse-Prozedur werden ein Testchip und dazugehörige Prozeßtechnologie für Metall-Gate MOS-Strukturen auf Kurzdurchläufern und auf vollintegrierten Losen entwicklet. TiN-Gate Strukturen werden benutzt, um einen Prozeß zur TiN Atomlagen-Abscheidung zu charakterisieren, Probleme zu identifizieren und diese zu lösen.
Die thermische Stabilität reiner Metall-Gate Elektroden ist für eine DRAM Anwendung mit Grabenkondensatoren nicht ausreichend. Ein Stapel von 25 nm polykristallinem Silizium und 20 nm TiN wird als geeignete niederohmige Grabenfüllung identifiziert, die bis zu 1050 °C stabil ist. Um solch ein Materialsystem in einen DRAM zu integrieren muß das Metall in tiefen Gräben mit guter Kantenbedeckung abgeschieden werden. Zu diesem Zweck wird die Gasphasenabscheidung von Wolfram Silizid untersucht und das gleichzeitige Erreichen von guter Kantenbedeckung und einer thermisch stabilen Zusammensetzung als größte Herausforderung identifiziert. Ein Abscheide-Regime für solch einen Prozeß wird vorgeschlagen, das jedoch außerhalb des in dieser Arbeit zugänglichen Bereiches liegt. Auf der anderen Seite wird eine TiN Atomlagen-Abscheidung erfolgreich modifiziert um 70% Kantenbedeckung in Gräben mit einem Aspektverhältnis von 40:1 zu erhalten. Mit diesem Prozeß werden Grabenkondensatoren hergestellt, die eine 60 s Temperung bei 1050 °C aushalten und gleichzeitig die Anforderungen für den Leckstrom und die Kapazität erfüllen. Zusammenfassend ermöglichen die in dieser Arbeit präsentierten Ergebnisse die Herstellung von niederohmigen Grabenfüllungen von DRAM-Kondensatoren so wie sie für eine 70 nm Technologie-Generation benötigt werden.

Chapter 1

Introduction

In 1960, the first **M**etal-**O**xide-**S**emiconductor **F**ield-**E**ffect **T**ransistor (MOSFET) laid the foundation for an unprecedented growth of semiconductor industry during the last 40 years. A MOSFET is a particularly cheap and simple device that can be scaled down in size while at the same time its performance improves. It took around 10 years before the one-transistor **D**ynamic **R**andom **A**ccess **M**emory (DRAM) cell and the first micro-processor were invented in 1968 and 1971, respectively. Storing and processing of information are the two main building blocks of electronics which led to a very rapid development of micro-processors and DRAMs as described in many text books [107, 39]. The anticipated future developments of both technologies are briefly stated in the following.

1.1 Dynamic Random Access Memory

Many devices of today's information technology like **P**ersonal **C**omputers (PCs) and **P**ersonal **D**igital **A**ssistants (PDAs) require a memory with fast read- and write-times. The only mature circuits fulfilling these requirements are DRAMs with one-transistor memory cells as shown in Fig. 1.1.

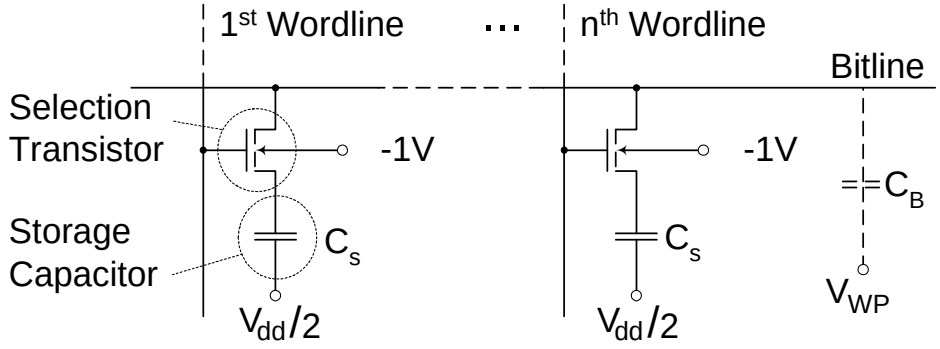

Fig. 1.1: Schematic of a DRAM cell array. At every cross point of bitline and wordline there is a DRAM cell consisting of a transistor and a storage capacitor with capacitance C_S. C_B indicates the parasitic bitline capacitance.

Information is stored in form of two distinct charge levels of the capacitor. To read out the information, the wordline is switched to high and thus opens the selection transistor so that charges from the capacitor can flow into the bitline. The voltage of this bitline is then evaluated by so-called sense-amplifiers that submit the digital information to the output buffer of the DRAM and write back the same information into the cell. A DRAM cell consisting of one selection transistor and a storage capacitor has many leakage paths by

which the stored charge diminishes in the course of time. For this reason, all information has to be refreshed after a certain period of time which is 64 ms in a standard product.

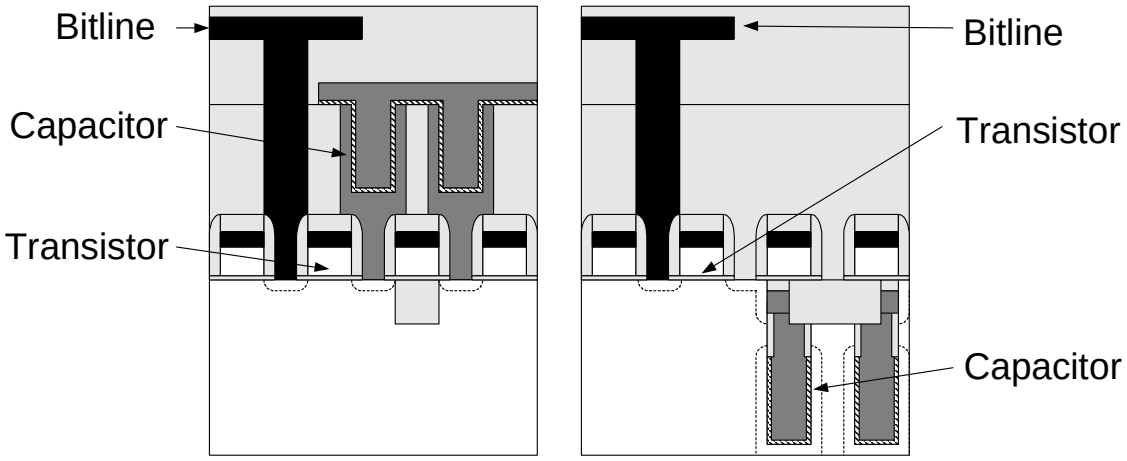

Fig. 1.2: Cross section of a stacked capacitor DRAM cell (left) and of a trench capacitor DRAM cell (right). White areas denote silicon, isolation is shown in light grey, capacitor electrodes in dark grey, metals in black and node dielectric by a hatched region.

The main driving force for DRAM-development has always been reduction in cost per bit, which means that the memory cells have to be as small as possible. The cell size is usually measured in terms of F^2, where the design rule F is defined as half pitch. The pitch is the smallest period of a periodical structure that can be printed with the lithography employed. With the commonly used sensing scheme, 8 F^2 is the smallest possible cell size which is also reached in state-of-the-art products. While the cell size reduces quadratically with decreasing design rule, it has been found out in production that a constant capacitance of approximately 25 fF is needed for every generation. At the same time, the leakage current of the node dielectric should not exceed 1 fA/cell at the operation voltage used which means that the dielectric thickness cannot be reduced below a certain value. To reach the capacitance for smaller design rules, an area increase by building threedimensional structures is indispensable. In principle, this can be realized by etching holes into the substrate (trench technology) or by building a stacked structure (stacked technology) as shown in Fig. 1.2. Both versions are used in production, but only the trench technology will be described in more detail here. The trench depth continuously increased over the past years in order to meet the capacitance target. For sub-100 nm generations the capacitance will be enhanced by **H**emispherical **S**ilicon-**G**rains (HSG), deep trenches with **A**spect **R**atios (ARs) higher than 60:1 and high-k materials for the node dielectric [34]. In parallel, the resistance of the inner electrode which is usually fabricated with poly-crystalline silicon (polysilicon) increases drastically [67]. According to the SIA roadmap, the polysilicon will have to be replaced by a metal from the 70 nm generation onwards. A more detailed explanation is given in the following.

In a DRAM the time available between applying a signal to the wordline and reading the information at the bitline is given by the joint electron device engineering council (JEDEC) agreements. As shown on the left hand side of Fig. 1.3 it is composed by the delay time of the word line τ_{WL}, the delay time of the transistor $\tau_{\mathrm{T}} \approx 2 * C_{\mathrm{S}} * (R_{\mathrm{C}} + R_{\mathrm{BS}})$, the delay time of the capacitor $\tau_{\mathrm{C}} \approx 2 * C_{\mathrm{S}} * R_{\mathrm{DT}}$ and the delay time of the bitline τ_{BL}. Here, C_{S} is the storage capacitance, R_{C} the channel resistance, R_{BS} the buried strap resistance and R_{DT} the deep trench (DT) resistance. The collar serves to switch off the vertical parasitic **F**ield-**E**ffect **T**ransistor (FET) between the substrate electrode (buried

plate) and the buried strap that connects capacitor and transistor. The exact shape of the collar has a strong influence on the DT-resistance so that no single design rule can be calculated from the time point at which polysilicon is no longer sufficient.

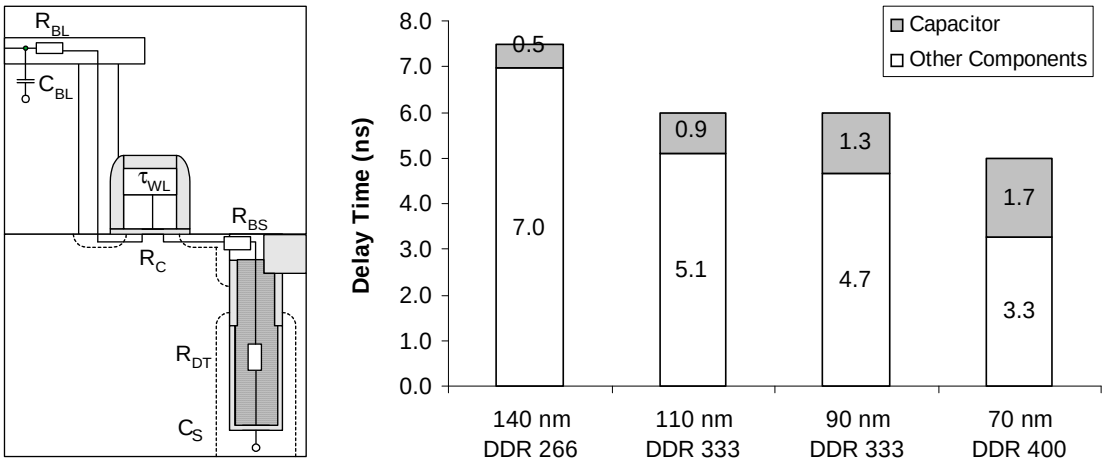

Fig. 1.3: Schematic and equivalent circuit of a DRAM deep trench cell (left). The right hand side shows the future trend of the cell delay time and of the capacitor delay time.

The right Panel of Fig. 1.3 shows the anticipated future development of the overall delay time and the calculated DT delay time for one specific cell layout with polysilicon as deep trench fill. While the DT delay time can be neglected for the currently produced technology (140 nm) it accounts for more than a third of the overall delay time for the 70 nm generation. The only way to reduce the delay time of the other components is to shorten the bitline. This, however, requires additional chip area and reduces the benefit of the new technology generation. It is therefore indispensable to replace the polysilicon by low resistance materials like metals. All metals considered here have a much lower resistivity than polysilicon. Hence, it might be sufficient to replace only parts of the fill with metal. The most promising approach is depositing a stack of a thin polysilicon layer and then a thin metal layer inside the DT. So, the well-known dielectric/polysilicon-interface can be employed further on and stress in the electrode will have a less deteriorating affect on the node dielectric. **Chapter 3** describes a test structure for the direct measurement of the DT-resistance. Deposition of metals in high-aspect ratio trenches is described in **Chapter 4** and deep trench structures are characterized in **Chapter 6**.

1.2 MOS Field-Effect Transistor

MOSFETs are massively employed as switching devices in many electronic circuits. It is remarkable that this device has not been modified significantly during the last 20 years while the performance increased mainly by scaling down device dimensions. Materials used in production even remained unchanged since the 70ies, after the aluminium-gate has been replaced by polysilicon.

The excellent properties of thermally grown silicon-dioxide (SiO_2) allowed for continuous shrinking of gate length and oxide thickness as shown in Fig. 1.4. There is, however, a physical limit to the reduction in oxide thickness. The absolute limit is 8 Å, which corresponds to 3 atomic monolayers that are required to obtain stoichiometric SiO_2. On the other hand, 12 Å are considered to be the minimum thickness for a production-worthy process [111]. Below this thickness, replacement of the well-known silicon-dioxide by a material with higher dielectric constant (high-k) will be necessary. This reduces gate leak-

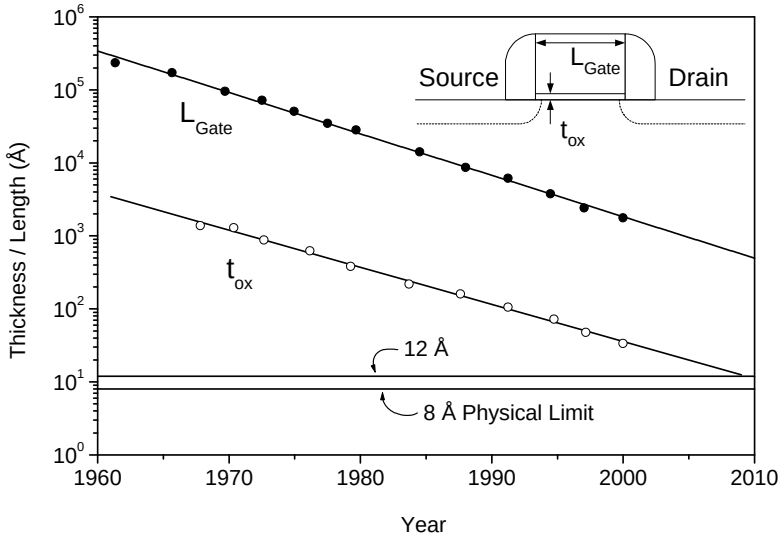

Fig. 1.4: Minimum gate lengths and oxide thicknesses in integrated circuits as they developed through the last 40 years. Also shown are the physical limit of 8 Å for the gate oxide thickness and the assumed practical limit of 12 Å.

age which in turn keeps the stand-by power at an acceptable level. There are a number of binary metal-oxides considered as high-k material for gate-dielectric application [112]. It might, however, be insufficient to replace the dielectric film since a significant amount of the equivalent dielectric thickness stems from the depletion regions in gate and substrate. While the substrate depletion region of 3-6 Å is difficult to avoid, gate depletion can be overcome by introducing a metal electrode [111]. Two approaches are possible with respect to the metal gate. On the one hand, a single midgap material like TiN can be used for n-MOS and p-MOS (Fig. 1.5). Due to the silicon-bandgap of 1.1 eV this leads to threshold voltages of around 0.5 V which are far too high for an expected voltage supply of 1 V necessary to turn on the devices [111]. On the other hand, two metals with different work function can be used as shown in Fig. 1.5 [64]. A selection of suitable metals is given in Section 1.3.

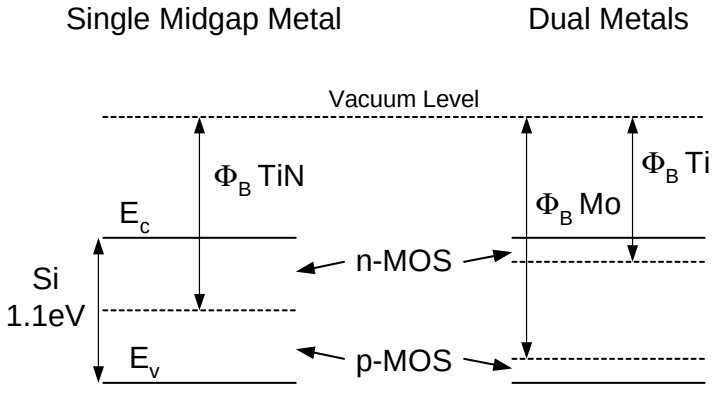

Fig. 1.5: Band diagram of a single midgap metal like TiN for gate electrode application (left). The right hand side shows the same schematic for a dual metal gate with Ti for n-MOS and Mo for p-MOS, respectively.

There are a number of ways to deposit the metal on the gate dielectric. While **P**hysical **V**apor **D**eposition (PVD) can damage thin dielectrics, **C**hemical **V**apor **D**eposition (CVD) has widely been used for this purpose. Recently, **A**tomic **L**ayer **D**eposition (ALD) has been introduced to semiconductor industry and exhibits potentially excellent step coverage and

a well defined deposition rate. The influence of this technique on the interface properties of TiN on tunnel-oxides is characterized in **Chapter 6**.

1.3 Advanced Electrode Materials

Today, polysilicon is used in DRAM-semiconductor technology as gate electrode and as inner electrode of the storage capacitor. Reasons for this are the reproducibly good interface characteristics to SiO_2, low mechanical stress, low defect density, high breakdown voltage as well as a work function that can be defined by the doping level. Drawbacks of polysilicon are the depletion region at the interface and, compared to most metals, a lower work function and higher resistivity. Highly-doped silicon can have a resistivity of 0.5 mΩcm, but due to very demanding deposition requirements and the need to use arsenic instead of phosphorus, a minimum value of 5 mΩcm is reached for the inner electrode of a deep trench DRAM electrode.

Metals considered for the described application have to withstand high frontend processing temperatures. Refractory metals are thermally stable up to very high temperatures, but are usually not chemically stable on SiO_2 or silicon. Their silicides and nitrides, on the other hand, exhibit that chemical stability and could be useful as metal electrodes. Table 1.1 summarizes the material properties of the most interesting metals.

Material	ΔE_c [eV]	ϕ [eV]	Resistivity [$\mu\Omega$cm]	Comment
Si	3.15 [109]	4.1±0.1[120]		
W	3.7 [109]	4.75 [104] 4.7±0.1 [120]	8 [55]	Midgap
Ti		4.4 [64, 106]	60 [25]	n-MOS
Ta		4.25 [104]	12 [110] 15 (bcc-Ta),160(β-Ta) [99]	
Mo		4.64 [104] 4.72 [64]	26 [3]	p-MOS
WN	3.48±0.2 [23]	5.00 [104]	200 [60]	
TiN	3.67±0.2 [23]	4.95 [104] 4.65 [89]	120 [54] 50 [54]	Midgap
TaN	3.40 [99]	5.41 [104]	1400 [65]	
MoN		5.33 [104]	400 [3]	
WSi_2			50 [87] 60 [23] 30 [72]	

Table 1.1: Material properties of the most promising metals considered for electrode material.

TiN is very promising as midgap gate-electrode [89] and is stable on SiO_2 up to 850 °C [23]. Beside thermal stability, the existence of a CVD or ALD process with good step coverage is crucial for the integration of such materials into DRAM capacitors. These processes exist for WN, TiN and WSi_2 all of which are investigated in this study.

This thesis is structured as follows: **Chapter 2** describes electrical characterization methods that are suitable to investigate interface properties of MOS-structures and a newly developed method to extract physical oxide thickness and flatband potential from tunnel-oxide MOS-capacitors. **Chapter 3** presents a new process flow that allows for

the fabrication of metal gate transistors on thin dielectrics. Test structures required for electrical analysis are described and were included in the test chip developed during this study. **Chapter 4** shows the exploration of technology development needed to fabricate the test structures for planar devices and DT capacitors which is explored in **Chapter 4**. Most metal gates were fabricated on RTP-oxides which are not as well characterized as the commonly used furnace-oxides. **Chapter 5** investigates special properties of RTP-oxides as a reference. In **Chapter 6** metal-gate MOS-structures as well as DT-capacitors with metal fills are finally characterized in detail. **Chapter 7** summarizes all results.

Chapter 2

Characterization Methods

This chapter describes the theoretical background of characterization methods which are suitable for analysis of metal-insulator-semiconductor (MIS) structures. Using simulations, it is shown how physical parameters of the electrode, the dielectric and the interface can be extracted from experimental data. Measurements were carried out on an 8 inch prober Cascade Summit 1200 with guarded chuck. A parameter analyzer HP4156B, an LCR-Meter HP4284, a switch HP5120 and a function generator HP8110A were used and controlled either by Metrics ICV or by HP Basic. Chuck temperature was set between 0 °C and 200 °C by a Temptronics controller. This setup allowed measurements in the fA-regime at the needles as well as at the chuck.

2.1 CV-Measurements on MIS-Structures

Capacitance-voltage (CV)-measurements determine the capacitance of an MIS-structure as a function of applied bias, V_G. On the one hand there are dynamic measurements at low (LF-CV) and high (HF-CV) frequencies and on the other hand static CV-measurements. Only the former are described in the following, because they yield the most interesting information.

2.1.1 Dynamic CV-Measurements

During dynamic CV-measurements an AC signal of 50 mV amplitude is applied to the gate and superimposed by a DC component which is swept through a range of several volts. Using an LCR-meter the impedance of the system is determined at each point. This can be done at low frequencies where the system is in equilibrium and at high frequencies. A classical description allows simulation of the LF-CV curves shown in Fig. 2.1a [74]. The time constant of majority carriers in accumulation is around 1 ps so that charge carriers can follow the typical frequencies between 1 kHz and 1 MHz. In inversion, however, minority carriers have generation and recombination times between 0.01 s and 0.1 s and cannot follow the signal any more, resulting in the LF-CV curve in Fig. 2.1a. Usually, LF-CV curves require measurement frequencies of a few Hz, but highly doped substrates or poor interfaces allow measurements in the kHz-regime. Generation time can be reduced by illumination, but this influences the voltage drop across the substrate and hence distorts the CV-curve. If the gate bias is swept faster than 100 mV/s, no equilibrium is reached in the depletion zone leading to deep depletion (DD) [93].

Classical simulations do not yield satisfactory results for modern dielectrics of a few nanometer thickness. For this reason, all simulations were carried out with *Medici*, a com-

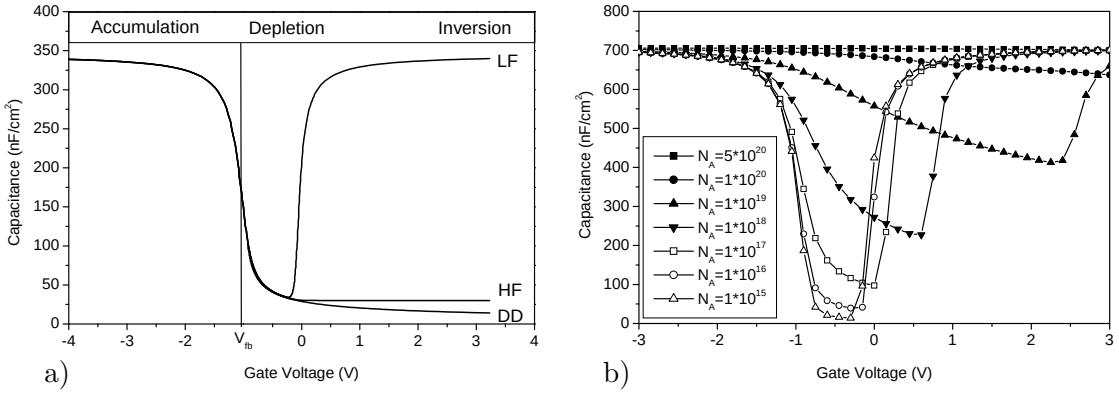

Fig. 2.1: a) Simulated CV-curves of an NMIS-structure with 10 nm Oxide for high (HF) and low frequency (LF) and for deep depletion (DD). Fig. b) shows HF-CV curves of a MIS-structure with 5 nm Oxide and metal gate for different substrate doping levels, N_{A}. Doping levels are given in cm^{-3}.

mercially available device simulator. Using quantities gained from simulation, a method is presented to extract physical parameters from measurements [96]. This extraction is very fast and does not require further simulation, making this method applicable for a large number of samples or even for the generation of wafer maps. The first quantity extracted from CV-curves is the flatband potential, V_{FB}, which depends on the work function difference, ϕ_{MS}, between the two electrodes and upon the different charges in the dielectric [93]:

$$V_{\mathrm{FB}} = \phi_{\mathrm{MS}} - \frac{Q_{\mathrm{f}}}{C_{\mathrm{diel}}} - \gamma\frac{Q_{\mathrm{m}}}{C_{\mathrm{diel}}} - \gamma\frac{Q_{\mathrm{ot}}}{C_{\mathrm{diel}}} - \frac{Q_{\mathrm{it}}(\phi_{\mathrm{s}})}{C_{\mathrm{diel}}} \tag{2.1}$$

Here, Q_{f} is the fixed charge near the silicon-dielectric interface, Q_{m} is the mobile charge, Q_{ot} the fixed dielectric charge and Q_{it} the surface potential-dependent interface charge. The factor γ accounts for the distribution of charge inside the dielectric. During analysis of new electrode materials, the work function difference, ϕ_{MS}, is of particular interest. Since the work function of the substrate is very well known, the work function of the gate can be extracted easily. This is necessary to predict the threshold voltage of a transistor and the gate leakage current.

In order to determine the work function, samples with different thicknesses of the gate dielectric can be prepared. A plot of the flatband potential versus dielectric thickness and extrapolation to 0 nm thickness yields the value of the work function difference. Doping variations in the substrate have to be taken into account [93]. Comparing the flatband potential of a homogeneously doped substrate with an epi-substrate which has a different doping concentration at the surface results in:

$$V_{\mathrm{FB}}(\text{epi wafer}) = V_{\mathrm{FB}}(\text{uniform wafer}) \pm \frac{kT}{2q}\ln\left(\frac{N_{\mathrm{sub}}}{N_{\mathrm{epi}}}\right) \tag{2.2}$$

The flatband potential is determined by the flatband capacitance. For low substrate doping levels, quantum confinement at flatband conditions is negligible so that the latter remains unchanged for thin oxides [75]. The ratio of flatband, C_{FB}, and dielectric capacitance, C_{diel}, is described by the following expression [93]:

$$\frac{C_{\mathrm{FB}}}{C_{\mathrm{diel}}} = \frac{1}{1 + \frac{\varepsilon_{\mathrm{d}}}{t_{\mathrm{diel}}}\,347\,\sqrt{\frac{T/300}{n+p}}} \tag{2.3}$$

During the analysis described in this section, flatband and dielectric capacitances are determined iteratively. The dielectric capacitance is estimated from the measured data and used to determine the flatband voltage. Together with the data and simulated quantities this value is used to recalculate the dielectric capacitance. After only a few iterations both quantities are determined accurately. The capacitance of the MIS-structure varies very strongly near the flatband voltage and thus supports an accurate measurement. Fig. 2.1b shows CV-curves as a function of substrate doping. The variation in capacitance near the flatband potential is much smaller for high substrate doping levels so that these samples are less suitable for extraction of the flatband potential.

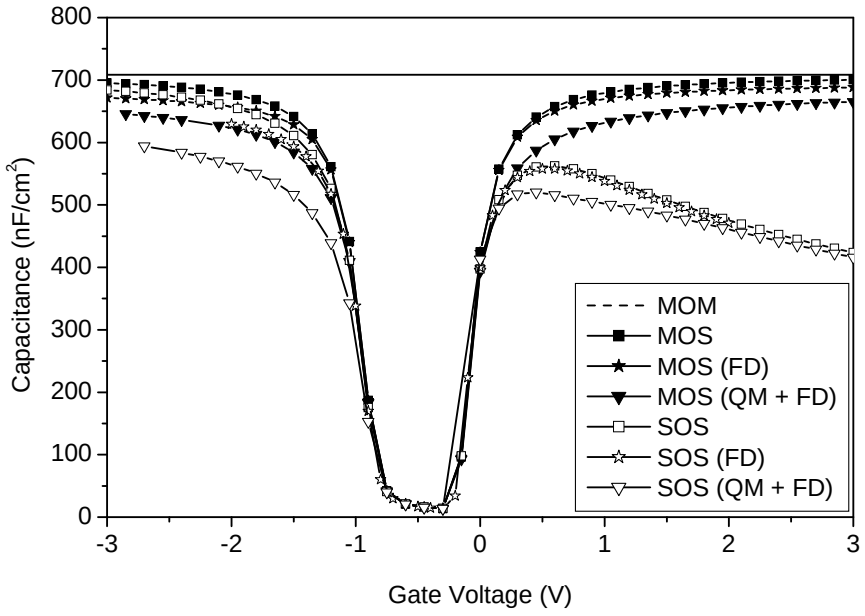

Fig. 2.2: Simulated CV-curves of MIS-structures with a 5 nm oxide. Shown is the influence of quantum mechanical corrections (QM), Fermi-Dirac statistics (FD) and polysilicon depletion in the gate electrode.

The following presents a method to extract the dielectric thickness of thin dielectrics. Fig. 2.2 shows simulated CV-curves of a 5 nm oxide for metal-oxide-metal (MOM), metal-oxide-silicon (MOS) and silicon-oxide-silicon (SOS) structures with and without the influence of quantum mechanical effect and Fermi-Dirac statistics. Van Dort's bandgap widening approach was used to approximate quantum mechanical effects [26]. An AC analysis yielded the MOS capacitance as function of applied bias. Even with relative thick oxides of 5 nm there are significant differences so that quantum mechanical effects and Fermi-Dirac statistics always have to be taken into account. Furthermore, silicon gate electrodes show a substantial reduction in capacitance for accumulation so that MOS and SOS structures have to be separated carefully. Simulation and fitting of measured CV-curves is rather tedious. Therefore, correction terms are determined in the following, which allow the direct extraction of physical oxide thickness from CV-curves. In literature there are already some contributions to the extraction of physical oxide thicknesses [61, 82, 83, 84, 9, 22, 81]. Given are values for the increase in oxide thickness as function of physical oxide thickness [51], and formulas to calculate the position of the charge centroid [122]. In this work, the effective oxide thickness increase is calculated for a fixed offset to the flatband potential for NMOS and PMOS structures. Metal gates and polysilicon gates with different doping levels have been examined. The fixed offset to the flatband potential allows for an automatic extraction of physical oxide thickness. The presented simulations

are able to describe all different systems and are in good agreement with literature values.

In most cases the dielectric thickness of a MOS-structure is extracted from the capacitance in accumulation. The difference between an exactly calculated capacitance in accumulation and the dielectric capacitance can be described by an additional capacitance in series to the dielectric one. This series capacitance accounts for incomplete accumulation, Fermi-Dirac statistics, quantum mechanical effects and the distance between charge centroid and interface between polysilicon electrode and substrate. Fig. 2.4 shows the oxide equivalent thickness, Δt_{ox}, of this capacitance as a function of gate bias, V_{G}, for different structures. Fig. 2.4a and b demonstrate simulations of SOS-structures as they are implemented in modern circuits. In Fig. 2.4a NMOS-capacitors were calculated while for Fig. 2.4b PMOS-structures were considered. Here, N and P always indicate the carrier type of the inversion layer so that NSOS is a structure with n-type polysilicon on oxide on p-type substrate. Panels c and d of the same figure are equivalent simulations of NMOS and PMOS capacitors with metal gate as characterized in this work. Panels e and f show simulations of structures as they are commonly used for storage capacitors in DRAMs. Usually both electrodes are highly doped with the same kind of dopand.

An automatic analysis of CV-curves requires a reference point, since the curve can be shifted by a change in work function. The flatband potential, V_{FB}, given in the figures is most suitable for this purpose. From the data in Fig. 2.4 the increase, Δt_{ox}, of the equivalent oxide thickness at the voltages $V = V_{\mathrm{FB}} - 1,5$ V for p-substrate and at $V = V_{\mathrm{FB}} + 1,5$ V for n-substrate have been summarized in Fig. 2.5a, c and e. A transformation of physical thickness into measured thickness as in Panels b, d and f allows for a direct analysis of measured data. Based on this method, a Matlab routine has been written and was used to handle the data presented in this thesis. Comparison with the IBM-Model of Lo et al. shows very good agreement for polysilicon gates [62]. However, the IBM-model is not suitable for metal electrodes. Wafer maps of a typical polysilicon sample with a 27 Å gate oxide are shown in Fig. 2.3 [96].

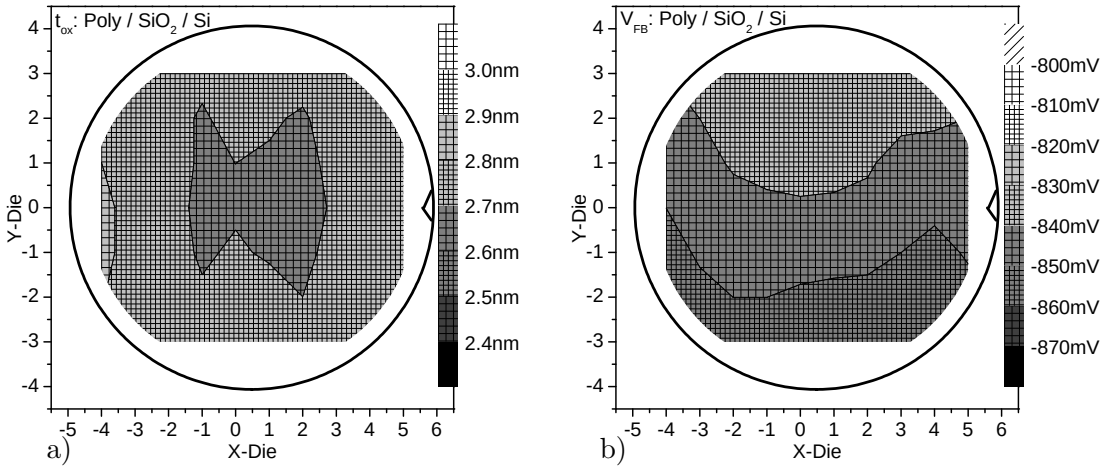

Fig. 2.3: Wafer maps of physical oxide thickness (a) and flatband potential (b) of MOS capacitors with polysilicon gate and 27 Å gate oxide. 58 point were measured on the wafer.

Extracted flatband values vary only 10 mV over the whole wafer and the thickness is determined with an accuracy of 0.1 nm. Thickness uniformity is mainly determined by the oxidation process while doping non-uniformities in the polysilicon result in flatband potential variations.

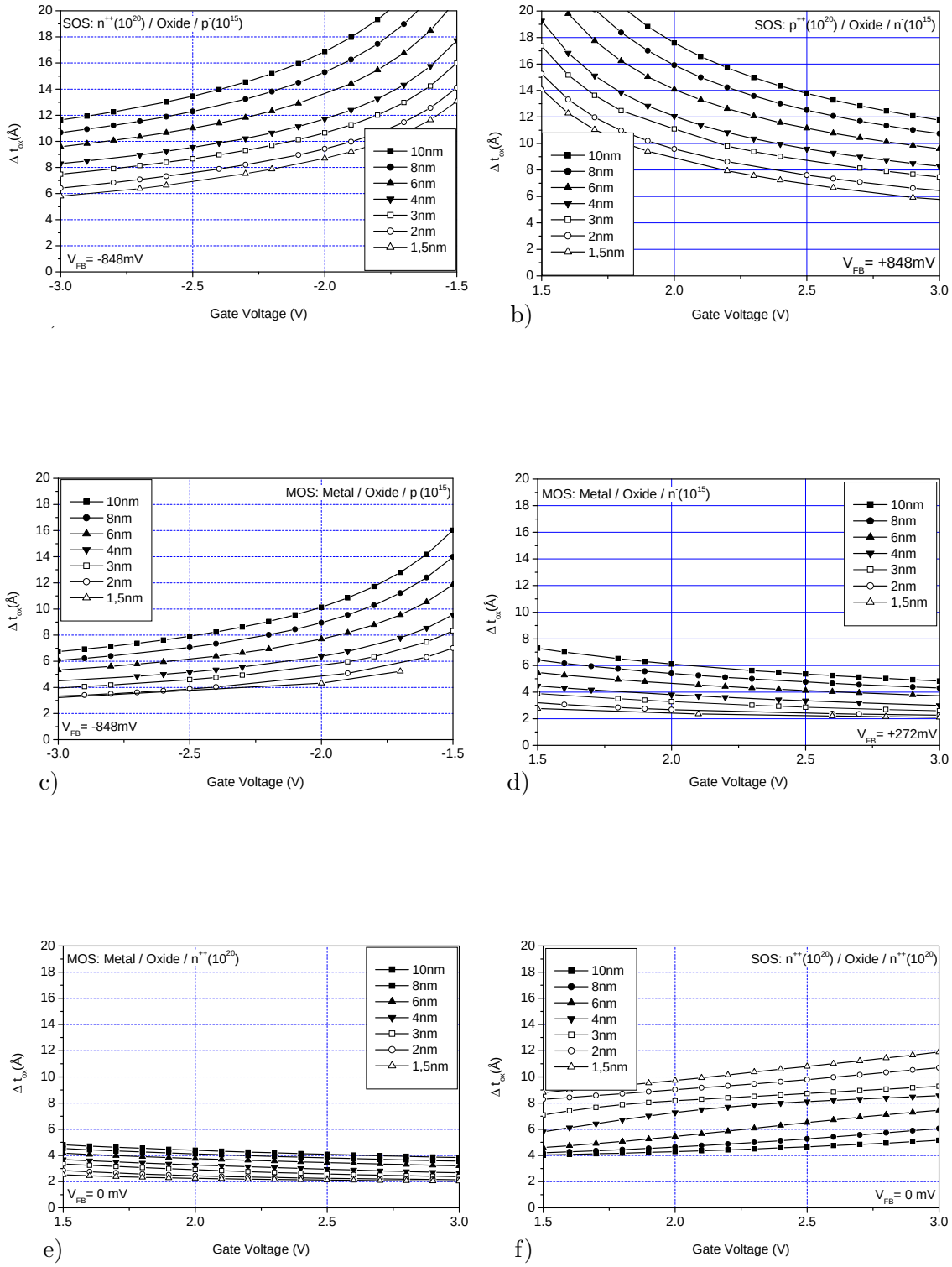

Fig. 2.4: Oxide equivalent thicknesses, Δt_{ox}, of correction terms which have to be added to the physical oxide thickness in order to coincide with the simulations. Shown are data for physical oxide thicknesses between 1.5 nm and 10 nm. Corrections account for incomplete accumulation, quantum mechanical corrections, Fermi-Dirac statistics and the distance between the charge centroid and the interface of polysilicon gate and substrate.

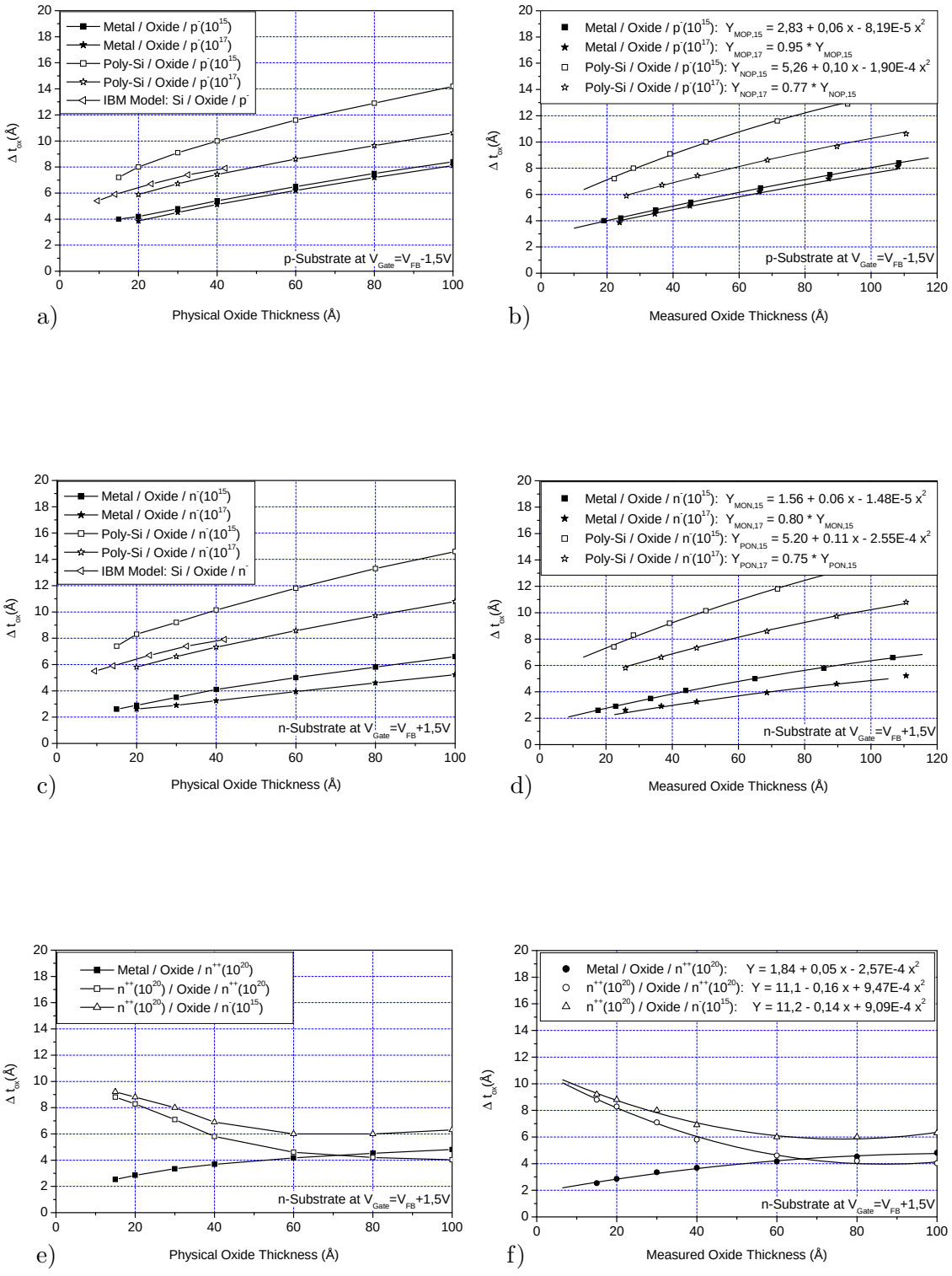

Fig. 2.5: Oxide equivalent thickness, Δt_{ox}, of correction terms as a function of physical oxide thickness (Panels a, c and e) or the measured oxide thickness (Panels b, d and f). Values have been extracted from data shown in Fig. 2.4 at a gate voltage of $V = V_{\mathrm{FB}} - 1,5$ V for p-substrate and $V = V_{\mathrm{FB}} + 1,5$ V for n-substrate. Shown are fit-parabolas to analyze oxide equivalent thicknesses between 1.5 nm and 10 nm. SOS-structures agree very well with the IBM-model.

Simulations did not show any differences for different dielectric constants as long as the equivalent thickness did not change. Hence, to a first approximation, the fit-parabolas can be used for any kind of dielectric with an equivalent oxide thickness between 1.5 nm and 10 nm. For very thin oxides, however, the leakage current through the dielectric can influence the capacitance measurements. High frequencies allow a higher leakage current and hence the measurement of thinner oxides. Fig. 2.6 shows the minimum measurable oxide thickness as a function of frequency which can be determined to better than 1% accuracy. Statements correlate to flatband potential $\pm 1,5$ V. This limitation is only valid for procedures which neglect influence of tunneling current. Measurements above 1 MHz require a detailed analysis of parasitic inductances and often a special measurement setup, so that 18 Ångström is the minimum measurable oxide thickness for simple setups.

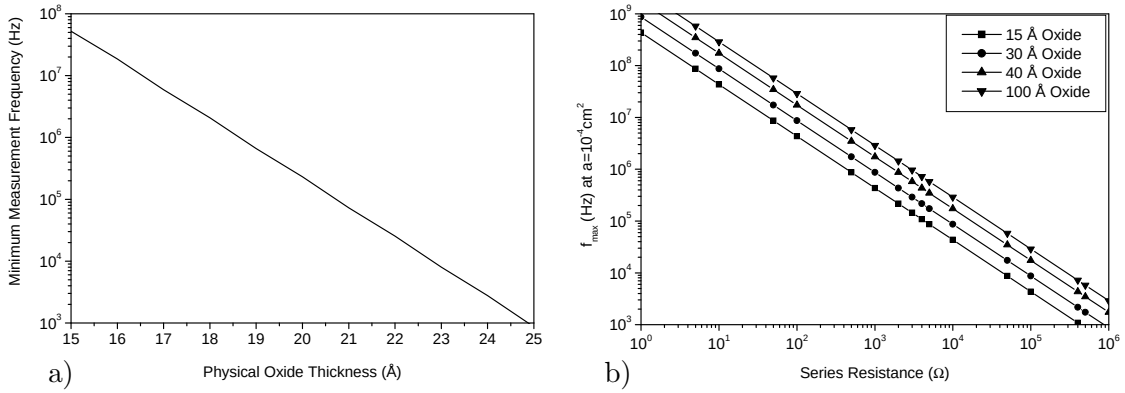

Fig. 2.6: Minimum oxide thickness of an SOS-capacitor, which can be determined with an accuracy better than 1% as a function of measurement frequency (a). Panel b) shows the maximum measurement frequency as function of series resistance with oxide thickness as parameter.

However, an increasing frequency also increases the influence of the series resistance. Fig. 2.6 b shows the maximum frequency as a function of series resistance, at which the dissipation factor $D = \omega R_s C$ is smaller than 0.1. A device area of 10^{-4} cm^2 has been assumed. Altogether, a small area with a small series resistance should be measured in order to be able to choose very high frequencies. Structures with a minimum series resistance of 1 Ω can be built, so that according to Fig. 2.6b measurement frequencies of more than 100 MHz could be used. In that way, oxide thicknesses down to 15 Ångström could be determined. For metal electrodes with higher work functions this value could even be reduced.

2.1.2 Extraction of Physical Parameters from CV-Curves

In addition to the equivalent oxide thickness of the dielectric, CV-curves also yield information on the interface charge density and work function difference between the two electrodes. Usually, quasi-static or low-frequency measurements are compared to high-frequency results or to simulations. Traps that follow the low-frequency signal but not the high-frequency one are calculated with the following formula [74]:

$$D_{\mathrm{it}} = \frac{\Delta C}{q} \cdot \left[1 - \frac{C_{\mathrm{HF}} + \Delta C}{C_{\mathrm{ox}}}\right]^{-1} \cdot \left[1 - \frac{C_{\mathrm{HF}}}{C_{\mathrm{ox}}}\right]^{-1} \tag{2.4}$$

Here, C_{LF} and C_{HF} are the low- and high-frequency capacitance respectively and $\Delta C = (C_{\mathrm{LF}} - C_{\mathrm{HF}})$. Usually, high-frequency measurements are carried out at 1 MHz while the low-frequency ones are done at 10 kHz [66, 1, 79].

A further means to determine work function difference and trap density evolves from equation 2.1. The flatband potential changes linearly with dielectric thickness. If samples of different oxide thickness are processed and the flatband potential is plotted as a function of oxide thickness, the slope gives the trap density and the extrapolation to 0 nm thickness yields the work function difference [120]. An example of such a measurement is presented in Fig. 2.7. The flatband potential has been extracted with the algorithm of Section 2.1.1. It is apparent that the substrate doping has to be known very well in order to obtain reliable data.

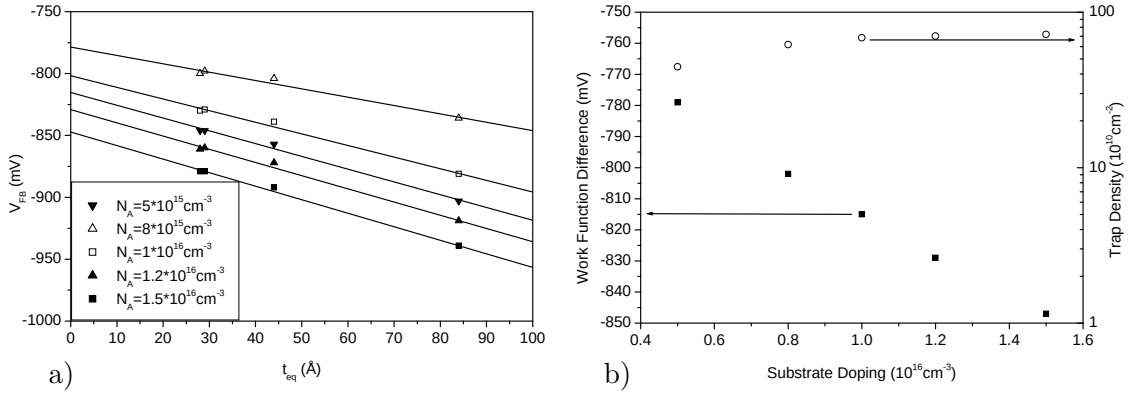

Fig. 2.7: Panel a) shows the measured flatband potential as a function of physical oxide thickness for different assumed substrate doping levels. Depending on this doping level slightly different values for trap density and work function difference are extracted (b).

The substrate doping can be determined from the oxide capacitance, C_{ox}, and the inversion capacitance, C_{inv} [93]:

$$N_{\mathrm{A}} = \frac{4\phi_{\mathrm{F}}}{q\varepsilon_{\mathrm{s}}\varepsilon_0 A^2} \frac{C_{\mathrm{inv}}^2}{(1 - C_{\mathrm{inv}}/C_{\mathrm{ox}})^2} \tag{2.5}$$

In this case a substrate doping level of $1.0 \pm 0.2 \cdot 10^{16}$ cm^{-3} has been determined. The extracted value depends only weakly on the oxide capacitance so that quantum mechanical effects can be neglected. Hence, a charge density of $7 \cdot 10^{11}$ cm^{-2} and a work function difference of -815 mV are found. According to equation 2.1 the gained charge density is the sum of the fixed oxide charge at the interface and the voltage-dependent trap density D_{it}. The latter can be extracted separately by charge pumping measurements described later in this chapter. A comparison gives the fixed charges in the oxide. Mobile charges can be identified from the hysteresis of the CV-curve, but are negligible in modern silicon-oxide and were below the detection limit in this work.

2.1.3 Parasitic Components during CV-Measurements

As already mentioned in Section 2.1.1, the measurement frequency determines whether a MIS-structure in inversion is in equilibrium or not. Fig. 2.8a shows CV-measurements of planar capacitors with WSi_x electrodes and NO-dielectric as a function of frequency. At 200 Hz typical LF-CV curves are measured, while the curve for higher frequencies look more like HF-CV curves. All measurements coincide in accumulation and depletion. The only exception is the measurement at 1 MHz where the influence of the parasitic inductance becomes dominant, since it increases quadratically with frequency.

During every measurement there is a parallel parasitic capacitance due to the cable connection and a series parasitic capacitance for example between silicon wafer and chuck.

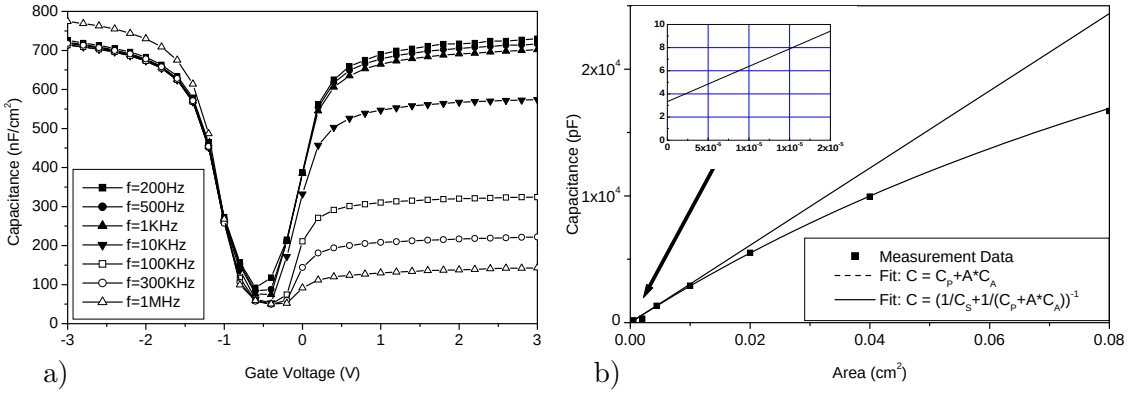

Fig. 2.8: CV-curves of an NMIS-structure for different frequencies (a). Panel b) shows the measured capacitance as function of capacitor area. Assuming one series and one parallel parasitic capacitance the data can be fitted very well.

Since these quantities are independent of the measured structure, they can be extracted by measuring a series of capacitors with different areas. Fig. 2.8b shows the measured capacitance as a function of area. The data can be fitted very well by assuming both parasitic capacitances. In this case, the extracted values for parallel and series capacitance were 3.2 pF and 55 nF, respectively. The parallel capacitance is relatively high since a switch has been used between LCR-meter and prober. To decrease the influence, the switch can be left out and the length of the cable, after merging the 4-point contacts should be as short as possible. In addition, the parasitic capacitance can be measured directly with non-connected needles using the *zero cancel* option of the LCR-meter. This value will be subtracted from all subsequent measurements. Most practical, however, is the measurement of capacitances with intermediate areas (10^{-3} cm^2), since then usually all parasitic effects are negligible.

2.2 IV-Measurements on MIS-Structures

Leakage current through an MIS-structure is an essential quantity during the development of transistors and capacitors. From IV-curves it is possible to determine the transport mechanism of charge carriers through the dielectric which yields information on interface traps. Ideally, the dielectric is free of defects so that the leakage current can be described by the well known Fowler-Nordheim formula. In this case it is possible to extract intrinsic properties like parabolic electron mass in the silicon-dioxide conduction band edge or the barrier height between electrode and dielectric. In the following, the two most important transport mechanisms through dielectrics and the most important aspects during the measurement are described.

2.2.1 Carrier Transport through Dielectrics

Dielectrics used in the semiconductor industry almost always have an amorphous structure. Despite the lack of a long-range order, a band diagram is typically used to describe physical effects. The density of states is not zero inside the band gap, but so called mobility edges exist. All states inside these edges are localized, all other states are non-localized and represent conduction and valence band. The density of localized states determines which conduction mechanism dominates.

In case of a negligible density of localized states the carrier transport can be described

by tunneling through a potential barrier. The following formula has been derived initially by Fowler and Nordheim, so that the mechanism is usually referred to as Fowler-Nordheim tunneling [31]:

$$J_{\mathrm{FN}} = A \cdot E_{\mathrm{ox}}^2 \cdot \exp\left(-\frac{B}{E_{\mathrm{ox}}}\right) \tag{2.6}$$

$$A = \frac{q^3}{16 \cdot \pi^2 \cdot \hbar \cdot \Phi_{\mathrm{b}}} \tag{2.7}$$

$$B = \frac{4 \cdot \sqrt{2 \cdot m_{\mathrm{ox}}}}{3 \cdot \hbar \cdot q} \cdot \Phi_{\mathrm{b}}^{3/2} \tag{2.8}$$

Here, E_{ox} is the electric field inside the oxide, Φ_{b} the barrier height between electrode and dielectric and m_{ox} the effective parabolic electron mass in the conduction band edge of the dielectric. This simple formula is valid only for triangular potential barriers, i.e. only for high gate voltages. A more sophisticated description of the leakage current is presented in Chapter 5 which is valid also for small gate voltages.

So far we assumed that the density of states inside the band gap of the dielectric can be ignored. If this is not the case, charge carriers can pass through the dielectric from one trap to another which is usually referred to as hopping-conduction. According to literature, deposited silicon-nitride has a defect concentration of 10^{19} cm^{-3} and shows such a conduction for small fields and low temperatures [85]. For a very small field, an ohmic conduction with a thermal activation energy of $q\Phi_{\mathrm{t}}$ has been observed, where Φ_{t} describes the depth of the defect compared to the conduction band edge. The potential near such a defect can be described by a Coulomb-potential which can be reduced by an electric field [7]. The measured current is then given by [106]:

$$J \propto E \cdot \exp\left(-\frac{q(\Phi_{\mathrm{t}} - \sqrt{qE/\pi\varepsilon_{\mathrm{d}}})}{kT}\right) \tag{2.9}$$

Here, E is the electric field and ε_{d} the dielectric constant of the insulator. Poole and Frenkel were the first to derive this expression giving the name to the so-called Poole-Frenkel emission [78, 32].

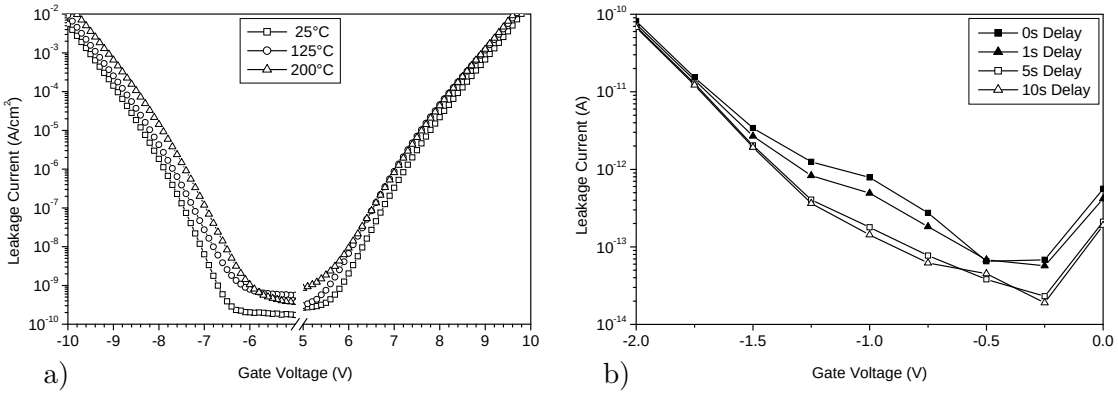

Fig. 2.9: Leakage current of a NMOS structure with 8.5 nm oxide for three different temperatures (a). Panel b) shows IV-curves of an NO-dielectric for different measurement parameters. All curves have been measured with long integration time and delay times between 0 s and 10 s.

It should be noted that the tunneling current depends on intrinsic properties of the dielectric only and not on the barrier height between electrode and insulator. Therefore, the mechanism is referred to as volume-limited transport. Both conduction mechanisms

described here always run in parallel, but depending on material, temperature and electric field usually one of them predominates. Measuring the temperature dependence of the leakage current reveals which transport mechanism can be neglected. Fig. 2.9a shows IV-curves of an MIS-structures with 8.5 nm oxide for three different temperatures. The small influence of temperature indicates Fowler-Nordheim tunneling, which has only a second-order dependence due to change in Fermi-level and in thermal energy kT of the electrons.

2.2.2 Measuring IV-Characteristics

During the measurement of an IV-characteristic, the gate voltage is swept step-wise and the current measured at each point. Starting at a fixed point in time after applying the voltage, the so-called *delay time*, the current is integrated over a period of up to several seconds. For too short delay times, transient currents can strongly affect the results. Fig. 2.9b shows IV-curves of a standard NO-dielectric for different delay and integration times. In the range of the operation voltage of -1 V values differ almost one order of magnitude. Therefore, NO-data presented in this work have been measured with a long integration time and a 5 s delay. Capacitors with an area of 10^{-3} cm^2 are ideal, since they facilitate a good resolution of the current density at small applied bias.

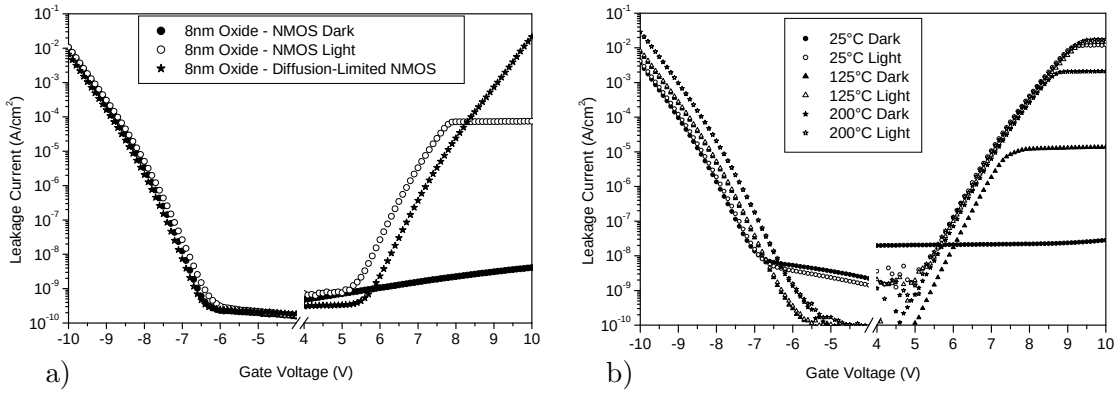

Fig. 2.10: Leakage current of an integrated MIS-structure with 8.5 nm oxide measured at a STI-limited capacitor with and without illumination and at a diffusion-limited capacitor (a). Panel b) shows the leakage current of simple planar capacitors with 8.5 nm oxide at different temperatures with and without illumination.

The type of the test structure can have a large impact on the measurement results. Fig. 2.10a shows IV-curves of the same dielectric measured at shallow trench isolation (STI)-limited NMOS-capacitors with or without illumination and at diffusion-limited capacitors. In accumulation, all values are identical so that very simple planar capacitors can be used to study the leakage current in accumulation. In inversion, however, only diffusion-limited capacitors give reliable results. STI-limited capacitors can only generate very few minority carriers for tunneling so that the current stays in the nA-regime also for high voltages. Illumination generates sufficient minority carriers but has also an influence on the voltage drop across the substrate so that the IV-curve is shifted by approximately 0.5 V. This shift can be seen also by narrowing of measured CV-curves. Fig. 2.10b shows IV-curves of simple planar NMOS-capacitors at different temperatures with and without illumination. The temperature dependence can be studied in accumulation and gives results similar to fully integrated structures. In inversion, the shift of the IV-curve due to the change in the voltage drop across the substrate can be seen for different temperatures as well as for cases with and without light. As a summary, only diffusion-limited capacitors should be

used to study leakage current of an MIS-structure in inversion. In accumulation, simple planar capacitors give the same values, while STI-limited structures are preferred to study the intrinsic reliability of the dielectric.

2.3 Charge Pumping

Back in 1969 Brugler and Jespers were able to show that a net current flows between source/drain and substrate of a MOS-transistor, if a pulsed voltage signal with certain properties is applied to the gate [11]. This phenomenon is usually referred to as charge pumping (CP) and Elliot had been the first one to apply this method to systematically explore interface states between silicon and silicon-dioxide in MOS transistors [28]. Fig. 2.11 shows the schematic of the setup, more details are found in standard text books [93, 113].

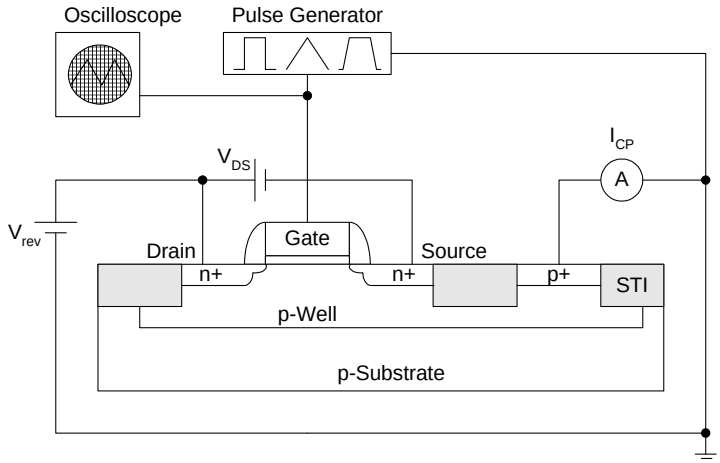

Fig. 2.11: Schematic setup for charge pumping measurements. Shown are a cross section of the device under test, a pulse generator to drive the gate and an ampere meter to measure the pump current.

The method can be applied to small structures where source and drain are shortened and set to a small reverse bias. A trapeziform AC voltage is applied to the gate as shown in Fig. 2.12. The amplitude at the gate has to be high enough to bring the substrate subsequently in inversion and accumulation. Depending on substrate doping and oxide thickness this is at least 3-5 V.

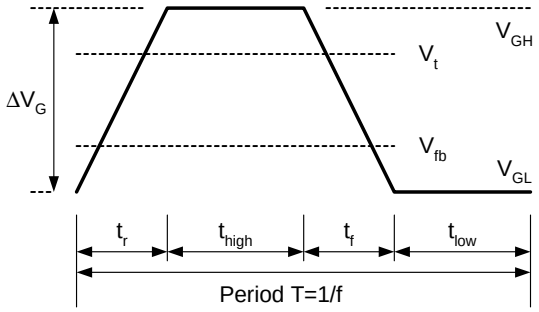

Fig. 2.12: Schematic of the trapeziform measurement signal which is applied to the gate. Raise and trailing times are 20% of the cycle time, i.e. $\alpha = 0.2$ and $\beta = 0.2$.

For very thin oxides the leakage current overlies the pump current, which makes a simple analysis of the data impossible. Pump cycles were investigated in order to evaluate other methods of data handling. The signal analysis presented in the following was

conducted by A. Avellán (TUHH) with an IV-transform circuit of reference [5]. CP-measurements were performed on p-channel MOS transistors with 10 μm channel length and 1000 μm channel width. Fig. 2.13 shows the measured source/drain, gate and bulk currents during one pump cycle with 33 kHz at a sample rate of 2,5 GHz [6]. A trapeziform AC voltage of 3 V amplitude was applied to the gate. Source and drain were shortened and put to 0 V. During a second measurement, the pump signal was applied to source/drain and substrate and the gate current was measured.

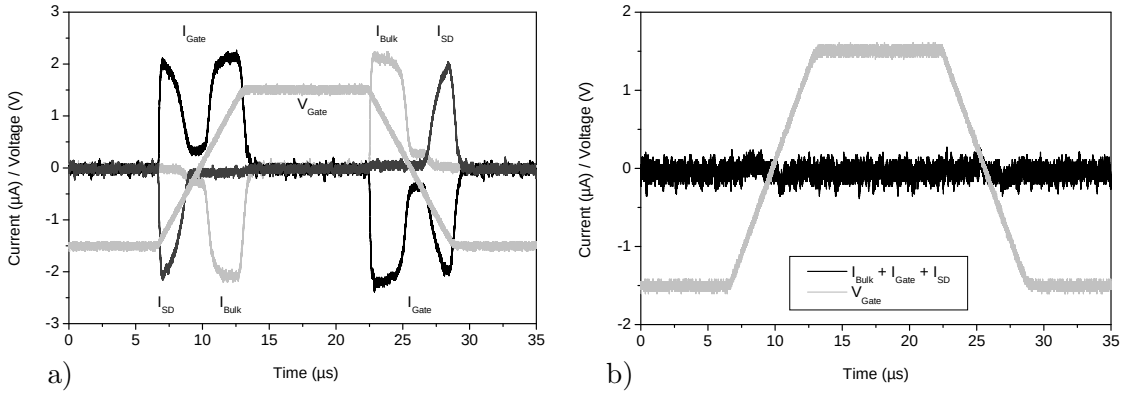

Fig. 2.13: Measured source/drain, gate and substrate currents and gate voltage during one pump cycle (a). The sum of the three current components is zero as shown in Panel b. Graphs were taken from reference [6].

After 0 μs the substrate is depleted completely. In inversion, channel and interface traps at the interface are filled with minority carriers. During the transition to accumulation, carriers in the channel and in shallow traps start to flow into the source/drain region. This process is completed after around 9 μs. A little later the deep traps and the space charge region are filled by majority carriers from the substrate. The recombination time of majority carriers with trap charges is very small in accumulation, so that all traps are filled with holes when the threshold voltage is reached at around 11 μs. The capture time, τ_{c}, of the holes is given by

$$\tau_{\mathrm{c}} = \frac{1}{\sigma_p v_{\mathrm{th}} p_{\mathrm{s}}} \tag{2.10}$$

where σ_p is the capture cross section, v_{th} the drift velocity of charge carriers and p_s the hole concentration at the interface [93]. At flatband potential $p_s \approx 10^{17}$ cm^{-3} and with $v_{\mathrm{th}} = 10^7$ cm/s and $\sigma_p = 10^{16}$ cm^2 the capture time is 10 ns and decreases rapidly with further reduction of gate voltage. The large current flowing between 11 and 14 μs does not contribute to the pump current and only charges the space charge region in the substrate. The gate can supply charge carriers very fast so that there are always mirror charges to those in the substrate. The sum of all currents is zero all the time as shown in Fig. 2.13b. The transition to inversion is analogous with the only exception that shallow traps are filled from the substrate and deep traps are filled from the source/drain region. In each pump cycle the current flows into the substrate during the rising edge and into the source/drain during the trailing edge. However, such a current is not detectable during time-resolved measurements, which leads to the conclusion that the pump current always is many orders of magnitude smaller than the displacement current. Therefore, time-resolved measurements are not suitable for investigating charge pumping phenomena.

To perform CP-measurements on tunnel oxides, two alternatives were proposed recently [69]. In the first approach, the leakage current component in one pump cycle is

estimated from the static tunnel current. However, this method requires a quasi-static distribution of electrons in the substrate and an additional measurement of the static gate leakage. During the second approach, a CP-measurement is done at a very low frequency, and the result is subtracted from all subsequent measurements. At low frequencies the contribution of the pump current is negligible so that the current is dominated by gate leakage. Despite the fact that measurements below a few kHz are sometimes difficult, this procedure has shown to be the most practical one.

In the description above, the transition from shallow to deep traps is determined by the rise and fall times of the gate signal. The emission time of traps, $\tau_{\mathrm{em,e}}$, can be expressed in the following way [93]:

$$\tau_{\mathrm{em,e}} = \frac{\exp((E_{\mathrm{c}} - E_1)/kT)}{\sigma_n v_{\mathrm{th}} N_{\mathrm{c}}} \tag{2.11}$$

Here, E_{c} and E_1 are energy levels of the conduction band and of the trap under consideration, σ_n is the capture cross section and v_{th} the drift velocity of the electron and N_{c} the density of states in the conduction band. Shallow traps have only a limited time period, $\tau_{\mathrm{em,e}}$, between reaching flatband potential and threshold voltage to emit electrons to the source/drain region. Thereafter they recombine with holes from the substrate. Hence, the rise time defines the energy level, E_1, which separates shallow from deep traps. Only deep traps contribute to the pump current, which is measured at the substrate. During the transition to inversion the fall time of the gate signal defines an emission time for holes:

$$\tau_{\mathrm{em,h}} = \frac{\exp((E_2 - E_{\mathrm{v}})/kT)}{\sigma_p v_{\mathrm{th}} N_{\mathrm{v}}} \tag{2.12}$$

Thus, during a pump cycle, only traps within an energy band $\Delta E = E_1 - E_2$ are measured. In the following, different methods are described to use this relation to interpret measured data.

2.3.1 Classical Charge Pumping

One way to analyze the data gained is to determine the energy band, ΔE, from measurement parameters and calculate the mean trap density via the CP-current [93]:

$$I_{\mathrm{CP}} = q A_{\mathrm{G}} f D_{\mathrm{it}} \Delta E \tag{2.13}$$

Groeseneken et al. developed a method in which frequency as well as rise and fall times of the gate signal are varied simultaneously and the change of pumped charge per cycle is determined as a function of frequency [33]. In this way the trap density per energy level can be determined. Using a trapeziform gate signal such as shown in Fig. 2.12 the emission times are given by the following expression:

$$t_{\mathrm{em,e}} = \frac{|V_{\mathrm{FB}} - V_{\mathrm{T}}|}{|\Delta V_{\mathrm{G}}|} \cdot t_{\mathrm{f}} \tag{2.14}$$

$$t_{\mathrm{em,h}} = \frac{|V_{\mathrm{FB}} - V_{\mathrm{T}}|}{|\Delta V_{\mathrm{G}}|} \cdot t_{\mathrm{r}} \tag{2.15}$$

The pumped charge per cycle, Q_{CP}, is approximated by the following expression [33]:

$$Q_{\mathrm{CP}} = 2q\overline{D_{\mathrm{it}}}A_{\mathrm{G}}kT\left[\ln(v_{\mathrm{th}}n_{\mathrm{i}}\sqrt{\sigma_{\mathrm{n}}\sigma_{\mathrm{p}}}) + \ln\left(\frac{|V_{\mathrm{FB}} - V_{\mathrm{T}}|}{|V_{\mathrm{G}}|}\sqrt{t_{\mathrm{f}} \cdot t_{\mathrm{r}}}\right)\right] \tag{2.16}$$

This formula becomes very simple when rise and fall times are proportional to frequency, i.e. if $t_{\mathrm{r}} = \alpha/f$ and $t_{\mathrm{f}} = \beta/f$. The pump current per cycle as a function of the logarithm of frequency is a straight line with the slope:

$$\frac{dQ_{\mathrm{CP}}}{d\log f} = \frac{2qkT\overline{D_{\mathrm{it}}}}{\log e} \cdot A_{\mathrm{G}} \tag{2.17}$$

With the knowledge of temperature and area, interface trap density can be calculated readily. From the frequency, f_0, at which the pump current vanishes, the capture cross section can be calculated:

$$\sqrt{\sigma_{\mathrm{p}}\sigma_{\mathrm{n}}} = \frac{1}{v_{\mathrm{th}} \cdot n_{\mathrm{i}}} \cdot \frac{|\Delta V_{\mathrm{G}}|}{|V_{\mathrm{FB}} - V_{\mathrm{T}}|} \cdot \frac{f_0}{\sqrt{\alpha\beta}} \tag{2.18}$$

Using this method, the temperature dependence of the capture cross section can easily be determined. There are a number of ways to measure the pump current. The base level can be held constant while the top level is raised bit by bit. If full inversion is reached by the top level, pump current increases rapidly and saturates at further increased voltage in the CP-current, I_{CP}. Saturation might not be reached if no reverse bias is applied to source/drain regions. This effect is due to channel electrons which cannot flow back to source/drain and it mainly appears for large geometries [93].

In a similar way, the top level can be held constant while the base level is varied. Most popular, however, is a method where the amplitude is kept constant and the base level is swept through the voltage region. An example of such a measurement is shown in Fig. 2.14a. Trap density and capture cross section are calculated from the maximum of the pump current as pictured in Fig. 2.14b. In this case we gain $D_{\mathrm{it}} = 1.0 \cdot 10^{10}\mathrm{eV}^{-1}$ cm^{-2} and $\sqrt{\sigma_p\sigma_n} = 5 \cdot 10^{-16}$ cm^2.

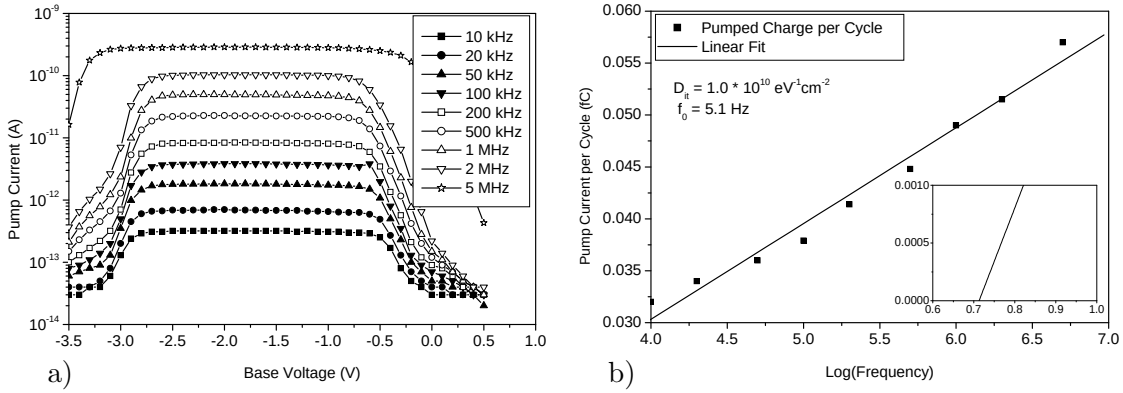

Fig. 2.14: Measured charge pumping currents as a function of base level with measurement frequency as parameter (a). Trap densities are extracted from the frequency dependence of the maxima (b).

The CP-method allows for a localized determination of trap densities. The active area can be varied with the reverse bias at the source/drain [33, 21]:

$$A_{\mathrm{G}} = W_{\mathrm{eff}} \cdot \left[L_{\mathrm{eff}} - 2 \cdot \sqrt{\frac{2\varepsilon_{\mathrm{s}}}{qN_0}(V_{\mathrm{r}} + 2\phi_{\mathrm{F}})}\right] \tag{2.19}$$

In addition, oxide traps near the interface can be examined [30, 8].

2.3.2 Drain-Current Charge Pumping

Recently a method has been proposed which is conducted in a way similar to classical charge pumping but applies a potential between source and drain [35]. This method has the advantage of the measured current being many orders of magnitude larger than the classical CP-current. The name *drain current charge pumping*, however, is somewhat misleading, because the charge pumping current should not play any role for this technique. Rather, the frequency dependence of the threshold voltage is measured. The gate voltage, V_{G}, is given by [35]:

$$V_{\mathrm{G}} = \phi_{\mathrm{s}} + \phi_{\mathrm{MS}} - \frac{1}{C_{\mathrm{ox}}}(Q_{\mathrm{i}} + Q_{\mathrm{d}} + Q_{\mathrm{it}}) \tag{2.20}$$

Here, ϕ_{s} is the surface potential and Q_{i}, Q_{d} and Q_{it} are inversion, depletion and interface charges per unit area. With constant gate voltage, interface traps thus lead to a shift in surface potential and consequently to a shift in threshold voltage. Contrary to the ideal case, traps cannot follow the signal at high frequencies and do not influence the threshold voltage. At low frequencies, traps have to be charged and emptied which leads to a shift in V_{T}. With constant gate potential, this shift is amplified in a variation of the drain current. Therefore, the top level of the gate voltage determines the sensitivity of the method.

In the original publication by Haddara, an ideal top level $V_{\mathrm{G}} - V_{\mathrm{T}} = 0.3V$ of the gate voltage has been proposed while the threshold voltage was estimated to be 0V. This values might have been estimated too low, since the typical value for V_{T} of a standard CMOS process lies usually somewhere around 0.5 V. Our observation revealed an optimum top level to lie below the threshold voltage, since the drain current varies exponentially in this regime.

If the measured drain current is plotted as a function of frequency while the rise and fall times are kept constant, a plateau is observed at low frequencies because all traps can follow the signal. For high frequencies, however, the threshold voltage is shifted because some traps cannot follow the signal anymore. The energy level of the traps measured is determined by the rise and fall time as described in the last section. Haddara predicted a second plateau in the drain current for high frequencies assuming that no traps can follow the signal at those frequencies [35]. However, this plateau has neither been observed by other authors nor in our own measurements [29, 35]. Actually it is not accessible experimentally since the cycle time cannot be less than the rise and fall times. The lack of such a plateau renders this method very sensitive to errors. Furthermore, the observed increase in drain current could not be attributed to a shift in flatband potential but rather to the classical charge pumping current. Altogether, this method does not appear to yield any new information.

2.4 Determination of MOSFET-Characteristics

Knowledge of some basic transistor properties is very valuable for the characterization methods presented so far. Apart from information on substrate doping and effective channel length it is possible to obtain information on the work function from the threshold voltage. These methods are described in the following.

2.4.1 Threshold Voltage of Transistors

The threshold voltage, V_T, of a transistor is an important quantity for the analysis of MOSFET data. There are different definitions of V_T given in literature [93]. Classically, threshold voltage is defined as that value at which the density of inversion charges in the channel equalizes the substrate doping, i.e. where $\Phi_s = 2\Phi_F$ [10]. This voltage is typically lower than the starting value for *Medici* simulations and also lower than experimentally extracted threshold voltages. There are various possibilities for the analysis of data which are more or less strongly affected by series resistance and reduced mobility. To reduce influence of the series resistance, measurements should be performed at small source/drain voltages, i.e. in the linear regime. The most simple way is drawing a tangent at the point of maximum slope of the transfer characteristics. The intercept of this tangent with the x-axis determines the threshold voltage. For long-channel MOSFETs this procedure yields reliable data as long as the source/drain resistance can be neglected. For many LDD-transistors the latter assumption is not valid and leads to an underestimation of the threshold voltage [93]. The choice of the appropriate method depends on the extraction of further quantities for which the threshold voltage is needed. In production, where uniformity and roll off are monitored, V_T is usually determined as the voltage at which a certain current $I_{DS} = (W/L)I_0$ is flowing [39], i.e.

$$I_0 = \mu_n \frac{\sqrt{\varepsilon_s q N_A}}{2\phi_F} \left(\frac{kT}{q}\right)^2 \tag{2.21}$$

This method requires the knowledge of some process-specific parameters. For new processes, the entire transfer characteristic has to be analyzed. The threshold voltage has been defined by that value, at which the drain current deviates from the exponential subthreshold behavior by a fixed percentage. As a further alternative, the derivative of the transfer conductance with respect to gate voltage $\partial g_m/\partial V_{GS}$ can be plotted as a function of gate voltage. The maximum of this curves is placed at $V_{GS} = V_T$ [118].

2.4.2 Determination of Substrate Doping

The threshold voltage of a MOSFET depends on the substrate voltage as follows [93]:

$$V_T = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F + V_{SB}} \tag{2.22}$$

where $\gamma = (2q\varepsilon_s\varepsilon_0 N_A)^{1/2}/C_{ox}$. This dependence can be employed to determine the doping profile of the substrate. If V_T is plotted as a function of $\sqrt{2\phi_F + V_{SB}}$, the slope yields the doping profile as a function of depth:

$$N_A = \frac{\gamma^2 C_{ox}^2}{2q\varepsilon_s\varepsilon_0} \tag{2.23}$$

$$W = \sqrt{\frac{2\varepsilon_s\varepsilon_0(2\phi_F + V_{SB})}{qN_A}} \tag{2.24}$$

Since Φ_F depends on N_A, the procedure has to be applied iteratively. A good starting value for NFETs is $2\phi_F = 0.6$ V [93]. The Fermi-potential is recalculated after one run using $\phi_F = kT/q \ln(N_A/n_i)$. Comparisons with SUPREM3-simulations from the literature showed a good agreement [93]. An operation amplifier circuit as shown in Fig. 2.15 is suitable for these kinds of measurements.

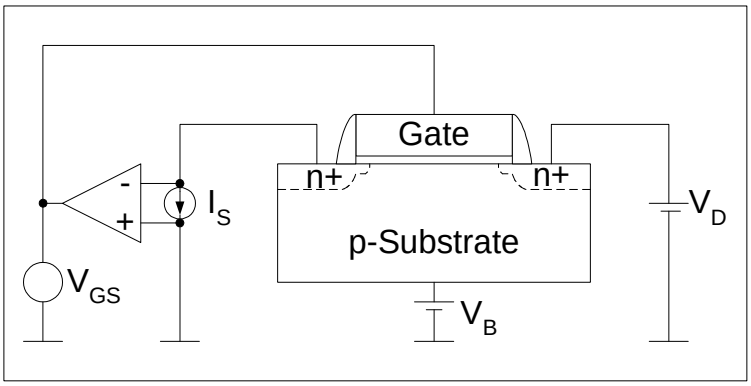

Fig. 2.15: Schematic setup of the measurements to determine the substrate doping.

A fixed current of approximately 1 μA is applied to the source while the substrate voltage is swept through a few volts. V_{T} is subsequently measured at each substrate bias. The exact value of the source current has only minor influence on the measurement results since only the change in V_{T}, but not V_{T} itself is used for analysis. In most cases, the measurement setup can be simplified significantly because modern parameter analyzers have built-in operation amplifiers. It is then sufficient to shorten source and gate, apply a current and measure the voltage at this unit. Drain can be set to zero volts while the substrate bias is varied. Data presented in this work were measured in this way.

2.4.3 Analysis of the Transfer Characteristics

In addition to threshold voltage, parameters like effective geometry, effective mobility, subthreshold slope and source/drain resistance are essential to describe the device. A method to derive these quantities from the transfer characteristics is described in the following. For comparison, different other methods can be found in the literature [73, 71]. To determine effective geometry at least two transistors of different lengths of widths are needed. The effective area is important to analyze CP-data. LDD transistors like those analyzed for part of this work can have an effective channel length larger than the metallurgical distance between source and drain because the channel can reach into the LDD-regions.

For high gate voltages with $(V_{\mathrm{GS}} - V_{\mathrm{T}}) \gg 0.5 V_{\mathrm{T}}$ the drain current can be described in the following way [93]:

$$I_{\mathrm{D}} = \frac{W_{\mathrm{eff}}\, \mu_{\mathrm{eff}}\, C_{\mathrm{ox}} (V_{\mathrm{GS}} - V_{\mathrm{T}}) V_{\mathrm{DS}}}{(L - \Delta L) + W_{\mathrm{eff}}\, \mu_{\mathrm{eff}}\, C_{\mathrm{ox}} (V_{\mathrm{GS}} - V_{\mathrm{T}}) R_{\mathrm{SD}}} \tag{2.25}$$

This equation is the basis for the determination of R_{SD}, μ_{eff}, L_{eff} and W_{eff}. In equation 2.25 the effective mobility can be approximated by $\mu_{\mathrm{eff}} \approx \mu_0/(1 + \Theta(V_{\mathrm{GS}} - V_{\mathrm{T}}))$. The measured resistance $R_m = V_{\mathrm{DS}}/I_{\mathrm{DS}}$ resolves in [93]:

$$R_m = \frac{L - \Delta L}{W_{\mathrm{eff}}\, \mu_0\, C_{\mathrm{ox}} (V_{\mathrm{GS}} - V_{\mathrm{T}})} + \frac{\Theta (L - \Delta L)}{W_{\mathrm{eff}}\, \mu_0\, C_{\mathrm{ox}}} + R_{\mathrm{SD}} \tag{2.26}$$

Usually, R_m is plotted as a function of $1/(V_{\mathrm{GS}} - V_{\mathrm{T}})$. Subsequently, the slope $m = (L - \Delta L)/W_{\mathrm{eff}}\, \mu_0\, C_{\mathrm{ox}}$ is plotted against different nominal transistor lengths and extrapolated to $m = 0$. Thereafter, ΔL and μ_0 can be determined.

In a similar manner, values for R_{m} can be plotted against m at the point at which $1/(V_{\mathrm{GS}} - V_{\mathrm{T}}) = 0$ and Θ and R_{SD} can be determined thereafter. For high accuracy the

channel lengths should vary by a factor of at least 10. A measurement over a large voltage range may be needed since the effective channel length depends on gate voltage. In general, extraction of parameters at low gate voltages yields the most reliable determination of V_{T} [103].

2.4.4 Analysis of Transistor Parameters

The measured subthreshold slope, S, is well suitable to determine the depletion capacitance. For an ideal transistor, S is given by the following expression:

$$S = \frac{kT}{q} \cdot \ln(10) \left[1 + \frac{C_{\mathrm{D}}}{C_{\mathrm{ox}}}\right] \tag{2.27}$$

If interface traps are present, S increases by:

$$\Delta S = \frac{kT}{q} \cdot \ln(10) \cdot \frac{qD_{\mathrm{it}}}{C_{\mathrm{ox}}} \tag{2.28}$$

This part can be determined by the knowledge of interface traps and oxide thickness from CP- and CV-measurements and can be subtracted from the measured subthreshold slope. The depletion capacitance C_{D} calculated via equation 2.27 can be applied together with theoretical values given in Section 2.8 to evaluate the depletion zone in the gate electrode. This method is most suitable for thick gate oxides.

2.5 Summary

Measurement techniques suitable to analyze metal electrodes were presented in this chapter. A method was developed to extract interface properties from CV-measurements on tunnel-oxide MOS structures. It allows fast interface analysis on a large number of measurements which is needed to generate wafer maps. Such maps are crucial to optimize a new metal deposition process. The chapter concluded with means of extracting physical parameters from fully integrated MOS transistors.

Chapter 3

Test Structures for Characterization

This chapter describes the layout and process flow of the test structures needed for the characterization methods described in Chapter 2. The standard CMOS process is unsuitable for the development of metal electrodes, so that a new process flow is defined which requires less effort in single process development. As a consequence, the classical gate stack etch is replaced by a simpler step which avoids the very demanding anisotropic dry etch of the gate material with high selectivity to the underlying insulator. This is particularly useful when state of the art tunnel-oxides are examined.

Many metals create a high stress during thermal treatment driven by a phase transition after deposition and by their difference in thermal expansion coefficient with respect to silicon. Hence, a reduction in thermal budget after gate deposition is indispensable. Here, this aim was reached by performing source/drain implants and belonging activation anneals before gate deposition. As a consequence, no self-aligned processes with deep-submicron structures can be fabricated. This, however, is not required for the investigation of material properties. If a new layer is included into an existing process, it is important to verify that this layer can be aligned to an already existing one. Alignment is done with special structures in the kerf of each chip. Space is very limited in this kerf so that only certain combinations of layers can be aligned to each other.

As described in Chapter 1, DRAM capacitors are one main application for metal electrodes. Process flow and test structures are described that allow the fabrication of simple deep trench capacitors where all high-temperature process steps can be performed before metal deposition. This is essential to study the thermal stability of the system.

3.1 Standard CMOS Process

The standard CMOS process is described in many textbooks and thus is demonstrated only briefly here. There are some variations due to CMOS-generation and manufacturer which influence device properties. The present work was based on a 0.25 μm process as shown in Fig. 3.1. At the beginning *active areas* (AA) were defined by the AA-mask and etched trenches filled with oxide. These trenches isolate active areas from each other and are usually referred to as *shallow trench isolation* (STI). Subsequent well implants define substrate doping of NFETs and PFETs and are followed by the deposition of gate insulator and gate electrode. After structuring and reoxidation of the gate, *lightly doped*

drains (LDDs) and pockets are implanted through the so called XP- and XN-layers. LDD implants facilitate a low source/drain doping at the transition to the channel, while pockets improve short channel performance through tilted implants under the gate. Side wall spacers of 100 nm thickness prevent dopands of the source/drain implants from diffusing too far into the channel. The doping of the gate is defined by these implants as well, so that the gate of NFETs and PFETs have different work functions.

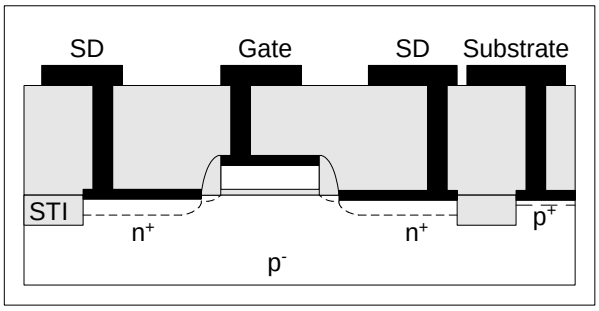

Fig. 3.1: Cross section of a 0.25 μm transistor as used in this work. The process is characterized by LDDs, pockets and titanium silicide and was processed up to the first metal layer.

Such a process is usually referred to as *dual work function* CMOS. A titanium silicide layer is formed to improve conductivity of gate and source/drain connections. The process is completed by standard metallization up to the first metal layer and by a passivation with photo imide. For simplicity, the latter is not shown in Fig. 3.1.

3.2 Process Flow of MIS-Structures

In the following, process flows developed during this work are described. Structuring of metal can be done either by dry or wet chemical etch or by *chemical mechanical polishing* (CMP). For both methods a process flow has been established in order to make best use of existing process experience with different materials. Beside gate electrodes also substrate electrodes were considered as they are found for example in DRAM capacitors.

3.2.1 Planar Capacitors with Metallic Substrate Electrodes

DRAM memory chips require capacitors with a capacity of 25 fF and a maximum leakage current of 1 fA per cell, whereas the available area is reduced from generation to generation. Beside surface enhancement techniques like *hemispherical silicon grains* (HSG) and new dielectrics metal electrodes are well suited to reach this aim.

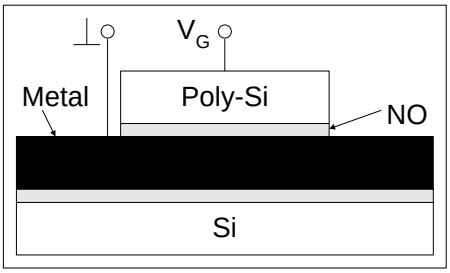

Fig. 3.2: Test structure for investigation of metal substrate electrodes. The gate electrode can by structured by a classical dry etch.

In order to modify the existing system as little as possible, at first only the bottom electrode and subsequently only the gate electrode are processed in metal. The former

is done by depositing a stack of metal, dielectric and polysilicon on blanket wafers and structuring the polysilicon with a standard dry etch process to form large area capacitors as shown in Fig. 3.2. No special masks are necessary since the standard *gate conductor* (GC) layer can be used for this purpose.

3.2.2 Planar Capacitors with Metal Gate Electrodes

This section presents processes wherein the top electrode is structured by CMP or by non-critical dry etch processes. A development of a usual dry etch process is only worth doing, if the material system has proven its functionality. Fig. 3.3 shows two variations facilitating a relatively simple fabrication of the test structures. In both cases a thick oxide is deposited in which large area holes are etched using a special mask (Fig. 3.3a). This mask is not part of the standard CMOS-process, but should be alignable to the AA-mask for sake of simple integration into the whole process. During this work the XP-mask was dedicated to this purpose and hence the thick oxide is referred to as XP-oxide in the following. The XP-mask is usually used for LDD implants of the PFETs but for our process these implants can also be performed via the source/drain layers.

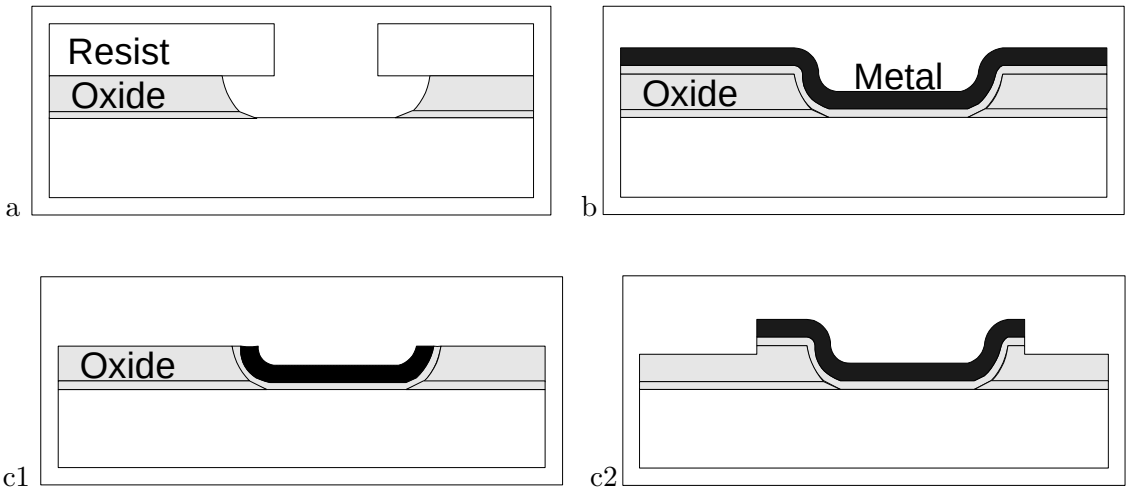

Fig. 3.3: Schematic process flow for the fabrication of metal gate electrodes. Openings which define the capacitor area are etched into a thick oxide (a). Subsequently, dielectric and gate electrode are deposited (b) and structured by either CMP (c1) or by dry etch (c2).

Etching of the XP-oxide is followed by deposition of dielectric and metal (Fig. 3.3b). Now there are two ways to structure the electrode. Firstly, the metal on the XP-oxide can be removed by a CMP-step as shown in Fig. 3.3c1. Support structures in the large area opening help to prevent over-polishing in these areas. Secondly, metal can be structured by a dry etch process as presented in Fig. 3.3c2. For this purpose, the GC-layer is appropriate which should be drawn somewhat larger than the openings in the XP-oxide. The etch process of the metal has to be neither selective nor anisotropic.

3.2.3 Modified CMOS Process with Metal Gate

The modified CMOS process is identical to the standard process up to gate oxidation. Then, deposition of the XP-oxide is followed by source/drain implants as shown in Fig. 3.4a. As described in the last section, holes are etched into the XP-oxide and gate dielectric as well as gate electrode deposited and structured (Fig. 3.4b). The titanium silicide (TiSi)-formation is omitted in the modified process, while the remaining metallization is analogue

to the standard process. Passivation (TV) protects the chip from erosion. A cross section of the final structure can be seen in Fig. 3.5.

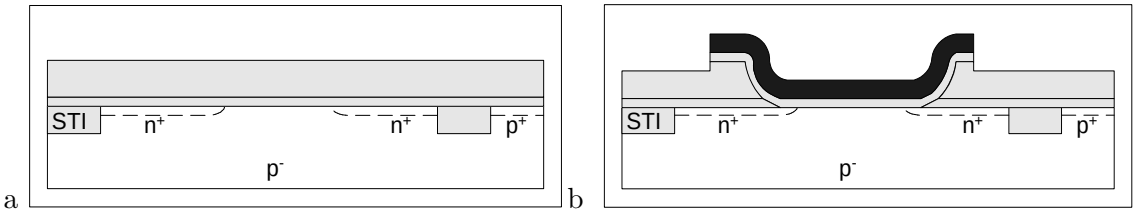

Fig. 3.4: Process flow of the modified CMOS process. Panel a) shows a cross section after well- and source/drain-implants and after deposition of the XP-oxide. Subsequently, openings are etched into the XP-oxide and gate dielectric as well as gate electrode are deposited and etched (b).

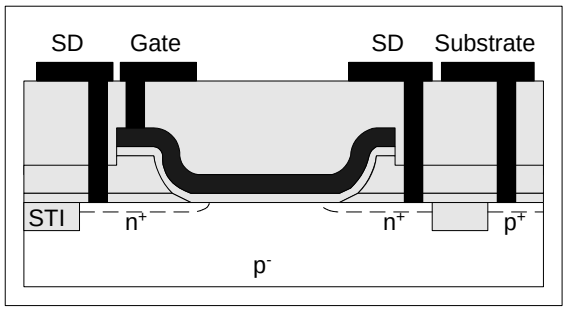

Fig. 3.5: Cross section of a transistor processed with the modified CMOS process.

3.2.4 Deep Trench Capacitors with Metal Electrodes

In a typical trench DRAM cell the charge is stored in a single deep trench capacitor. Single devices can be connected only when a full process is used to fabricate the test structure. Running a full process, however, is very expensive and not suitable for the development of the capacitor.

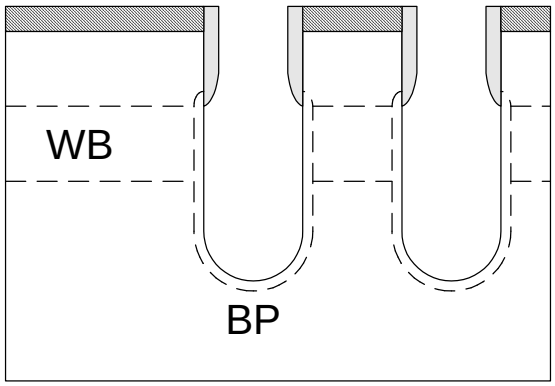

Fig. 3.6: Cross section of a deep trench short loop structure during processing showing deep trenches with oxide collar (grey), buried plate and buried well (hatched lines).

For this reason, so-called deep trench short loops (DTSL) are used where many trenches are connected in parallel. This process is very simple and allows for a short learning cycle. At the beginning, deep trenches are etches into the silicon-substrate and a so-called oxide-collar formed at the upper part of the trench. Below the collar, the trench surface is doped with n-type species like Arsenic defining the substrate-electrode (buried plate) of

the capacitor. All substrate electrodes are connected via a buried well (WB) implantation as shown in Fig. 3.6. The buried well is connected to the wafer surface by an additional n-well implant (WN) through the pad nitride (hatched in Fig. 3.6). A lithography step is done before this implant in order to reduce the pad nitride thickness as shown in Fig. 3.7a. The latter is required to reach a high impurity concentration at the wafer surface necessary for an ohmic contact during the electrical measurement.

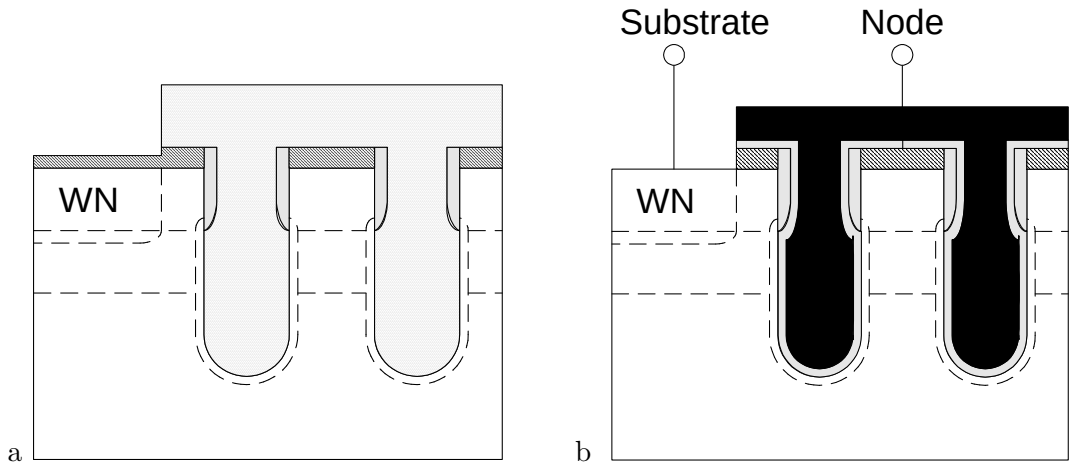

Fig. 3.7: Cross section of a deep trench short loop structure during processing after WN-implant (Panel a) and after deposition and structuring of node-dielectric and electrode (Panel b).

After resist strip, node dielectric (grey) and top electrode (black) are deposited and structured as shown in Fig. 3.7b. The top electrode can be polysilicon, metal or a polysilicon/metal stack.

3.3 Required Test Structures

This section describes the layout of the test structures which are needed for the development and which were implemented in the test chip developed in this work. Many of these test structures are equally suitable for characterization of dielectrics and of silicon process material.

3.3.1 Test Structures on Short Loops

Short loops are silicon wafers which are led through a part of the whole process only. This reduces the amount of analysis methods available, but shortens drastically the learning cycle during the development of metal electrodes or dielectrics. Most simple are planar capacitors which require only one mask as shown in Fig. 3.2.

For material characterization small capacitors of some 10^{-4} cm^2 area are sufficient for intrinsic properties and large area capacitors with 0.1 cm^2 for the determination of extrinsic defects. For process development and determination of parasitic effects more structures with different areas and varying area/circumferance ratios should be placed on the chip. Test structures as in Fig. 3.8a are most suitable for patterning with CMP. Support structures reduce dishing effects during polishing and depending on the CMP process should have a separation of 100 μm to 400 μm from each other. If neither a well developed dry etch process nor a CMP process for structuring of the electrodes are available, capacitors can be manufactured using two layers as seen in Fig. 3.3c2. The second layer can also be used to form a measurement pad with the benefit of the measurement needle

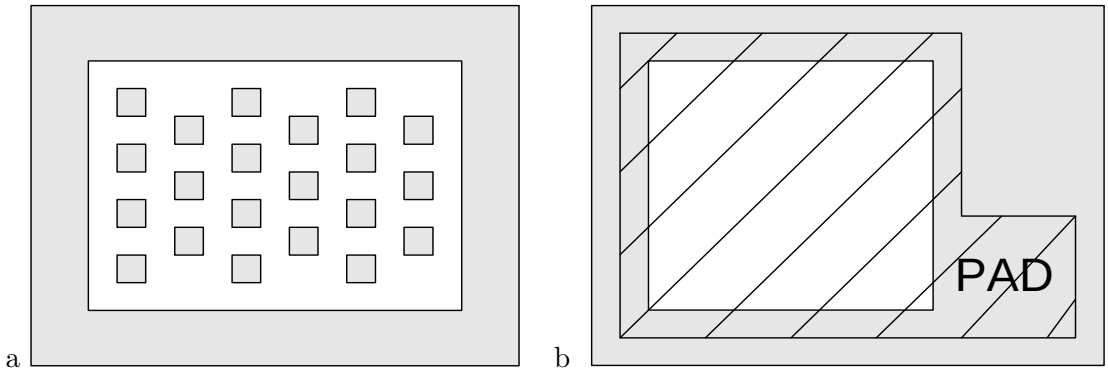

Fig. 3.8: Typical layouts of planar capacitors which are structured by CMP (a) or by dry etching (b). Rectangles drawn in (a) serve as support structures, which reduce dishing during CMP process. The XP-layer is shown in gray while the GC-layer is hatched.

not being placed on the active area. A layout of such a structure is shown in Fig. 3.8b. Only one additional mask was needed to facilitate all structuring methods mentioned so far. Since this layer is an implantation mask, no layers can be aligned on it. Therefore, the capacitor shown in Fig. 3.8b has to be fabricated with three layers. The AA-mask serves in defining the alignment marks, while XP- and GC-layers structure XP-oxide and electrode, respectively.

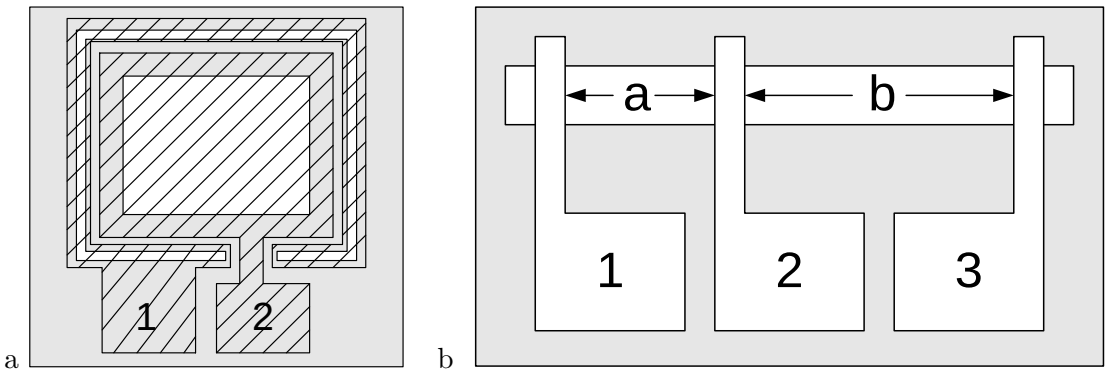

Fig. 3.9: Layout of a planar capacitor with guard ring (a) and of a structure to determine contact resistance (b). The XP-layer is shown in gray while the GC-layer is hatched.

For some analysis methods like C-t measurements, a guard ring is required as shown in Fig. 3.9a. In this way, diffusion of minority carriers from the capacitor to the surrounding area can be avoided. The guard ring is contacted via pad 1 and the gate electrode via pad 2. For the sake of clarity the substrate contact has been omitted in the picture. Some electrodes like the gates in a state of the art DRAM process consist of more than one layer like a polysilicon and a metal. An ohmic contact between these layers is indispensable for good device performance. The contact resistance can be measured with a structure shown in Fig. 3.9b. For each contact two pads might be drawn to allow 4-point measurements. The measured resistance between pad 1 and 2 and between pad 1 and 3, respectively are plotted as a function of the contact distance (a or a+b) and extrapolated to a distance of 0 μm. The gained value is simply twice the contact resistance.

3.3.2 Test Structures on Fully Integrated Wafers

A full CMOS process with one or more metal layers allows for the fabrication of single devices like transistors. Diffusion-limited capacitors as in Fig. 3.10a facilitate measure-

ments of LF-CV curves since minority carriers are always supplied by the diffusion region. The latter is connected via pad1 and the gate via pad2. The substrate contact has been omitted in the figure. IV-curves in inversion are produced from the same structures while extrinsic defects are usually examined on structures like those presented in Fig. 3.10b.

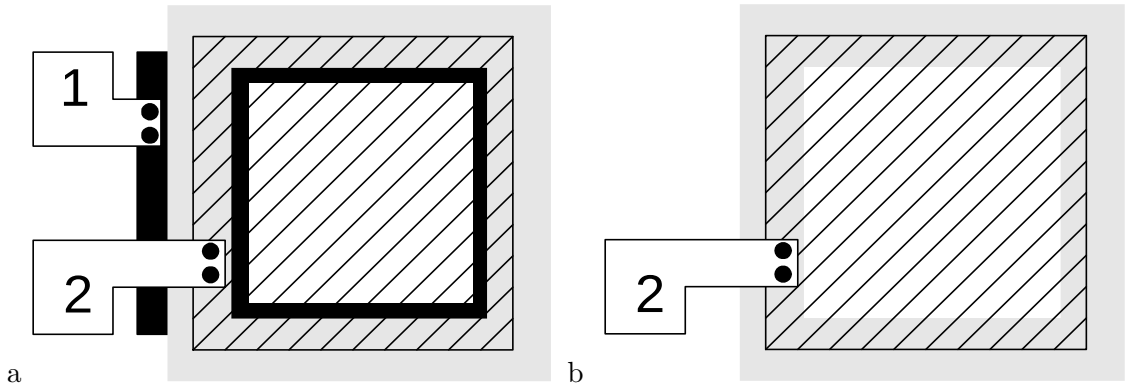

Fig. 3.10: Layout of a diffusion-limited (a) and a STI-limited (b) capacitor. The XP-layer is shown in gray while the GC-layer is hatched, the source/drain region being black and the metal layer being white. Contact holes are sketched by black dots.

In modern CMOS processes an extrinsic defect density of less than 0.01 defect per square centimeter is required, so that the area needed has to be somewhat larger than 100 cm^2 to determine defect density. The test chip developed in this work has an active area of 56 cm^2 for n-type as well as for p-type so that only a few wafers are necessary to investigate defect densities with high resolution.

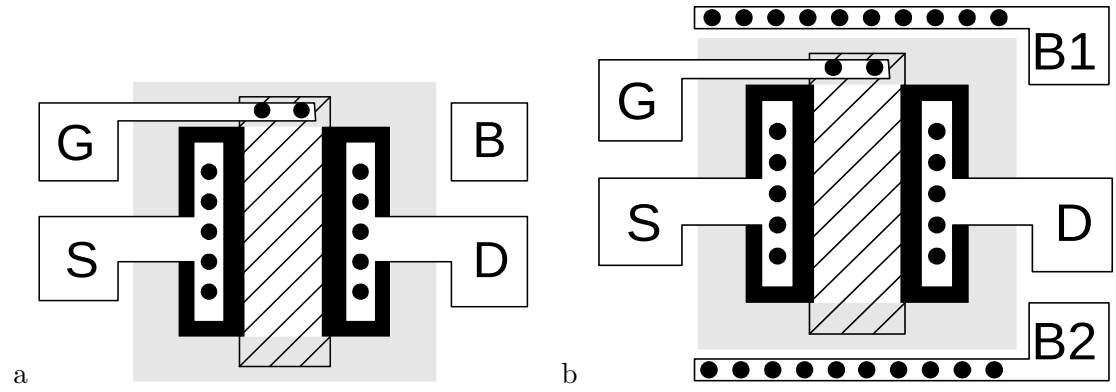

Fig. 3.11: Typical layout of single transistors which are suitable for parameter extraction and charge-pumping measurements (a). Panel b) shows a transistor with 2 substrate contacts for electrical localization of a soft breakdown. The STI is shown in gray while the GC-layer is hatched, the source/drain region being black and the metal 1 layer being white. Contact holes are sketched by black dots.

A series of transistors are required for the measurements described in the Chapter 2. Most suitable are NMOS and PMOS single transistors having four separate pads which are not shared with other test structures. A typical layout is given in Fig. 3.11a. Panel b) shows a version with two substrate pads. In this way, a soft breakdown can be localized orthogonal to the channel direction. A series of different lengths and widths is needed to determine the effective geometry, while large areas are preferred for charge-pumping measurements. For the latter, however, the channel length should not exceed 10 μm. For static measurements long channel transistors are most suitable when intrinsic parameters of the materials are to be extracted. Extrinsic defects of transistors are usually determined

on so-called multi-transistors in which some 10^6 transistors with the size of the considered logic generation are switched in parallel. The measurement averages over all transistors and indicates the moment at which the first transistor breaks down.

3.3.3 On-Chip Signal Amplification

High-accuracy techniques are needed to investigate physical effects on single devices. Important quantities like oxide thickness and flatband potential can be extracted for example from capacitance-voltage curves [96].

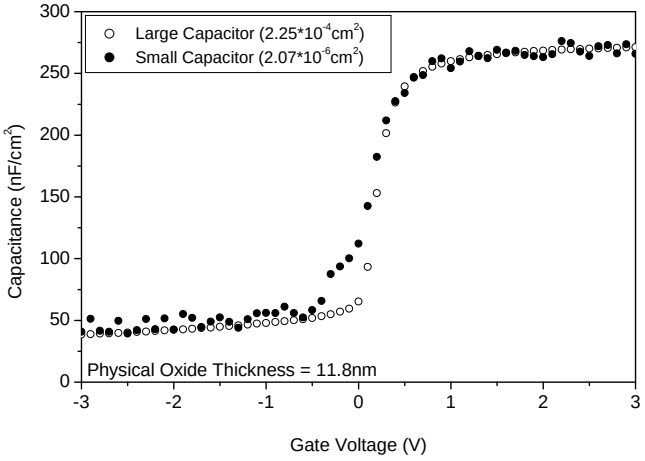

Fig. 3.12: Capacitance-Voltage characteristics as determined with standard high-frequency capacitance measurements. Measurements on small area capacitors (closed circles) lead to a significant noise level.

The latter are usually measured with an LCR-meter, where the impedance is determined by applying an AC potential to one electrode and measuring the resulting AC current at the other. Using standard measurement setups, significant noise is included when determining capacitances well below 1 pF as seen in Fig. 3.12. Here, the absolute value of the smaller capacitance is approximately 500 fF in accumulation.

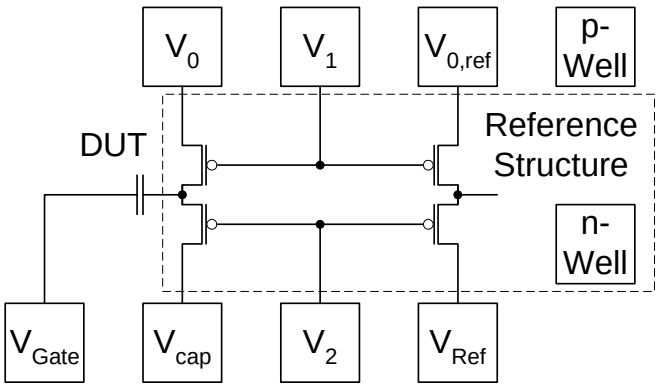

Fig. 3.13: Layout of the test structure used to measure capacitance-voltage characteristics. The DUT is a simple MOS capacitor with n-type substrate.

Recently, a charge-based capacitance measurement (CBCM) structure was proposed to measure interconnect capacitances with Femto-Farad resolution [19, 105]. Based on this method, a new technique was developed which allows measurement of whole capacitance-voltage curves with comparable resolution [94]. In addition a means to extract parasitic components during measurements is given which allows accurate measurements without

the need of reference structures. Test structures were fabricated with a standard 0.35 μm CMOS-process with 11.8 nm physical gate oxide thickness. A layout of the proposed test structure is shown in Fig. 3.13. The device under test (DUT) is a simple MOS-capacitor with n-type substrate which is used to pump current between V_0 and V_{Cap}.

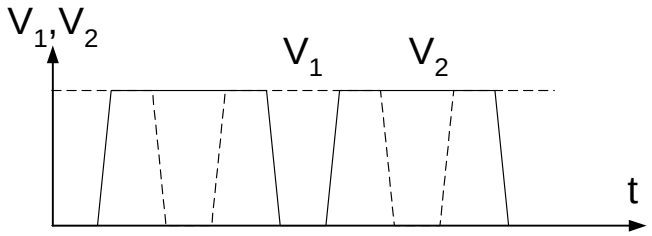

Fig. 3.14: Trapezoidal voltage signal as applied to the two PFETs used in the test structure.

Signals as shown in Fig. 3.14 were applied to two PFETs leading to a periodical variation of the DUT substrate potential between V_0 and V_{Cap}. If not stated otherwise, V_0 was set to 0 V and V_{Cap} to -100 mV. Trapezoidal pumping signals with an amplitude of 2 V and an offset of -1 V had a frequency, f, between 100 Hz and 1 MHz, a cycle time of $0.7/f$ and rise- and fall-times of $0.1/f$. Transistors were used only to switch the substrate potential of the MOS device between V_0 and V_{Cap}. Gate voltage, V_{Gate}, was swept from 3 V to -3 V to drive the device from accumulation to inversion and pump currents were measured successively at every voltage step. The effective gate voltage, V_{G}, can then be written as:

$$V_{\mathrm{G}} = V_{\mathrm{Gate}} + (V_0 - V_{\mathrm{Cap}})/2 \tag{3.1}$$

while the capacitance of the DUT is given by:

$$C_{\mathrm{DUT}} = \frac{I_{\mathrm{Cap}} - I_{\mathrm{Ref}}}{f * (V_0 - V_{\mathrm{Cap}})} \tag{3.2}$$

where I_{Cap} is the current measured at the pad labelled V_{Cap} and I_{Ref} is the reference current as determined by a separate method described below.

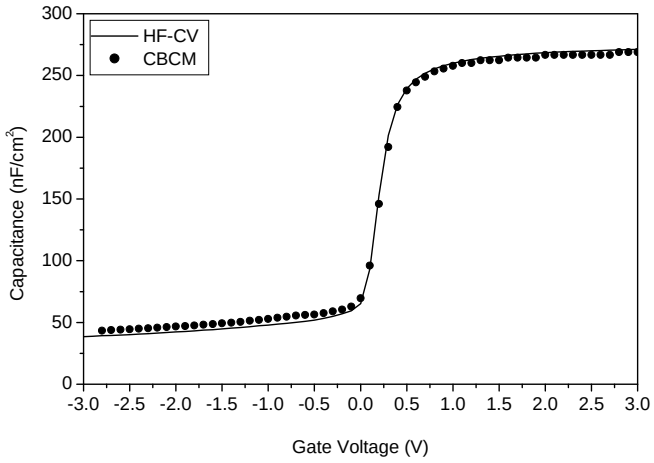

Fig. 3.15: CV-curves of a $2.25 \cdot 10^{-4}$ cm^2 capacitor as measured with standard high-frequency (HF) measurements and with the charge-based capacitance measurements (CBCM) described in the text.

Excellent agreement between standard high-frequency (HF) measurements and charge-based capacitance measurements (CBCM) indicate the basic functionality of the proposed method (Fig. 3.15). One main advantage of the new test structure could be higher

accuracy and less noise during analysis of small capacitances. The author, therefore, investigated capacitors with an area of $2.04 \cdot 10^{-6}$ cm^2 which have a maximum capacitance of around 500 fF. While large-area capacitors can be examined without post-treatment of data, reference measurements to determine parasitic effects are indispensable for smaller capacitors. Possible sources of parasitic currents can be understood by examining the cross-section of the test structure shown in Fig. 3.16.

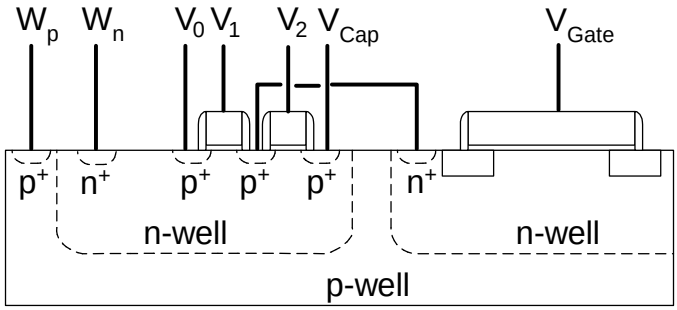

Fig. 3.16: Cross section of the proposed test structure for charge-based capacitance measurements.

Interface traps of the transistors lead to significant charge pumping currents which are proportional to the gate area of the PFETs. Hence, the latter should be as small as possible. In our case, transistors had an area of $1.8 \cdot 10^{-6}$ cm^2 which is comparable to the investigated DUT area. PFETs with a smaller gate area might be advantageous for this technique. A second source of error is the pn-junction between n-well and p-substrate. In an optimized test structure, the two contacts of the DUT might be exchanged by each other in order to eliminate this effect, although the substrate potential has to be set below -3 V in this case. A reference structure as shown in Fig. 3.13 can be used to measure the parasitic pumping current directly.

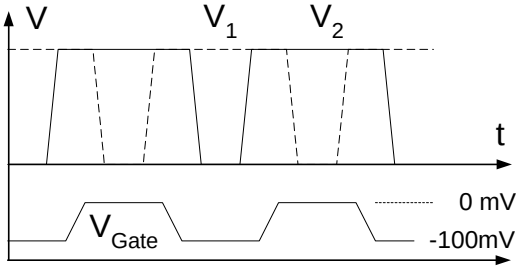

Fig. 3.17: AC voltage signal as applied to the gate of the DUT during the measurement of parasitic components.

Alternatively, reference measurements can be performed on the DUT as described in the following. To eliminate any charge-pumping contribution from the DUT, an AC signal can be applied to the gate which is setup in a way, that the voltage drop across the MOS capacitor is always 0 V as shown in Fig. 3.17. The AC signal at the gate has to be set up very carefully. Measured pumping current as function of the duty cycle (DC) is plotted in Fig. 3.18. Three distinct levels are observable: one below 20% DC indicating the overall pumping current for a static gate voltage of -100 mV, one above 80% DC indicating the pumping current for a static gate voltage of 0 V and a third level around 50% DC. This third level is identical to the pumping current resulting from all components but the DUT. The value at 50% DC was used as reference value for all subsequent measurements and was subtracted from all currents measured. CV-curves of small area capacitors measured at frequencies ranging from 100 Hz to 1 MHz are shown in Fig. 3.19a and compared to a

HF-CV curve obtained at 1 kHz. Values correlate well to HF-CV data but show less noise. In addition, analysis of interface traps is possible by examining the frequency dependence of the capacitance around -0.25 V.

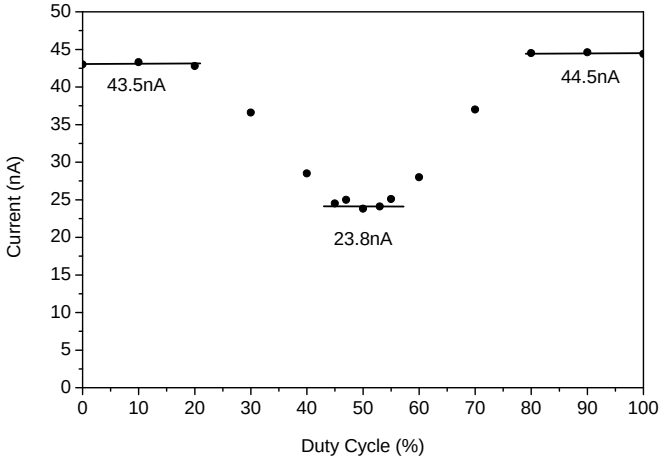

Fig. 3.18: Measured pumping current as function of the duty cycle of the signal applied to the gate of the DUT. The level around 50% was used as reference current.

The presented method is well suited to investigate degradation effects due to electrical stress. Fig. 3.19b shows CV curves for the same structure after applying an electrical stress. The stress has been induced by applying a high-voltage signal for several seconds to the DUT. As a consequence, the generation of interface charge which causes the shift in the CV-curve can be accurately determined by this method. Some variations in inversion and accumulation capacitance indicate that the test structure can be further optimized to reduce parasitic effects. Such optimizations could include, for example, reducing the area of the PFETs and exchanging gate and substrate contacts of the DUT.

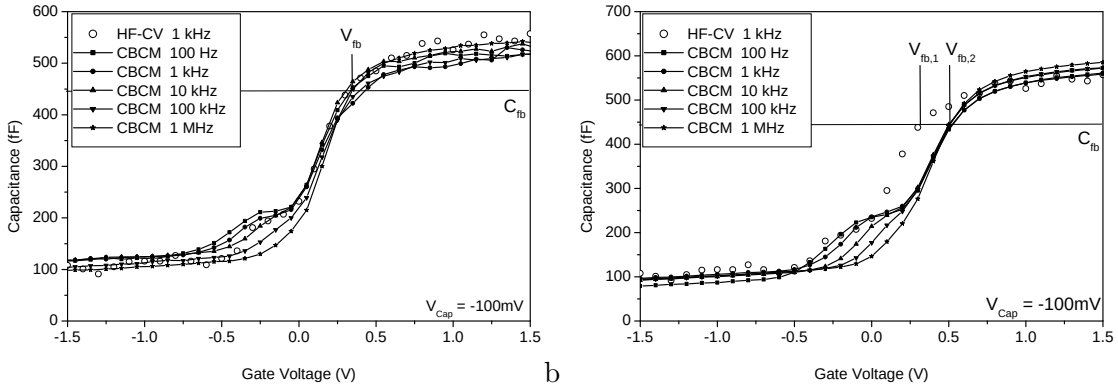

Fig. 3.19: Capacitance-voltage curves as determined by charge-based capacitance measurements at frequencies ranging from 100 Hz to 1 MHz (closed symbols). Good agreement with standard high-frequency measurements (open circles) is apparent. The structures were exposed to electrical stress leading to a flatband potential shift (b) as compared to unstressed samples (a).

The main noise component stems from the measurement devices of the test structure since the measured currents are rather high. As an estimate it is assumed that the noise level is roughly proportional to the area of the measurement devices. The fact that the measurement devices and the DUT have a comparable size in the presented investigations might suggest that the resolution of this technique scales down with device dimensions. A Femto-Farad resolution might therefore me achievable as long as the size of the measurement devices does not exceed the one of the DUT.

The amplification of small current is much more demanding. Commercial parameter analyzers have a resolution of 1 fA whilst many amplification circuits have a much higher off-current. At single transistors current amplification could facilitate CP-measurements at very small structures. However, no such structures were examined in this work.

One main reason to integrate metals into a CMOS process is the low resistivity. For planar applications like the gate electrode the resistance can simply be determined by 4-point measurements. If on the other hand metals are used as an inner electrode of a trench capacitor, no simple measurement is possible since the lower end of the metal fill cannot be connected (Fig. 3.20a). A ring oscillator (RO) circuit has been developed which determines the time constant, T_{RC}, of the trench fill as shown in Fig. 3.20b.

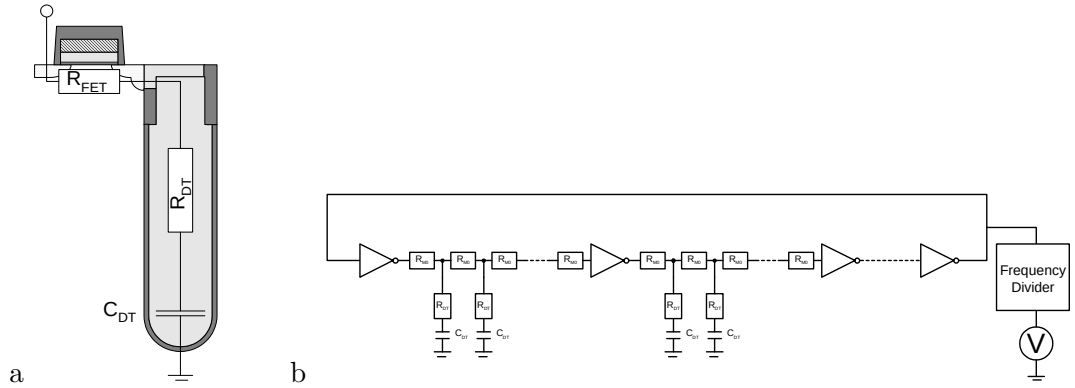

Fig. 3.20: Schematic of a trench capacitor (a). Since the lower end of the metal fill cannot be connected, the series resistance, R_{DT}, has to be determined via a dynamic measurement. An example of such a measurement structure is given in Panel b).

The oscillator frequency should be variable over a wide range of frequencies. For frequencies smaller than $1/T_{\mathrm{RC}}$, the measured capacitance can contribute to the RC-delay of the RO and thus reduces the oscillator frequency. For frequencies higher than $1/T_{\mathrm{RC}}$, the measured capacitance is separated electrically from the RO and does not reduce the frequency. Thus, T_{RC} can be determined directly from the nonlinearity of the oscillator frequency. A separate measurement of the capacitance yields the series resistance R_{DT}.

3.3.4 Requirements for the Mask Layout

The requirements to the mask layout are rather low, since only relatively large structures are needed for material characterization. Masks were fabricated with 0.35 μm design rules, but partly have deep ultra violet (DUV)-pellicles in order to use the masks with existing 248 nm steppers. *Pellicles* are protective covers which protect the masks from mechanical damage. A standard logic mask set from *AA* up to the first metal layer is sufficient for the described structures in which even the masks for the LDD implants and the salicide blocking mask can be omitted. After fundamental material properties are characterized and a material system has been proven to be feasible, further development has to be done on product-related test chips. These test chips should have design rules which correspond to the product the new material system is applied to.

3.4 Summary

Test structures suitable to characterize metal electrodes were described in this chapter. New test structures and a new process schematic were developed which allow the fabrication of metal gate CMOS transistors with low thermal budget after metal deposition. These test structures include simple capacitors suitable for short loops as well as fully integrated devices. Full integration facilitates on-chip signal amplification which might be required to study physical effects on single devices. A method to determine low-signal capacitance-voltage characteristics concludes this chapter.

Chapter 4

Technology Development for Fabrication of Metal Electrodes

This chapter discusses process technology that is required to fabricate the test structures mentioned so far. Planar capacitors as described in the first half of this chapter will be used to study interfaces and basic physical effects. The second half presents technology development that enables the application of metal- and polysilicon/metal-electrodes to deep trench DRAM capacitors.

4.1 Process Technology of Planar Test Structures

The following describes the process technology for planar test structures with metal substrate or gate electrodes which is required for interface characterization and the study of basic physical effects.

4.1.1 Capacitors with Metal Substrate Electrodes

As mentioned in Section 3.2.1, metal bottom electrodes are advantageous for the application in DRAM capacitors. In stacked capacitors as well as in trench capacitors metal is deposited at least partly on silicon. The interface between these two layers has to be thermally stable and has to supply an ohmic contact. To reach the latter, a minimum active dopant concentration of $5 \cdot 10^{19}$ cm^{-3} is required. In the following, deposition of substrate electrode and dielectric on blanket wafers is described. Deposition and structuring of the gate electrode will be described in Section 4.1.2.

Deposition and Cleaning of Substrate Electrodes

Using oxidized nitride (NO) as dielectric, only silicides can be used as metal bottom electrodes as will be explained in the following section. Presented results are therefore limited to tungsten silicide. A stack of 100 nm polysilicon and 55 nm tungsten silicide (WSi_x) has been deposited on oxide in an *Applied Materials Centura* single wafer tool using first silane and then tungsten hexafluoride (WF_6) and dichlorosilane (DCS) as precursors. The polysilicon layer is doped with phosphorus to a level of 10^{20} cm^{-3}, whereas the topmost 20 nm are undoped to avoid a tungsten-rich deposition at the beginning of the WSi_x-process. For some experiments the metal has been deposited directly on oxide. Two annealing atmospheres were tested to bring the WSi_x in the final tetragonal phase. Rapid thermal processing (RTP) at 1000 °C for 60 s was either done in an oxygen-rich atmosphere or in a pure nitrogen. Before NO-deposition, a hydrofluoric acid (HF)-dip

removed SiO_2 from the surface. Fig. 4.1 shows transmission electron microscope (TEM)-images of both treatments after deposition of NO and polysilicon. A thick oxide generated during recrystallization of WSi_x in oxygen is observed at the grain-boundaries which could not be removed by the HF-dip (A). Chapter 6 will show that this oxide leads to a reduction in capacitance. Enhanced oxidation at grain boundaries and the formation of metal oxides which cannot be removed by HF have already been reported by Lee et al. [56].

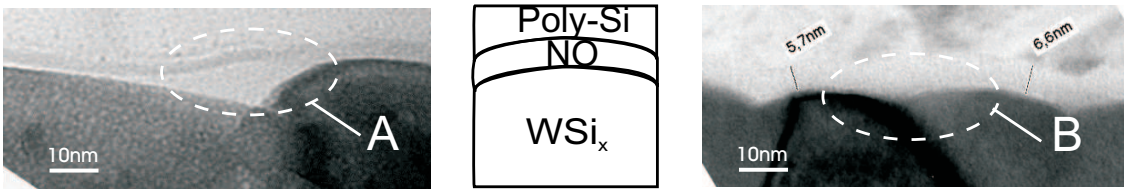

Fig. 4.1: TEM-images of the WSi_x/NO-interface. Recrystallization of WSi_x has been performed in oxygen-rich (left) and pure nitrogen (right) atmospheres. A sketch of the structure is shown in the middle. Annealing in oxygen leads to oxide-residuals at the grain-boundaries (A) while a pure nitrogen-atmosphere yields clean interfaces (B).

If annealing is done in pure nitrogen, no oxide is formed at the grain-boundaries, and the NO is deposited smoothly on WSi_x (B). This process will in the following be referred to as *passivation*.

Deposition of NO on WSi_x

A main question for the usage of WSi_x as substrate-electrode is whether the defect-healing mechanism of the NO works similarly as with silicon-electrodes. Usually, the substrate is nitrided in ammonia and subsequently a 4 nm thick low pressure chemical vapor deposition (LPCVD)-nitride is deposited. Reoxidation leads to growth of thick oxide at the defect sites in the nitride which avoids high leakage currents at those places. It is crucial to know whether pure SiO_2 is formed during oxidation of metal electrodes and in which way these processes depend on the stoichiometry of the WSi_x. Silicon is required in the substrate electrode for the formation of SiO_2. Hence, substrate electrodes can only be metal-silicides or metal-silicon-nitrides. Metal-oxides are usually not suitable for defect-healing mechanisms, since most of them have a high dielectric constant leading to a locally increased electric field at that site.

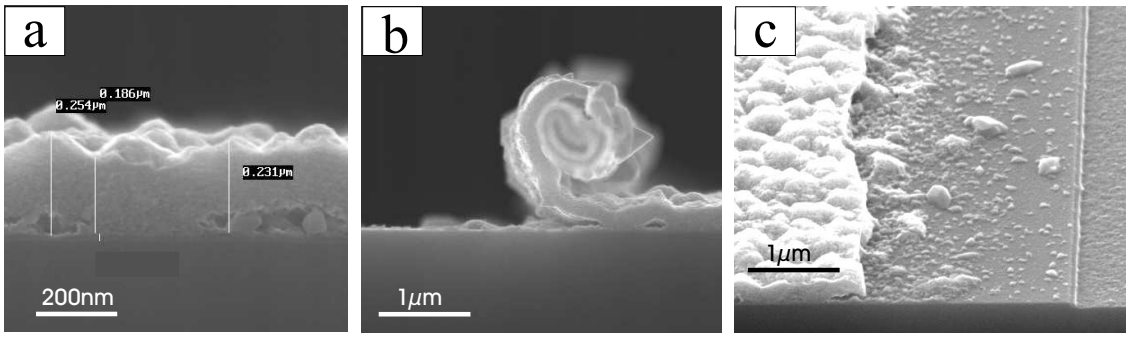

Fig. 4.2: SEM-image of a 60 nm $WSi_{1.9}$-layer, which was oxidized at 960 °C. During oxidation, the layer thickness has increased significantly (a) and high stress resulted in roll ups (b) and peeling (c).

Fig. 4.2 shows scanning electron microscope (SEM) images of a $WSi_{1.9}$-layer which has been oxidized at 960 °C. 60 nm thick homogeneous WSi_x mutated to a very rough layer with a factor 4 increase in thickness (a). This layer incorporates huge stress so that

it rolls up (b) or peels off completely (c). The oxidation mechanism can be studied very well with the defect-healing mechanism of NO. Table 4.1 summarizes the results of six different samples.

Layer	Substrate	Stoichiometry before annealing	Stoichiometry after annealing	Visible defects
WSi_x	Si	2.7	2.1	No
WSi_x	Si	2.3	2.1	No
WSi_x	Si	1.9	2.0	No
WSi_x	SiO_2	2.7	2.7	No
WSi_x	SiO_2	2.3	2.3	No
WSi_x	SiO_2	1.9	1.9	Yes

Table 4.1: Overview of tungsten-silicide samples at which the NO reoxidation has been tested. Visible defects indicate generation of tungsten-oxide.

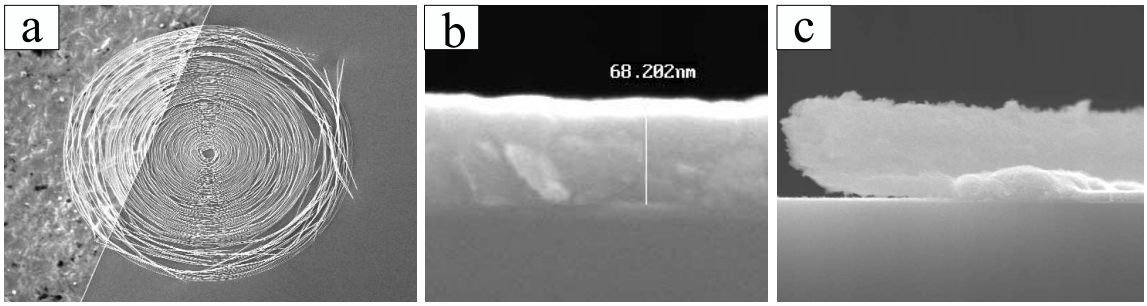

Fig. 4.3: Typical defect after reoxidation of node-nitride on a tungsten-rich WSi_x-layer. Panel a) shows a light microscope image of a defect with 1 mm diameter. In defect-free areas, the layer stack is fine (b), while WSi_x oxidizes at the defects, builds up high stress and peels off (c).

Light microscope- and SEM-images of a typical defect are presented in Fig. 4.3. The peeling that occurs at defects exposes further regions of the WSi_x-layer to the oxidizing atmosphere, so that the initially point-like defect grows to 1 mm in diameter. Such defects were observed only for WSi_x stoichiometries smaller than 2.0. This might lead to the conclusion, that WSi_x-layers with a stoichiometry of 2.0 or higher have a defect-curing mechanism similar to that of pure silicon. Polysilicon has been deposited and structured on top for electrical characterization of such layers.

4.1.2 Capacitors with Metal Gate Electrodes

If the gate-electrode is to be structured by dry etch, the process starts with AA-lithography and pertinent AA-etch. Silicon wafers are typically 10^{16} cm^{-3} p-doped. Samples, which are aimed to behave similar to DRAM-capacitors are produced on n-type substrates which are further doped by gas phase doping (GPD) to a level of $2-5\cdot10^{19}$ cm^{-3}. A reoxidation is followed by a XP-oxide deposition in which holes are etched through a resist mask using buffered hydrofluoric acid (BHF). Wet-chemical etches are mostly isotropic and at this point have the advantage, that the XP-oxide thickness increases only slowly at the transition from the opening to the point of full thickness. An abrupt increase would lead to a very high local stress in this region during subsequent high-temperature treatments. A schematic cross-section of the test structure after BHF-etch is shown in Panel 1) of Fig. 4.5b.

Even the moderate transition leads to a higher stress as shown in the TEM-images of Fig. 4.4. Panel a) shows a significant thinning of the gate oxide from approximately 3 nm

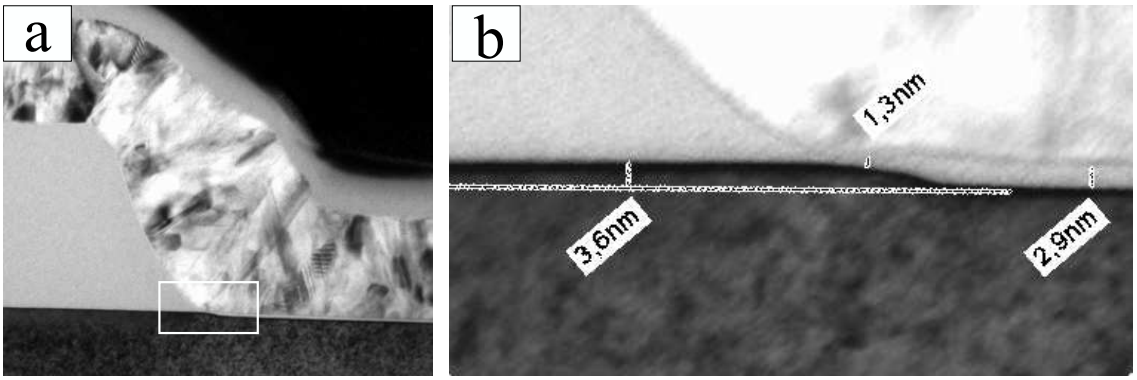

Fig. 4.4: TEM-image of the XP-oxide edge. Panel a) shows the transition from the XP-oxide (left) to the opening (right). The white rectangle indicates the detail presented in Panel b).

to 1.3 nm. Such thinnings drastically increases the gate-leakage current and would make all electrical results useless. Fig. 4.5a shows the measured leakage current of a 4 nm gate oxide with different XP-oxides. 300 nm silane oxide with or without underlaying thermal oxide leads to high leakage currents, while 30 nm thermal oxide yields good values. A stack of 7 nm thermal oxide and 55 nm TEOS is even preferable. TEOS is a SiO_2 which is deposited using tetraethylorthosilicate (TEOS) as precursor. Fig. 4.5a indicates that samples completely without XP-oxide show good leakage currents but as mentioned before this process only works on polysilicon gate-electrodes.

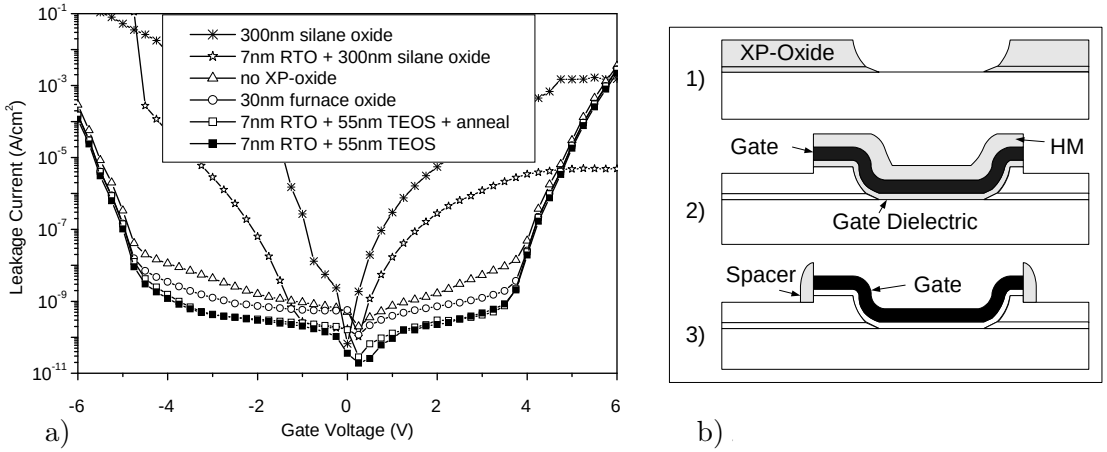

Fig. 4.5: Gate-leakage of a 4 nm thick thermal oxide for different XP-oxides (Panel a). The measured leakage current for the 300 nm silane-oxide is proportional to the perimeter and not to the area of the test structures indicating that the main gate-leakage results from the edge of the XP-oxide. Panel b) shows a schematic cross section of the test structure after BHF-etch of the XP-oxide (1), after structuring of the gate electrode (2) and after deposition of spacers and removal of the nitride hard-mask (3).

In summary, 300 nm silane-oxide is unsuitable for oxide characterization, while 55 nm TEOS on 7 nm thermal oxide give good results. In the latter case, leakage currents depend only on area and not on the perimeter of the capacitors, so that high leakage currents at the edges can be ruled out. This stack has been used as XP-oxide for all samples with SiO_2 as dielectric. The standard preclean before gate oxidation removed that stack completely and had to be exchanged by a less aggressive clean. Deposited dielectrics like NO or aluminium-oxide (Al_2O_3) do not show the thinning effect at the edges of the capacitor, since deposition rate depends only weakly on the stress in the substrate. After deposition of gate dielectric and electrode, the latter can be structured by dry etching using a resist

mask. As an alternative, a nitride hard mask deposited and structured directly on the metal might be preferable. After resist strip, the metal is then structured via the patterned silicon-nitride on top as shown in Panel 2 of Fig. 4.5b.

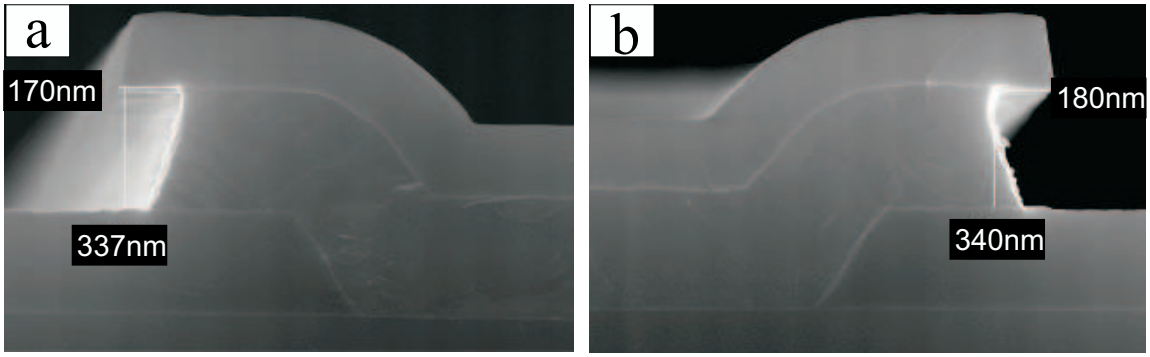

Fig. 4.6: SEM-image of the edge of a typical test structure. Shown are XP-oxide (bottom), a 340 nm thick polysilicon layer (middle) and 200 nm silicon-nitride on top. The under-etch below the hard mask was around 180 nm.

Fig. 4.6 shows a 340 nm thick polysilicon-layer which has been structured by an almost isotropic SF_6-etch using a nitride hard mask. There are many isotropic etch-chemistries for metals which are selective to silicon-nitride, so that the integration scheme can be used for a number of metals under consideration. Usually, LPCVD-nitride is used for this purpose, but a plasma enhanced (PE)CVD-nitride is deposited at lower temperature leaving more flexibility in subsequent temperature treatments. Problems occurring during this process are summarized in Fig. 4.7.

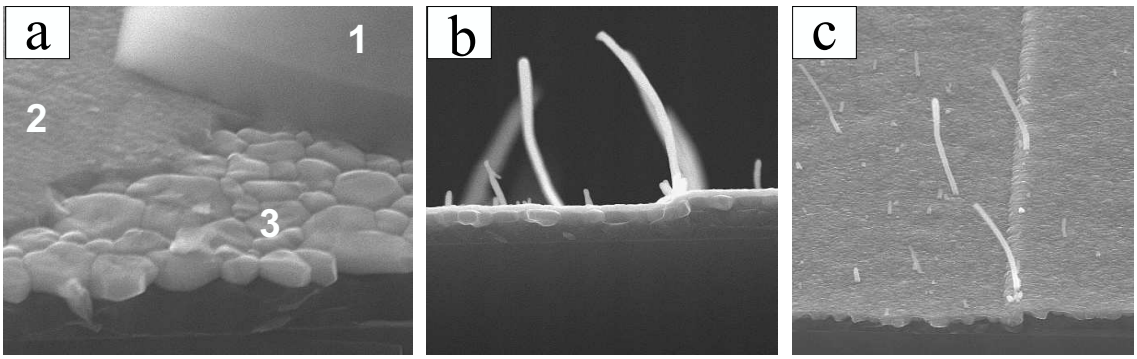

Fig. 4.7: SEM-images of process problems occurring during the usage of a 30 nm Oxide/170 nm plasma enhanced (PE)CVD-nitride stack as hard mask. Panel a) shows intact areas (1) as well as regions where nitride (2) or oxide (3) peeled off. Tungsten-rich needles that grew during annealing are demonstrated in Panels b) and c).

30 nm silicon-oxide serves as etch stop for the strip of the nitride-mask later on. During annealing both nitride and oxide can crack off the WSi_x (areas 2 and 3 in Panel a). In addition long tungsten-oxide needles are formed. Due to these problems, no further samples have been prepared with this hard mask. All samples analyzed in Chapter 5 and 6 were fabricated with a LPCVD-nitride hard mask or the gate was structured directly with a resist mask. Usually, a wet-chemical clean is performed after gate-stack etch. A 5-minute clean with $H_2O/NH_4OH/H_2O_2$ works very well for polysilicon- and WSi_x-gate electrodes. Tungsten-nitride (WN) on the other hand is etched very fast with this chemistry, as shown in Fig. 4.8.

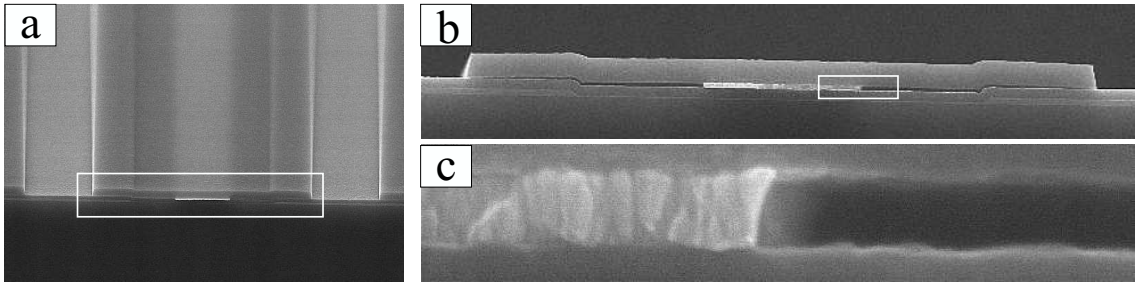

Fig. 4.8: SEM-images of a test structure with WN gate electrode and silicon-nitride hard mask after a 5-minutes $H_2O/NH_4OH/H_2O_2$ clean. WN has been etched 4 μm per edge (Panels a and b). The etch rate is a factor of 20 higher than the WN-etch rate on blanket wafers indicating an anisotropic etch behavior. The clean edge in Panel c) demonstrates the uniformity of this etch process.

The hard mask has to be removed at the end of the process, in order to contact the gate with a measurement needle. Hot phosphoric acid etches nitride very selectively to oxide and silicon. WSi_x, on the other side, is attacked quite strongly as shown in Fig. 4.9.

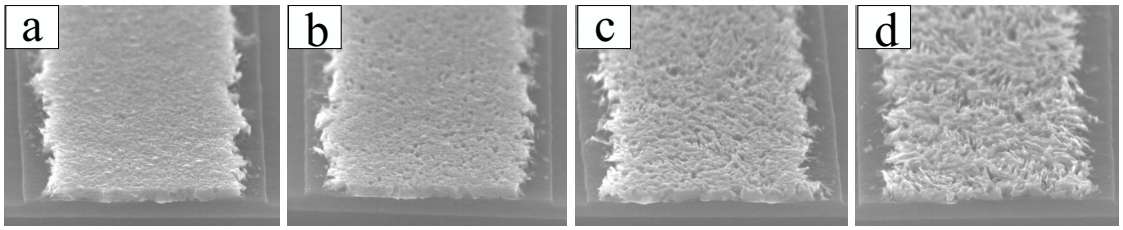

Fig. 4.9: SEM-images of a WSi_x-layer after nitride etch with different over etch times. All samples exhibit a strong erosion of the edges. 10% (Panel a) over-etch does not harm the WSi_x while 20% (b), 30% (c) or 40% (d) over-etch lead to severe damage.

Depending on the nitride over-etch time, the WSi_x is more or less strongly attacked, while the edges are always heavily eroded. The latter was avoided later on by the deposition of an oxide spacer prior to nitride etch (compare Panel 3 in Fig. 4.5b). After 10% over-etch (a), the WSi_x-layer is smooth while clear damage is visible for 20% (b), 30% (c) and 40% (d) over-etch.

If a CMP-process is available for the gate-electrode, fabrication of capacitors can be simplified as shown on the right hand side of Fig. 4.10. No AA-lithography is required for these samples and the process can be started directly with depositing the XP-oxide. However, a thicker XP-oxide is required to account for the dishing effects during polishing. A 300 nm silane-oxide is sufficient for this purpose, but as mentioned earlier, this layer is not suitable for thermal oxides. For deposited gate dielectric like NO or Al_2O_3, on the other hand, this process gave good results. As an example, Fig. 4.10 shows a microscope-image of planar capacitors having a WSi_x-gate that was structured by a CMP-process.

In summary, a process technology was developed that enables the fabrication of planar test structures with metal gate electrodes. Two options for structuring the top electrode were presented both of which are meanwhile regularly used for material development. While dry etching of the gate is more favorable for metal gates, the CMP-process has advantages for the processing of wafers that include high-k dielectrics. Due to contamination issues, the latter should be processed with as few process as possible after the deposition of the dielectric. Structuring with CMP only requires polysilicon deposition, anneal and the polishing step after dielectric deposition. The dry etch process, on the other hand, was used on samples that were integrated further to fully integrated wafers. These allow

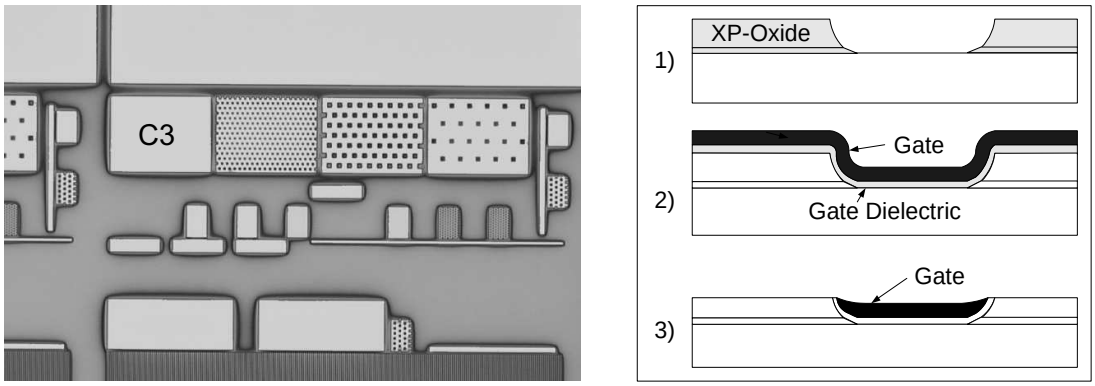

Fig. 4.10: Light microscope-image of a typical test structure which was structured by CMP (left). Visible are XP-oxide (dark grey) and tungsten-silicide (light grey) in the openings of the XP-oxide. For an electrical characterization the capacitor C3 with an area of $1.2 \cdot 10^{-3}$ cm^2 is well suitable. The right hand side shows a schematic cross section of planar capacitors after BHF-etch of the XP-oxide (1), deposition of gate dielectric and electrode (2) and after polishing (3).

the fabrication of transistors and diffusion- and STI-limited capacitors. The latter are regularly used for reliability analysis.

4.2 Process Development for Deep Trench Capacitors

Conformal deposition of metals into deep trenches is one of the major challenges when fabricating metal electrode capacitors for deep trench DRAMs. The main focus of this section, therefore, lies on the deposition and etching of metals in high aspect ratio trenches. TiN-, WN- and WSi_x-deposition techniques have been developed which resulted in a good step coverage suitable for DRAM application.

4.2.1 Deposition Techniques

Evaporation followed by condensation is the simplest deposition technique. The material has to be heated above the sublimation temperature in order to generate enough vapor pressure. Refractory metals and their silicides and nitrides have melting points above 2000 °C which makes this method unsuitable. During *sputtering* the material to be deposited is in a solid phase and bombarded with high-energy ions. The sputtered ions diffuse towards the silicon wafer and are adsorbed at the surface. This technique is regularly used in production since deposited layers have a high purity and stick well to the surface. A drawback is the limited step coverage on surfaces with high topology. For this application, deposition from the gas phase is more suitable. During this so-called *chemical vapor deposition* (CVD) the material is bound in more or less complex molecules. These *precursors* react at the substrate to generate a layer on the surface while gaseous reaction products are pumped out of the reaction chamber. A significant amount of impurities is usually incorporated in the layer, but potentially a good step coverage is feasible. Recently, *atomic layer deposition* (ALD) has been introduced into semiconductor industry, which is especially suitable for the deposition of thin layers with excellent step coverage.

Requirements on Deposited Layers

Since the elimination of the aluminium gates refractory metals such as tungsten or titanium as well as their nitrides and silicides were studied extensively for gate electrode application. The gate material has to stick well to the dielectric and has to survive high-temperature

source/drain activation anneals. Only minor stress is allowed at the interface to the dielectric and no impurities should diffuse into the latter. In addition, a good anisotropic dry etch with high selectivity to the gate dielectric has to be developed. Due to these requirements, only stacks of polysilicon and metals are in production so far. However, by optimized deposition techniques and reduced thermal budgets, these problems might be solved in the future. The application of metals in trench capacitors is even more demanding since the deposition has to be done on extreme topologies and the thermal budget of subsequent steps is always higher than in gate applications.

Plasma Processes for Deposition

Many processes like sputtering or the initiation of some metal-ALD processes require a plasma in the deposition chamber. The main features of such plasmas are described in the following.
A radio-frequency (RF) or DC-field is used to generate ions in a gas consisting of e.g. O_2-, Ar- or H_2-molecules. These ions are accelerated in the electric field to produce even more ions until a stable plasma with a constant ion density is burning. A so-called *magnetron* is sometimes used to increase the ion density by forcing the ions to circular motion. Between the plasma and the wafer surface the plasma bias drops across a dark space. The plasma bias increases with increasing RF-power and defines the maximum energy of the ions hitting the wafer surface. A minimum number of ions is required to keep a plasma burning. Besides a certain chamber volume, RF-plasma instead of DC-plasma and a high chamber pressure of some 10 mTorr can be used to increase the number of ions generated. RF-plasmas are very sensitive to water contamination, since the dipoles can strongly distort the electric field. Hence, modern tools often contain a separate degas chamber and cryo-pumps instead of turbo-pumps since the former have a much better pump efficiency for water. However, when reactive gases are used in the process, cryo-pumps are no possible choice [117]. During the initiation of a plasma there is an overshoot with very high energy ions which can harm the wafer surface. If the plasma is initiated between the target and a shutter which is opened afterwards, all very high energy ions are trapped by this shutter before the actual process starts.

4.2.2 Tungsten-Nitride ALD

Atomic layer deposition (ALD) exists for more than 20 years, but has been discovered only lately as a deposition technique for microelectronics. At least two different precursors are used that, in contrast to CVD, are not fed into the reactor at the same time. Deposition is divided into cycles, in each of which roughly one monolayer is deposited. The four phases of one cycle are sketched in Fig. 4.11.

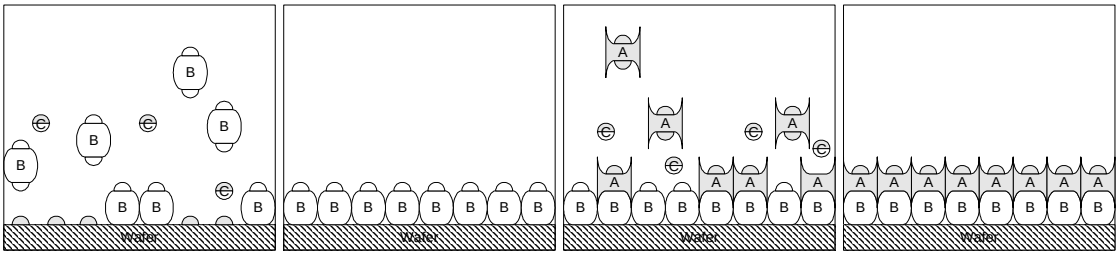

Fig. 4.11: Schematic of the 4 phases during an ALD-process. In the first step, the substrate has to be prepared in a way that precursor B can be adsorbed. The reaction product C is pumped out in the second step together with the remaining precursor B. During the third step precursor A is fed into the chamber while the remaining precursor with the reaction product C is pumped out in the fourth step.

Before deposition the surface has to be prepared in a way that one of the precursors can be adsorbed. This is usually done by nitridation or plasma activation inside the ALD-chamber. During phase "I" of each deposition cycle, the chamber is flooded with tungsten-hexafluoride (WF_6 - "B" in Fig. 4.11), which can react for example with hydrogen-terminated surfaces through the generation of HF (C). After deposition of one monolayer no further WF_6-molecules are adsorbed, since fluor-terminated surfaces cannot reduce these molecules. In phase "II" the chamber is purged with an inert gas like argon to completely remove the WF_6 and reaction products. Phase "III" is analogue to the first phase with the exception that ammonia (NH_3 - "A" in Fig. 4.11) is pumped into the chamber and adsorbed while generating HF. During the final phase "IV" the chamber is purged and the process starts all over again. The presented process has been optimized on a *Lynx2*-system built by *GENUS*. In this case, one cycle takes several seconds and deposits 0.4 monolayers of WN.

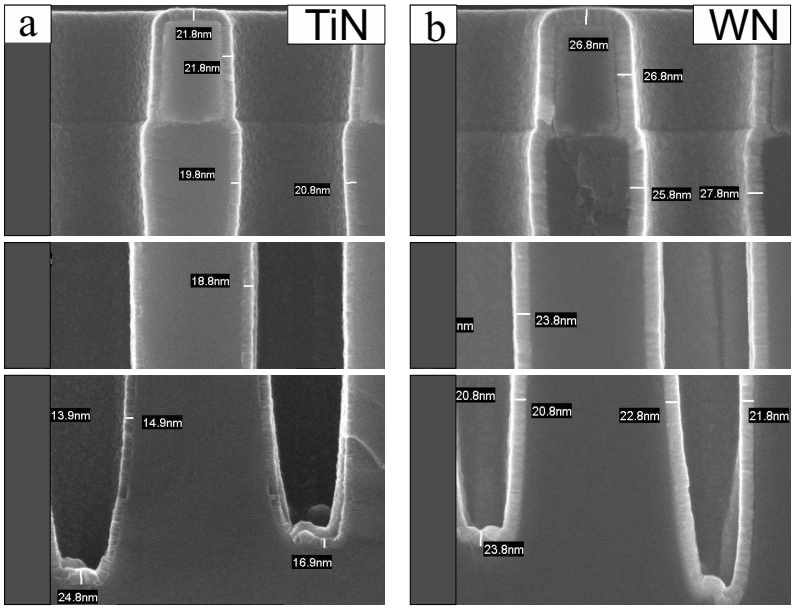

Fig. 4.12: SEM-images of ALD TiN (a) and WN (b) layers that were deposited in deep trenches. Trenches have an aspect ratio of 40:1 and are filled with very good step coverage.

If silicon is used as substrate, tungsten nitride (WN) cannot be deposited directly on top of it, since WF_6 is reduced by the substrate and thus etches the silicon. In this case, titanium-nitride (TiN) is deposited as a diffusion barrier before WN. For this application an ALD-process with titanium-tetrachloride ($TiCl_4$) and ammonia was used, while HCl was generated during the process instead of HF. Fig. 4.12 demonstrates the excellent step coverage of the processes on trenches with an aspect ratio (AR) higher than 40:1.

4.2.3 Tungsten Silicide Deposition

In this section CVD of WSi_x into features with high ARs is discussed. While the focus is on deep trench plugs of DT-DRAMs (as manufactured by Infineon Technologies) all results presented here also apply to vias of stacked (St)-DRAMs with only minor modifications. All simulations that will be presented have been developed and conducted by Georg Schulze-Icking of Infineon Technologies in Munich [95].

Key requirements for a WSi_x-plug are thermal stability in contact with polysilicon up to frontend processing temperatures (T≈1400 K), good step coverage in trenches with aspect ratios AR≥50:1, and sufficiently uniform stoichiometry along the depth of the DT. The

only thermodynamically stable phase in contact with silicon is WSi_2. While silicon-rich WSi_x-films ($x > 2.3$) exhibit a smooth film-to-substrate interface [91], they are thermally unstable and segregate the excess silicon as silicon crystallites or increase the thickness of the underlying polysilicon [108]. Tungsten-rich films ($x < 2.0$), on the other hand, are also thermally unstable and lead to the consumption of the underlying polysilicon layer during anneal steps later in the process [20]. Thus the composition of tungsten silicide layers in frontend processes should always be larger than 2 and ideally around 2.3. Note that slightly less stringent limitations apply to vias in St-DRAMs since those are not subject to high temperature anneals.

This section demonstrates a procedure to choose appropriate precursors and process conditions for tungsten silicide CVD into structures with extreme aspect ratios. An optimized process with both good step coverage and suitable stoichiometry for frontend applications will be proposed.

Growth Kinetics

Generally, CVD can be divided into four steps: gas phase reactions of precursors in the deposition chamber (A), adsorption of reactive molecules at the surface (B), surface diffusion of adsorbed species (C), and finally the surface reaction leading to bulk deposition (D). While each of these steps can be present in a CVD process it has been shown that surface diffusion (step C) is negligible for tungsten deposition [92]. Furthermore, the remainder of this section concentrates on feature scale aspects, assuming that step A has already taken place.

Under the above assumptions, steps B+D determine the (local) deposition rate and thus the step coverage of the process. At low temperatures, the deposition rate is generally limited by the surface reaction rate. This condition is referred to as the "reaction-limited" process regime. On the other hand, at high process temperatures the surface reactions can be so fast that the transport of precursors to the surface is the deposition rate (DR)-limiting step. This regime is generally referred to as "diffusion-limited" (or equivalently "transport-limited"). This latter regime is usually employed for deposition on planar wafers since under these conditions the process temperature has only minor impact on the deposition rate, and also because the throughput is maximized. However, for wafers with high aspect ratio structures (such as DTs) rapid surface reactions lead to precursor depletion inside deep trenches and, therefore, to a poor step coverage. A simplified way to look at surface reactions is assuming first-order reactions only. In this case the surface reaction rate can be expressed by the sticking probability η for particle-surface collision. A large sticking probability (η=1) means that precursor molecules stick to the surface after the first collision. If, on the other hand, η is very small it takes an average of $1/\eta$ collisions before precursor molecules stick to the surface. Since particles can only reach the bottom of a DT after *many* collisions with the sidewalls it is evident that a good step coverage is only achieved if η is very small. This is equivalent to the extreme reaction-limited process regime.

In a gas the mean free path length λ is given by $\lambda = kT/(\sqrt{2}\sigma p)$ with k being the Boltzmann constant, T the gas temperature, σ the collision cross section of the molecules, and p the total pressure. Inside a DT, on the other hand, the mean path length between collisions is approximated well by the trench diameter, provided the latter is smaller than λ. In this so-called "Knudsen regime" the step coverage of a process has been calculated analytically [101]:

$$\frac{t_B}{t_S} = \frac{1}{\cosh\phi + \frac{\phi W}{2L}\sinh\phi} \qquad \text{with} \qquad \phi = \frac{L}{W}\sqrt{\frac{3\eta}{2}} \tag{4.1}$$

where η is the sticking coefficient and the other quantities are as explained in Fig. 4.13.

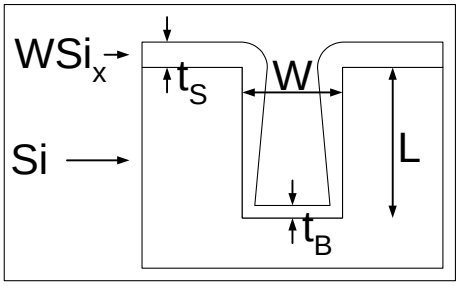

Fig. 4.13: Schematic of a layer deposited into a trench and definition of the variables in equations 4.1.

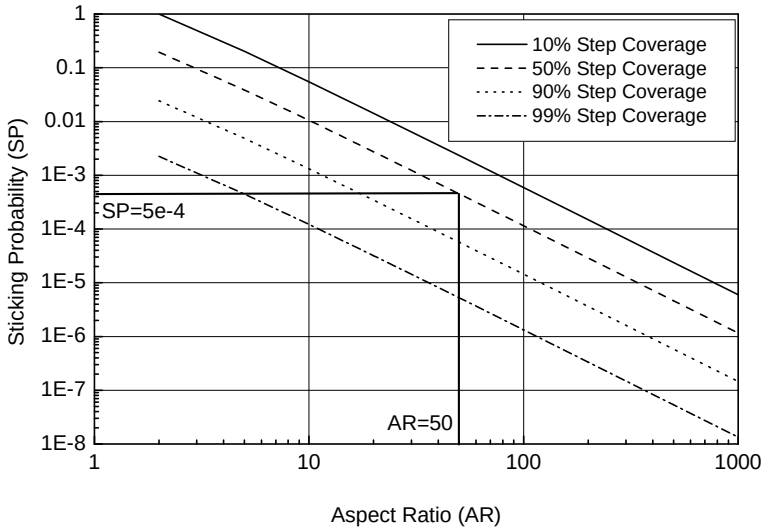

Fig. 4.14: Sticking probability η required to achieve a specific step coverage for a given aspect ratio as calculated using Equation 4.1.

Equation 4.1 indicates that for a given aspect ratio AR=L/W the step coverage solely depends on the sticking coefficient η. To illustrate this behavior, Fig. 4.14 shows the η required to achieve a specific step coverage for a given aspect ratio. If, for instance, a trench with AR=50 is to be filled with 50% step coverage, the sticking coefficient should not exceed $5 \cdot 10^{-4}$. Note, however, that as the deposited film moves inwards the AR continuously increases, resulting in a lower step coverage later on in the process [90].

Reaction Chemistry

While the above arguments provide general insights on the properties of a suitable DT-CVD process, detailed studies are required to evaluate a specific chemistry. This section discusses some chemistries commonly used for planar WSi_x-CVD with focus on their applicability for high-AR DTs.

As follows from Sections 4.2.3 and 4.2.3, a WSi_x DT-CVD process has to simultaneously meet two requirements. On the one hand, the *total* sticking probability of the precursors has to be lower than approximately $5 \cdot 10^{-4}$ in order to enable sufficient step coverage. On the other hand, the *individual* concentrations and sticking probabilities of the precursors need to be such that the composition of WSi_x is $x \approx 2.3$ everywhere inside the DT. In the following several chemistries with respect to these requirements will be discussed.

According to theoretical results [52], single-phase WSi_2 can be deposited using volatile organo-metallic tungsten precursors like $[H_2W(\eta^5 - C_5H_5)_2]$ where $\eta^5 - C_5H_5$ represents cyclo-pentadienyl. However, to the authors knowledge no experimental confirmation of this prediction has been published, let alone a production-worthy process. Mixed-phased layers are generally deposited using WF_6 or tungsten carbonyl ($W(CO)_6$) [53, 44, 36]. This section will only discuss WF_6 since it is the most commonly used tungsten precursor in the semiconductor industry. As silicon containing precursors, silane (SiH_4), disilane (Si_2H_6), and dichlorosilane ($Si_2Cl_2H_2$, DCS) are all commonly used and are, therefore, considered in the following.

For the CVD process using WF_6 and silane or disilane, it has been shown that the DR crucially depends on the occurrence of a radical chain reaction in the gas phase [27]. With a process temperature above the so-called "extinction temperature" (T_{ex}) the radical chain reaction rapidly saturates, and the concentration of reactive gas phase precursors does not depend on T. If, on the other hand, the reactor temperature is below T_{ex} the chain reaction ceases and the deposition rate drops to zero due to lack of reactive precursors [90]. It has been shown, however, that even in the case the *surface* temperature is well below T_{ex} deposition can be achieved by heating the gas inlet (and thus initiating the radical chain reaction) [90]. Using this technique, CVD of WSi_x at temperatures as low as 313 K has already been demonstrated. It is important to note that under these conditions the step coverage is still determined by the surface temperature rather than by that of the inlet.

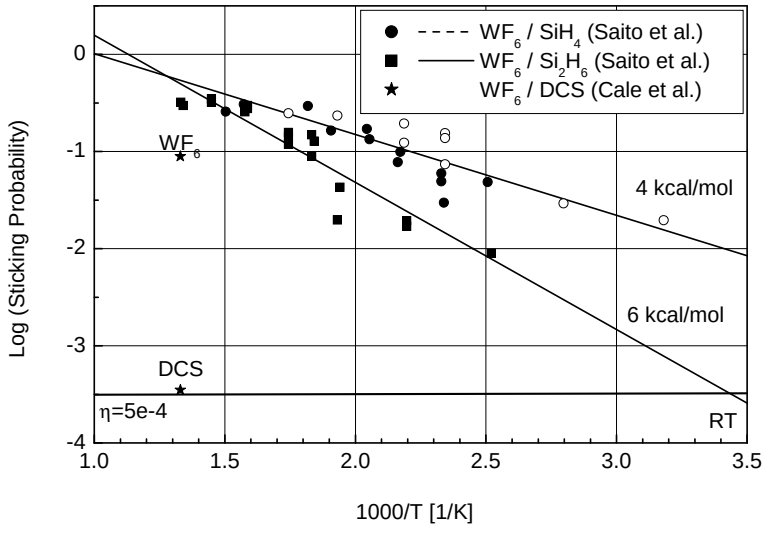

Fig. 4.15: Total sticking probabilities as a function of deposition temperature. Data is taken from [90] for the silane and the disilane processes and from [16] for the DCS process. The open and closed symbols represent depositions with and without preheating of precursors, respectively.

WF_6 and Silane

Recently, Shimogaki et al. suggested a model for the gas phase reaction between WF_6 and silane that is similar to the disilane reaction discussed below [100]. According to this model the radical chain reaction consists of the following gas phase reaction:

$$\begin{aligned} WF_6 + SiH_4 &\rightarrow WF_5(SiH_3) + HF \\ WF_{(6-y)}(SiH_3)_y + SiH_4 &\rightarrow WF_{(5-y)}(SiH_3)_{(y+1)} + HF \end{aligned} \tag{4.2}$$

These reactions describe the successive replacement of fluorine by SiH_3 at the WF_6-molecule. Assuming that deposition is mainly due to the resulting $WF_{6-z}(SiH_3)_z$ molecules [17], and that the above reactions have positive activation energy, this model correctly predicts an increasing Si content of the deposited layer for increasing temperature.

For the SiH_4/WF_6-system Saito et al. have determined the total sticking probability η as a function of temperature [90]. The observed behavior is plotted in Fig. 4.15. For this chemistry, η is higher than the required $5{\cdot}10^{-4}$ for all temperatures above room temperature. Additionally, the described chemical model predicts a composition close to x=1 for low temperatures. Both properties render this chemistry unsuitable for DT application which will, therefore, not be considered further in this article.

WF_6 and Disilane

While WF_6 decomposes at silicon surfaces [37] it has been shown experimentally that in the presence of sufficiently high concentrations of disilane (Si_2H_6) another reduction path dominates [27]. In this so-called "radical chain reaction", first proposed by Saito et al. [90], WF_6 is very efficiently reduced by an autocatalytic gas phase reaction. According to current understanding, the radical chain reaction is initiated immediately at the gas inlet by a thermal dissociation of WF_6 [27].

$$WF_6 \rightarrow \cdot WF_5 + \cdot F \tag{4.3}$$

Once this reaction has occurred, both resulting radicals can initiate a cascade of successive replacements of fluorine atoms by Si_2H_5-groups. In the case of the radical $\cdot F$ this is

$$\begin{aligned} \cdot F + Si_2H_6 &\rightarrow \cdot Si_2H_5 + HF \\ \cdot Si_2H_5 + WF_y(Si_2H_5)_{6-y} &\rightarrow WF_{y-1}(Si_2H_5)_{7-y} + \cdot F \qquad (y = 1...6) \end{aligned} \tag{4.4}$$

Similarly the $\cdot WF_5$ radical produced by the thermal dissociation (4.3) also can initiate a reduction chain reaction:

$$\begin{aligned} \cdot WF_5 + Si_2H_6 &\rightarrow WF_5(Si_2H_5) + \cdot H \\ \cdot H + WF_6 &\rightarrow \cdot WF_5 + HF \end{aligned} \tag{4.5}$$

Again, this pair of reactions results in the reduction of WF_6 by a Si_2H_5-group while the reactive $\cdot WF_5$ radical is restored.

The reactions (4.4) and (4.5) produce a significant amount of reactive $WF_y(Si_2H_5)_{6-y}$ precursors which can deposit at the surface. According to this model a higher Si/W ratio of the deposited layer is obtained by increasing the degree of reduction of WF_6. This can be achieved by heating either the substrate or the gas phase. Note, however, that in contrast to the WF_6/silane chemistry, here the film-forming species contains at least 2

silicon atoms per tungsten atom, even at low temperatures [27]. This favorable behavior is consistent with both, thermodynamic calculations [86] and experimental observations [90].

As has been shown by Kimbara, the minimum temperature (T_{ex}) to sustain the above radical chain reactions (and thus the deposition) depends on the surface-to-volume ratio (A/V) of the reactor [49]. Generally it is observed that increasing A/V also increases T_{ex}. According to [90] this dependence of the extinction temperature on reactor geometry is caused by a competing loss of reactive radicals to the reactor walls:

$$\begin{aligned} \cdot WF_5 + \cdot Si_2H_5 &\rightarrow WF_5(Si_2H_5) \\ 2 \cdot H &\rightarrow H_2 \\ 2 \cdot F &\rightarrow F_2 \\ \cdot H + \cdot F &\rightarrow HF \\ &\dots \end{aligned} \tag{4.6}$$

If the loss of radicals overcomes their production (reaction 4.3) the chain reaction ceases and the deposition rate drops to zero due to lack of reactive precursors [90].

Though the above gas phase chemistry model provides a thorough understanding of Si_2H_6/WF_6-based CVD, it has to be noted that no surface chemistry model exists for this process. Nevertheless, the experimentally obtained sticking coefficients plotted in Fig. 4.15 show that even at room temperature acceptable step coverage can only be achieved for moderate AR smaller than 50:1. Because of this limitation the author considers the Si_2H_6/WF_6 chemistry inappropriate for DT-CVD and will not discuss it further in this study. Note, however, that for medium AR trenches this process seems to be well suited, especially because of its favorable Si/W-ratio of the deposited layers.

WF_6 and Dichlorosilane

In contrast to the two systems described above, there exists no gas phase reaction model for CVD based on DCS and WF_6. For this chemistry it is apparently assumed that either there is no relevant gas phase reaction, or that the gas phase reactions do not vary noticeably in the process range explored so far. However, Cale et al. have developed a *surface* chemistry model which summarizes the present state of knowledge for this system [16]. It accounts for the deposition of the desired WSi_2(b) phase, but also for the deposition of Si(b) and a tungsten-rich, meta-stable W_5Si_3(b) bulk phase. This model, which its developers have implemented into their feature scale simulator *Evolve* [14], consists of the following three surface reactions:

$$\begin{aligned} SiH_2Cl_2 &\rightarrow Si(b) + 2HCl \\ WF_6 + 4SiH_2Cl_2 &\rightarrow WSi_2(b) + 8HCl + SiF_4 + SiF_2 \\ 5WF_6 + 11SiH_2Cl_2 &\rightarrow W_5Si_3(b) + 22HCl + 7SiF_4 + SiF_2 \end{aligned} \tag{4.7}$$

The deposition rate, R_j, of these reactions has been parameterized by:

$$R_j = k_j \exp\left(-\frac{E_j}{RT}\right)\left(\frac{p_D^{\beta_j} p_F^{\gamma_j}}{1 + K p_F}\right) \qquad j = Si(b), WSi_2(b), W_5Si_3(b)$$

where k_j is a prefactor, E_j the activation energy, p_F the WF_6 partial pressure and p_D the DCS partial pressure. β_j and γ_j are the respective reaction orders of DCS and WF_6. The rather unfamiliar factor on the right hand side has been introduced in [16] to account

for the experimental observation of Schmitz et al. that the deposition is inhibited by the reaction product SiF_4 [92]. Note that no inhibition has been observed for HCl which is produced in large quantities by the above reactions. The kinetic parameters for reactions (4.7) are summarized in Table 4.2 [15].

Table 4.2: Kinetic parameters for reactions (4.7) as given by Cale et al.

Phase	k_{oj}[mol/(cm^2 s mTorr$^{(\beta_j+\gamma_j)}$)]	E_j [kcal/mol]	β_j	γ_j	K [mTorr^{-1}]
Si	$1.3{\cdot}10^{19}$	90	2	0	0
WSi_2	$3.6{\cdot}10^{30}$	120	1	1	1000
W_5Si_3	$9.5{\cdot}10^{4}$	40	0.5	1	0

Because of the lack of literature data on the DCS/WF_6-chemistry at low temperatures it is quite difficult to decide whether this system is suitable for DT-CVD. Still, extrapolating the above model to low temperatures might already reveal some problems that occur at low temperatures.

Even without quantitative analysis it is evident that the relative activation energies of the above model favor the deposition of the tungsten-rich W_5Si_3(b) phase at low temperatures. This trend is consistent with experimental data which will be shown in section 4.2.3. Since a Si/W ratio of 0.6 is unacceptable for DT-CVD for stability reasons (see Section 4.2.3), it is apparent that this behavior is a major drawback of the DCS chemistry (especially compared to disilane). However, very high DCS and very low WF_6 partial pressures might suffice to counteract this effect.

A great advantage of the DCS/WF_6 chemistry as compared to silane and disilane is the much lower fluorine and chlorine bulk concentration, the low mechanical stress, and the better adhesion [108]. It has been observed that the F and Cl bulk concentrations increase with deposition temperature, a trend which is opposite to that in most other impurities [119]. However, both concentrations exhibit no measurable dependence of the individual flow rates.

A possible problem for a DCS/WF_6-based DT-CVD process is the experimental observation that the initial reduction of WF_6 is by the silicon-substrate rather than by DCS. This effect leads to a tungsten rich layer at the interface which is covered by other phases [37]. While these W-rich layers disappear at high deposition temperatures they might pose a problem at low temperatures.

Summarizing the DCS/WF_6-chemistry it can be stated that this system offers some advantages (low F/Si concentrations, low stress, good adhesion) but also includes some drawbacks (high W content at low T, W-rich interface). However, it is hard to decide about this chemistry since the model in [16] has only been calibrated to a narrow temperature range between 733 K and 813 K and experimental data for low T are sparse. The only available measurement of the *individual* sticking coefficients in Fig. 4.15 indicates that especially the value for WF_6 is far too high at 753 K. This implies that still much lower temperatures are required to fill a high AR feature with uniform stoichiometry.

Experimental

To fill the gap in the experimental data base and to determine the individual sticking coefficients of DCS and WF_6 as a function of temperature some first experiments were performed. The experimental setup and some preliminary results are presented in this section. All experiments presented here have been performed in an 8 inch Applied Materials *Centura* single wafer tool using WF_6 and DCS as precursors and Argon as carrier gas. In order to be consistent with earlier measurements [80] and with the above chemistry-

model the wafer temperature was assumed to be 50 K lower than the measured chuck temperature.

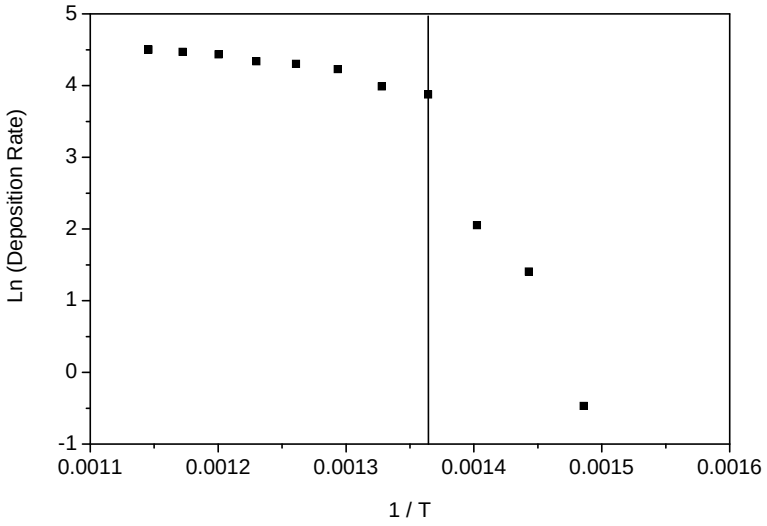

Fig. 4.16: Experimental deposition rate as a function of temperature for a DCS:WF_6 flow ratio of around 50:1 and 1.2 Torr total chamber pressure.

Fig. 4.16 shows an Arrhenius plot of the measured DR on blanket wafers for a DCS:WF_6 flow ratio of around 50:1 and a total chamber pressure of 1.2 Torr. The exponential decrease at low temperatures indicates that in this regime the DR is limited by thermally activated reactions. However, at temperatures above 750 K the deposition rate levels off to a maximum at 80 nm/min. In this regime, which is dependent on chamber design, total pressure, and the precursor flux, the DR is limited by the transport of reactive species to the surface. The transition between the two regimes is indicated by a vertical line.

In order to determine the step coverage in the respective regimes deposition experiments on structured wafers have also been performed. SEM images of the top of a DT are shown in Fig. 4.17. It is clearly visible that at 873 K only little deposition occurs inside the DT while at 693 K a smooth layer with good conformity is deposited.

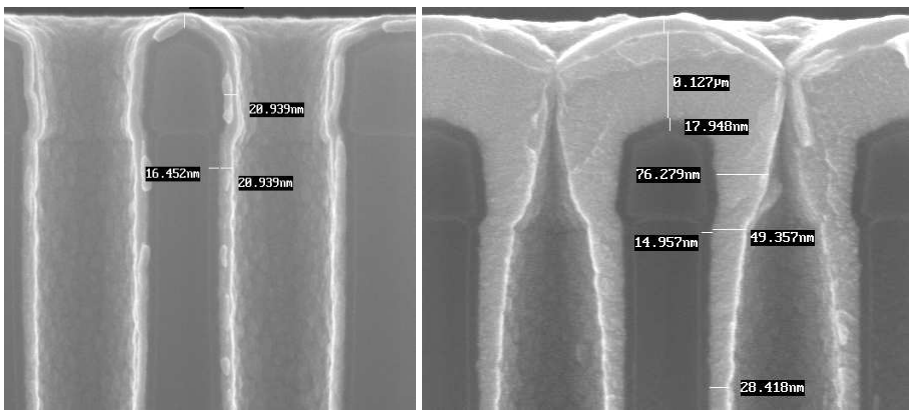

Fig. 4.17: Scanning electron microscope images of a WSi_x-layer deposited in DTs with an AR of 40:1. Layers were deposited at 693 K (left image) and at 873 K (right image).

While the above measurements yield the *total* sticking coefficient as a function of temperature (and thus the step coverage) they reveal nothing about the *individual* sticking probabilities (and thus the stoichiometry of the deposited layer). To access these additional experiments were performed.

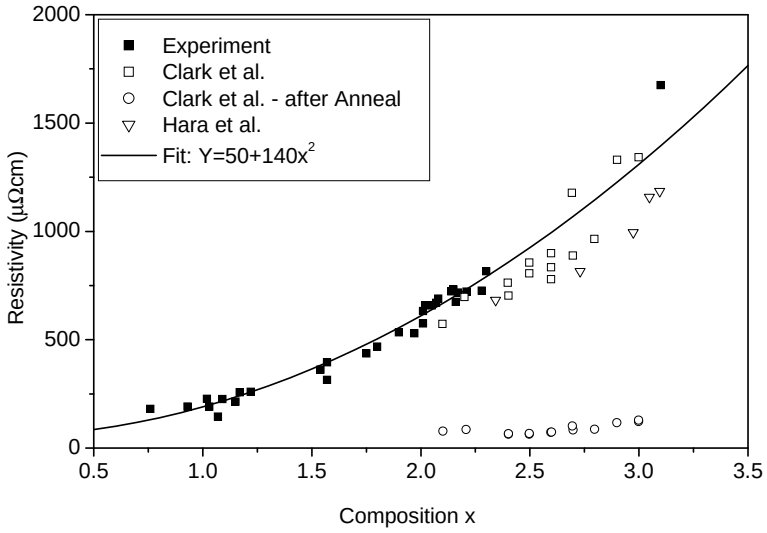

Fig. 4.18: Specific resistivity of tungsten silicide layers as a function of composition. Included in the graph are data from Clark et al. [37] and Hara et al. [24] as well as a quadratic fit to the data. Resistivity after an anneal shows only minor influence on composition.

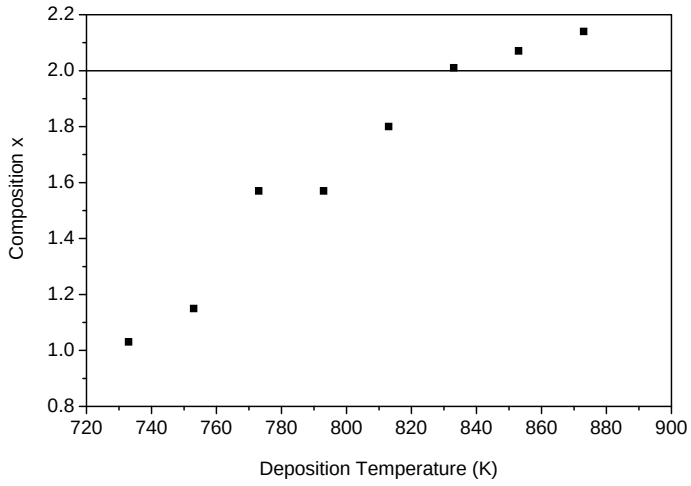

Fig. 4.19: Measured bulk composition as a function of deposition temperature for a DCS:WF_6 flow ratio of around 50:1 and a total chamber pressure of 1.2 Torr.

A common tool to measure bulk composition is Rutherford Backscattering (RBS). But, since this technique is both tedious and expensive the author applied the much faster and cheaper method of measuring the thickness using SEM and the film resistivity using a four-point probe. Proof that this simple method is sufficient to determine bulk stoichiometry of WSi_x-films is shown in Fig. 4.18 where data from [37] and [24] are plotted. As it is apparent, the specific resistivity of the as-deposited film depends quadratically on stoichiometry. This relation can thus be used to determine the Si/W ratio of as-deposited films from measurements of the specific resistivity. Note that this strong dependence is

lost after thermal annealing.

Using this technique, the bulk composition is measured as a function of deposition temperature. The obtained stoichiometries are presented in Fig. 4.19. According to the data, silicon-rich WSi_x is deposited at a temperature above 820 K while the silicon content steadily decreases to 1.0 at 723 K.

From the measured planar DR and stoichiometry the *individual* sticking probabilities of the precursors can be obtained from the kinetic gas theory as follows: given the surface temperature T_s and the partial pressure of a molecule p_j its deposition rate R_j is determined by the surface flux J_j and the sticking probability η_j:

$$R_j = \eta_j \cdot J_j \cdot \frac{M_d}{\rho_d} = \eta_j \cdot \frac{p_j}{\sqrt{2\pi M_j R T_s}} \cdot \frac{M_d}{\rho_d} \tag{4.8}$$

where M_j and M_d represent the precursor and bulk molar masses, and ρ_d the bulk density. Values for the species considered in this section are summarized in Table 4.3.

Table 4.3: Molar masses and bulk densities of species involved in the WF_6 reduction by DCS [16].

	Si	WSi_2	W_5Si_3	DCS	WF_6
M_j [kg/mol]				0.100	0.298
M_k [kg/mol]	0.028	0.240	1.004		
ρ_k [kg/m^3]	2200	9300	14550		

The sticking coefficients for DCS and WF_6 obtained from the above data and also from data by Raupp et al. [80] are shown in Fig. 4.20.

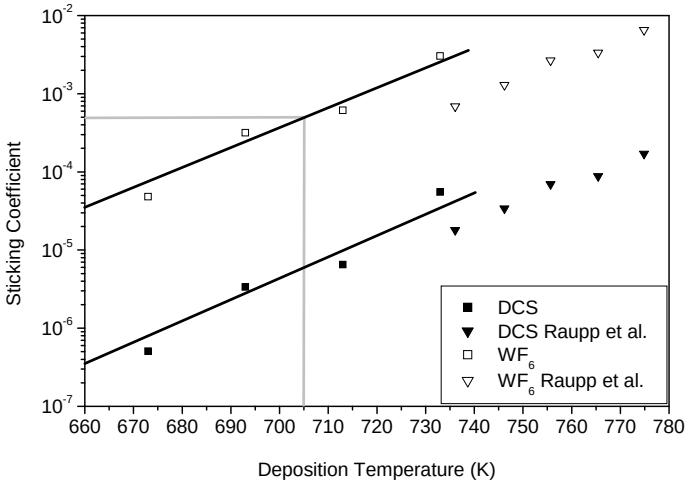

Fig. 4.20: Sticking coefficients (η) of DCS and WF_6 as a function of temperature, as determined from the deposition rate. Also shown are the η calculated from data by Raupp et al. [80].

Simulation

For additional insight into CVD using WF_6 and DCS, extensive simulations based on the surface reaction model by Cale et al. [16] were performed. The purpose of these studies is to identify process conditions under which both a good step coverage and a Si/W ratio of ≈2.3 are achieved along the total length of the deep trench.

Once calibrated, simulations are performed within minutes and give easy access to the step coverage and the stoichiometry inside the DT which is hard to measure experimentally.

However, it is important to remember that for DT-CVD the chemistry model in [16] needs to be pushed far beyond its validated parameter space. Thus the simulation results presented below should be considered with the appropriate care.

Two different types of simulations have been performed in this study, namely the deposition on planar wafers and into deep trenches. For the simulation of planar deposition it is assumed that the surface of the wafer is connected to a reservoir of constant precursor partial pressures via a diffusion channel. This channel is required to account for the transport limitation at elevated temperatures. For given process (reservoir) conditions detailed balance requires that the molar loss at the reactive surface (given by reactions 4.7, [16]) is equal to the diffusion flux through the transport channel which is governed by Fick's first law:

$$\begin{aligned} j_j &= -D_j \cdot \nabla n_j \\ D_j &= v_{\text{therm}} \cdot \lambda \end{aligned} \tag{4.9}$$

where ∇n_j is the concentration gradient of the molecules, $v_{\text{therm}} = \sqrt{2kT/m}$ the average thermal velocity and λ the mean free path of the precursors as described in section 4.2.3. Given the length of the depletion zone the precursor partial pressures at the wafer surface are varied until the loss through reactions equals the respective flux towards the surface. Here the collision cross section can be estimated whereas the depletion length has to be fitted to experimental data.

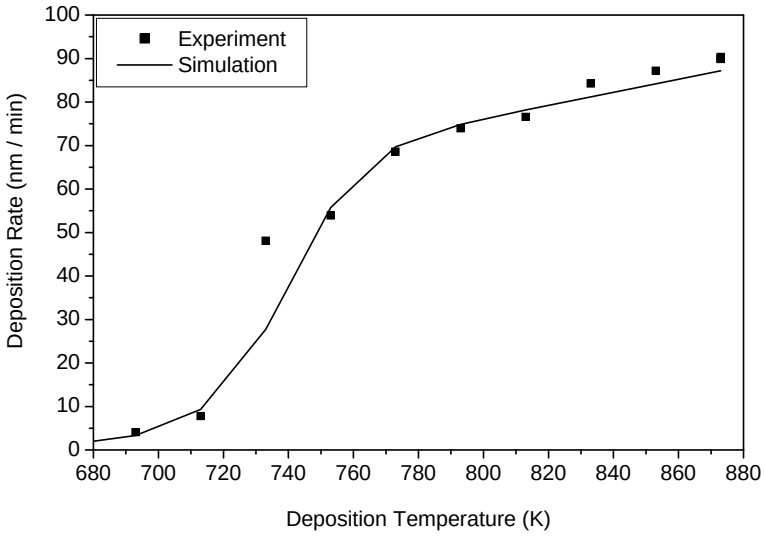

Fig. 4.21: Measured and simulated deposition rate as a function of temperature for a DCS:WF_6 flow ratio of around 50:1 and a total chamber pressure of 1.2 Torr.

Fig. 4.21 compares the simulated DR to the experimental data shown in Fig. 4.16. As can be seen the overall agreement with experiment is rather satisfactory. For the reaction-limited regime, which is important for DT-CVD, an Arrhenius plot of the deposition rate is shown in Fig. 4.22. Taking into account an offset of 50 K between measured chuck temperature and wafer temperature, this regime is also described reasonably well by the reaction rates given in [16]. But while the planar DR is simulated quite well, Fig. 4.23 shows that the simulated Si/W ratio as a function of temperature differs significantly from the experimental data. This indicates that the chemistry model in [16] is less accurate under the rather extreme process conditions studied here.

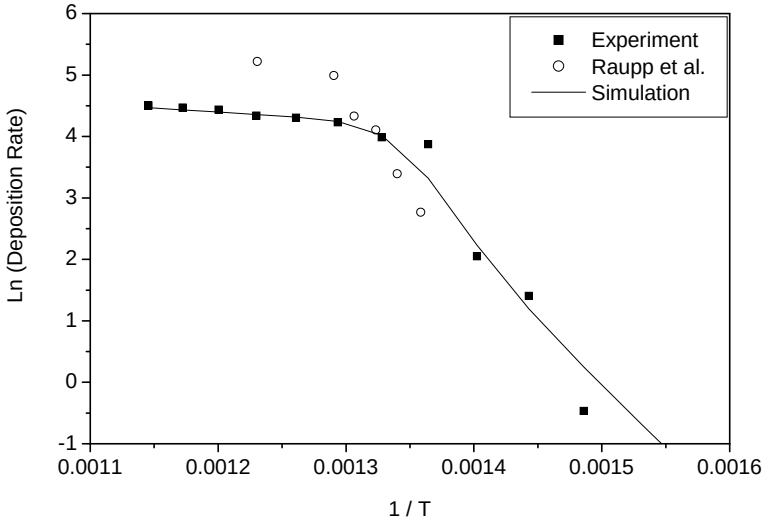

Fig. 4.22: Arrhenis plot of the reaction rate of Fig. 4.21.

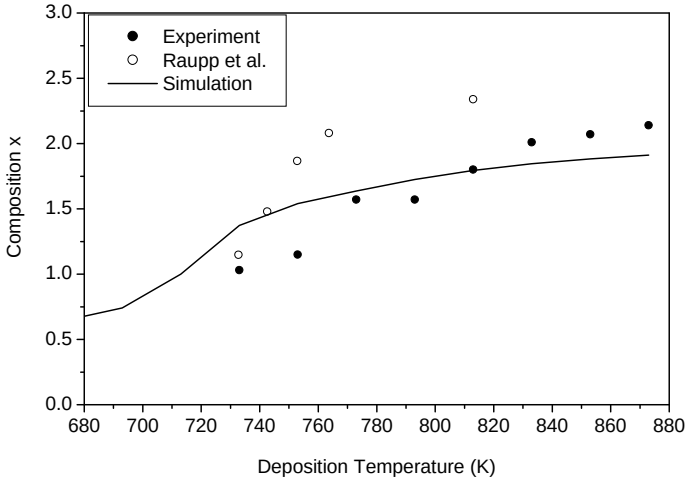

Fig. 4.23: Measured and simulated Si/W ratio x as a function of deposition temperature for a DCS:WF_6 flow ratio of around 50:1 and a total chamber pressure of 1.2 Torr.

Besides modelling planar deposition rates and Si/W ratio a simulator (*Knudsim*) has been developed which calculates the *local* DR and stoichiometry inside a circular DT. Here, the main problem (and the key to its solution) lies in the extreme AR of 50:1 or more. While simulators like *Evolve* give reliable results for deposition into comparably shallow features [16] they are generally unsuitable for structures with very high AR. This is due to the fact that an accurate calculation of the transport inside a DT requires either, massive CPU power, unrealistically much memory, or both. *Evolve*, for instance, solves the transport problem using the view-factor method which involves impractically large matrices for large AR. To avoid these problems the quasi-1D geometry of deep trenches has been exploited and it has been assumed that inside the DT the concentration is independent on radius. This is a reasonable approach since under typical CVD conditions today's trench diameters are orders of magnitude smaller than the mean free path in the gas.

Assuming this kind of 1D transport the diffusion inside the DT is also governed by Fick's first law (equation 4.9) but with the (Knudsen-)diffusion constant now given by:

$$D_j \approx v_{j,\text{therm}} \cdot d \tag{4.10}$$

where d is the diameter of the DT. As in the above case of planar deposition detailed balance requires that the molar loss due to surface reactions is equal to the diffusion flux. But while there the surface reactions occur *after* the transport through the depletion zone, inside the DT transport and deposition occur simultaneously.

In order to obtain the precursor concentrations θ_j of WF_6 and DCS the deep trench was divided into segments of length l and iteratively solve the equation for the steady state of $\theta_{j,i}$ at each discrete position i:

$$0 \equiv \frac{\partial \theta_j}{\partial t} = -\frac{D_j}{l}(2\theta_{j,i} - \theta_{j,i+1} - \theta_{j,i-1}) + \sum_{k=1}^{3} R_k(\theta_{1,2}) \tag{4.11}$$

where the reaction rates R_j are those from section 4.2.3. To solve equations 4.11 successive overrelaxation (SOR) was used until convergence is achieved. Given the precursor concentrations $\theta_{j,i}$ the *local* deposition rate (step coverage) and stoichiometry (x) are readily obtained from the reaction rates 4.7.

In order to study the impact of increasing AR during the process the above described *Knudsim* simulator was coupled to the feature scale simulator *Topsi*. In each timestep the *local* diameter d_i of the trench is obtained from the simulated front (*Topsi*). Using d_i the normal velocity for each front segment is calculated using *Knudsim*. The front is then propagated (*Topsi*) using the levelset algorithm [98] before the next timestep is computed.

Fig. 4.24 shows a comparison of the thus obtained structures with SEM images for three different temperatures and "standard" precursor concentrations. *"A"* shows a reaction limited deposition at 693 K, *"B"* a deposition at 713 K in the transition regime between reaction and diffusion limited and *"C"* a diffusion limited deposition at 783 K. As can be seen the step coverage and overall shapes of the evolved structures agree very well for all temperatures.

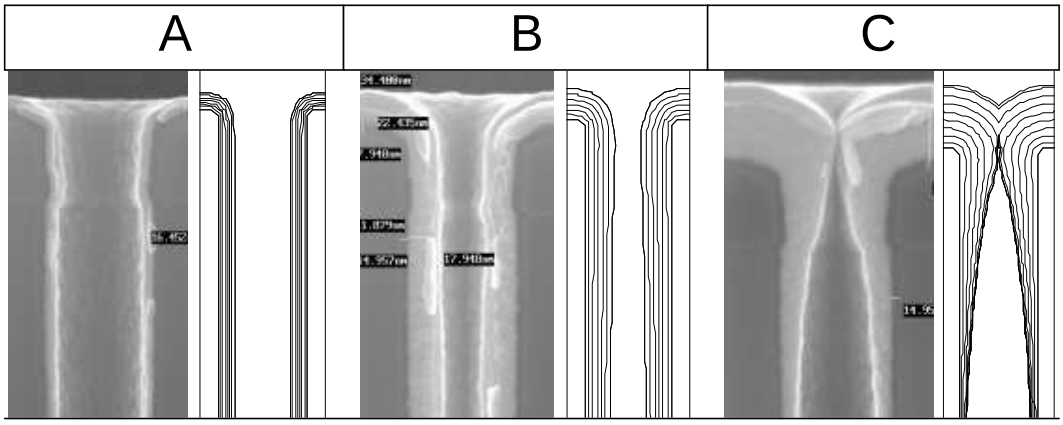

Fig. 4.24: Experimental and simulated structures for three different temperatures. *"A"* represents reaction limited CVD at 693 K, *"B"* the transition of reaction to diffusion limited deposition at 713 K, and *"C"* the deposition in the diffusion limited regime at 783 K.

In order to identify a process with both good step coverage and a uniform stoichiometry Si/W$\approx$ 2.3 an optimization of the free parameters p_{DCS}, p_{WF6}, and T with respect to

these properties has been performed. The simulations suggest such a regime below 650 K, a DCS partial pressure of 2-10 Torr and a WF_6 partial pressure of several mTorr.

Discussion

Already the general discussion of growth kinetics yielded that a DT-CVD process needs to be reaction limited in order to have a good step coverage inside high AR structures. Assuming an aspect ratio of 50:1 and a targeted 50% step coverage relation 4.1 (or Fig. 4.14) indicates that the total sticking coefficient η has to be lower than $5 \cdot 10^{-4}$. For the chemistry discussed here this in turn implies a deposition temperature well below 700 K as shown in Fig. 4.20.

While very low process temperatures might solve the problem of insufficient step coverage it has been shown experimentally that WF_6/DCS-based CVD produces tungsten-rich layers which are unsuitable for DT applications. However, the simulations presented in section 4.2.3 suggest that the higher deposition rate of W can be countered by (drastically) increasing the DCS partial pressure while decreasing the WF_6 concentration. While the proposed process conditions are far beyond the calibrated parameter space of the chemistry model in [16] the prospect of metallic DT plugs and vias has already triggered respective experiments which are currently performed.

Cale et al. have demonstrated that CVD using WF_6 and DCS is inherently transient in nature [16]. This means that the composition of the film changes during the deposition in a trench due to the change in AR. Compositional changes during the process are observed already with an AR of 5:1. When the opening of a structure shrinks the Knudsen diffusion of the reactants into the trench slows down and the deposition rate decreases. While this effect is correctly modelled by our combined *Knudsim+Topsi* simulator all above simulation results (with the exception of Fig. 4.24) were obtained without accounting for local and transient trench diameters. Thus the results presented here only represent the initial timestep in the process. If, however, the currently performed experiments turn out promising, further refinements of the parameters can be achieved by fully transient simulations.

For the extremely low sticking probabilities η required for DT-CVD the deposition rate necessarily is very low. For a first-order reaction and a given precursor pressure the deposition rate can be calculated using equation 4.8. For instance, deposition at partial pressure of 1 Torr and a sticking coefficient of $5 \cdot 10^{-6}$ does not exceed 10 nm/min. For economic reason a production worthy DT-CVD process therefore needs to be either performed in a batch reactor or in a single wafer tool at higher pressures.

Conclusion

Using experiments and simulations the reduction of WF_6 by DCS has been identified as a promising candidate for CVD of WSi_2-layers into deep trenches with aspect ratios of $\geq$50:1. The simulations presented here are based on a literature surface chemistry model and yield realistic step coverage over a wide range of temperature. Justified by these results a CVD process with both good step coverage is proposed that has a suitable stoichiometry inside the DT. While conclusive experimental confirmation is still missing our simulations suggest that stoichiometric films with a step coverage of more than 50% can be deposited below 700 K if the DCS partial pressure is 2-3 orders of magnitude higher than the WF_6 pressure.

4.2.4 Tungsten Deposition

Tungsten deposition into deep trenches is required when self-aligned tungsten-silicide bottom electrodes are to be fabricated. For this integration scheme pure tungsten is deposited into deep trenches having an oxide collar in the upper part. Tungsten silicide is formed during a subsequent anneal at those regions, where the tungsten is in contact with silicon. A wet etch removes pure tungsten that did not react selectively to tungsten silicide. A further anneal transforms the metal silicide into the final stable phase before the node dielectric is deposited. In the following, a tungsten deposition process is developed, that allows deposition of tungsten on patterned wafers without damaging the substrate.

Classical Tungsten CVD on Oxide

Usually, tungsten is deposited using the precursors WF_6 and H_2 while silane is required for initial seed generation. During SiH_4 and pulsed WF_6 supply a smooth tungsten-layer is deposited since the SiH_4-content increases after WF_6-supply leading to an increased seed generation [48]. In most cases tungsten-layers are poly-crystalline, but amorphous layers can be deposited by increasing the SiH_4 flow [18]. Above 400 °C, tungsten can be deposited directly on oxide yielding a relatively poor adhesion. The latter can be enhanced by the deposition of some sacrificial polysilicon which is consumed during further anneals while generating WSi_x [123]. Deposition of tungsten using WF_6 and hydrogen (H_2) leads to so-called *worm holes* in the silicon substrate mainly at Si/SiO_2 interfaces. At the end of these tunnels W-particles remain, which are neither consumed nor enlarged during the process [114]. Typically, TiN is used as a barrier to protect the substrate from the WF_6-attack [41]. If tungsten is deposited on oxide, already a 2 nm thick Si_3N_4 layer protects the silicon, but this is not sufficient to hinder F-diffusion into the SiO_2 [48].

Worm hole generation during reduction of WF_6 by H_2 is explained by dividing the deposition in three regimes [114]. In open regions, the silicon surface is exposed to the atmosphere, while half sealed and sealed regions are covered by an oxide. Gas molecules can only diffuse slowly to these areas, which in turn shifts the equilibrium of the chemical reactions:

- Open region A:

$$2WF_6 + 3Si(b) = 2W(b) + 3SiF_4 \quad (4.12)$$

$$WF_6 + 3H_2 = W(b) + 6HF \quad (4.13)$$

  If oxygen atoms are present at the surface, tungsten-fluorates which act as diffusion barrier for WF_6, are being generated: WF_5+(O)=WOF_3(b) + ... This leads to a self-limiting W-deposition.

- Half sealed region B: A higher concentration of subfluorides is expected at this point, which increases the deposition rate:

$$WF_n + 4n\ Si(b) = W(s) + 4n\ SiF_4 \quad (n=1...6) \quad (4.14)$$

$$WF_n + n\ H_2 = W(s) + n\ HF. \quad (4.15)$$

  In addition, oxifluoride as well as SiF_4, HF and WF_n are being generated. Since subfluorides are more reactive, the oxide is etched by WF_6, HF and subflurides.

- Sealed region C: A high concentration of HF leads to a fast etch of Si, during which worm holes are generated.

Silicon is always etched at sealed regions, so that an optimization of process parameters will not lead to a suitable process. Generation of HF has to be avoided, which can be done in two ways: either by using F-free precursors or by the elimination of hydrogen.

Tungsten CVD using Metalorganic Precursors

One group of F-free precursors are the metalorganic precursors like carbonyls. Kaplan and d'Heurle invented the tungsten deposition with tungsten-carbonyl [44]. As-deposited tungsten has a columnar structure, a grain size of 50 nm and a surface roughness of 5 nm. These properties remain unchanged after high temperature anneals [109]. Multiple deposition of W/Si has been achieved with hot filaments and $W(CO)_6$, but this method might not be suitable for good step coverage processes [36].

Recently, tungsten layers have been deposited successfully on very thin gate oxides [12]. Spontaneous decomposition of $W(CO)_6$ at the silicon-surface has not been observed, rather a critical density of metal seeds has to be formed at the surface before deposition will start. The content of carbon- (C) and oxygen- (O) impurities can be decreased in an UHV reactor by high deposition temperatures [53]. Metal-nitrides such as WN and TiN can also be deposited using metalorganic precursors, but often have a high specific resistivity [46, 41]. It is possible to deposit tungsten with other precursors like $(\eta^5{-}C_5H_5)_2WH_2$ and $(\eta^5{-}C_5H_5)W(CO)_3CH_3$, but deposition rates are usually much lower than with $W(CO)_6$ [102]. However, metalorganic precursors were not available during this study so no further discussion will follow.

Tungsten CVD without Hydrogen

If there is no hydrogen beside the WF_6, deposition is selflimiting in the region "A" of Section 4.2.4, since WOF_3 acts as a diffusion barrier for WF_6. Region "C" does not exist in this case. Due to thermodynamic calculations, WOF_3 should evaporate in UHV at temperatures above 300 °C, which would facilitate an unlimited deposition [114]. The pretreatment of the silicon surface determines the deposition kinetics. Fig. 4.25 shows SEM-images of tungsten deposition on two differently pretreated substrates.

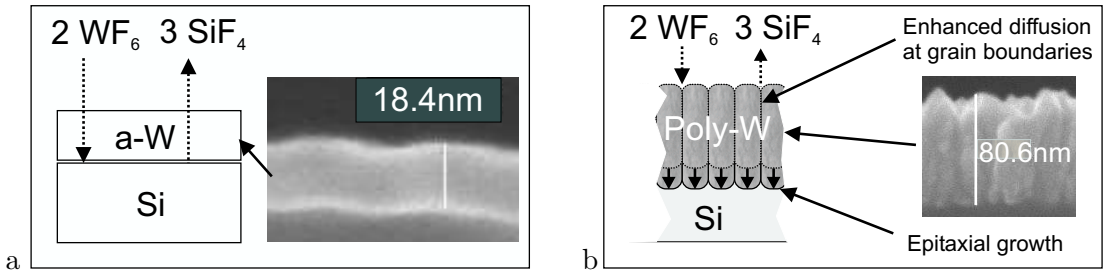

Fig. 4.25: Schematic and SEM-images of the tungsten deposition process on a hydrogen-passivated silicon surface (a) and on a chemically oxidized substrate (b).

Deposition on a hydrogen-passivated surface (a) leads to deposition of an amorphous layer, which acts as a diffusion barrier for WF_6 during subsequent deposition leading to a relatively thin layer. Panel a) also shows the schematic of the chemical reaction. More molecules are generated than consumed at the interface so that the process leads to bubble-formation and peeling. To reduce this trend, the process should proceed very slowly which can be reached by reducing the WF_6 partial pressure. Panel b) shows the same process on a Si-surface, which was prepared with an oxidizing wet-clean. During

this process, poly-crystalline tungsten which allows diffusion of molecules at the grain boundaries has been deposited. Therefore, the deposited layer is thicker than in Panel a). Smooth and thin layers, however, are required for the considered application, so that all further deposition has been performed on hydrogen-passivated surfaces.

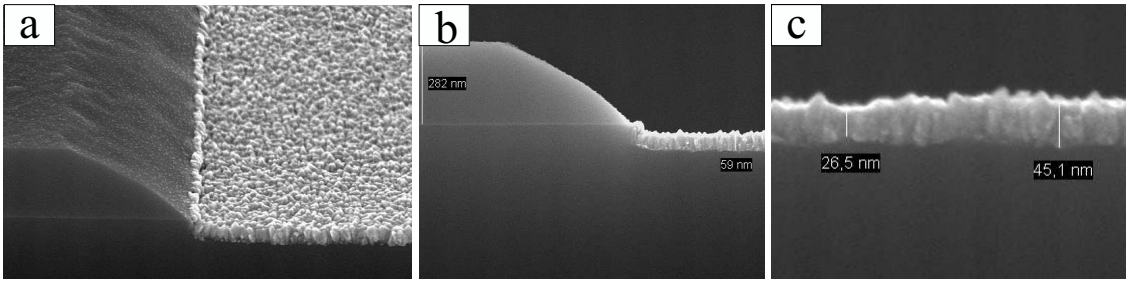

Fig. 4.26: SEM-image of a self-limiting tungsten deposition on prepatterned surfaces. The surface of the tungsten-layer is rough (a), but there is no enhanced substrate etch at the Si/SiO_2-interface (b). The layer thickness is still inhomogeneous and varies between 25 nm and 60 nm (c).

Fig. 4.26 shows selective deposition of tungsten on prepatterned substrates. A 300 nm thick oxide layer was deposited and openings defined by lithography were etched into it. Tungsten was deposited only in those openings, where silicon was exposed. The rough surface (Panel a) and the columnar structure (Panel b) indicate some oxide residuals before deposition. However, the selective deposition worked without enhanced substrate etch at the Si/SiO_2 interface (b), while the layer thickness varies between 25 nm and 60 nm (c). The usual deposition with SiH_4 instead of Ar would not work here since it is not selective [50].

4.2.5 Structuring of Tungsten Electrodes

Volatile compounds have to be generated during dry etch of materials in order to have a good material transport of the etched layer out of the system. During isotropic etching, volatile compounds are generated directly and can simply be pumped out of the system. Anisotropic etch processes are characterized by reaction products, which deposit on the wafer surface during the process and which have to be removed by ion bombardment. Very steep profiles can be reached by ions directed orthogonally to the surface. Volatile compounds are usually halogenides, while deposits consist of organic layers. At room temperature, WF_6 is the only volatile tungsten-halogenide, so that typically SF_6 is used as etch gas. Tungsten can be etched isotropically with an etch rate of 150 nm/minute and a selectivity of 30:1 to SiO_2 [42]. Since silicon is etched with the same chemistry, the etch of WSi_x using SF_6 has been tested in this work showing good results. With a SF_6/CHF_3-mixture, tungsten electrodes can be etched anisotropically without any undercut [45].

4.2.6 Polysilicon Deposition on Metals

To integrate metal-electrodes into a DRAM trench capacitor, polysilicon has to be deposited on metal. However, it is well known that metal-atoms can accelerate silicon deposition locally.

Fig. 4.27 shows an example of such effects for the polysilicon deposition on WSi_x. To inhibit the behavior demonstrated here, metal atoms at the surface of the WSi_x have to be passivated prior to polysilicon deposition. This has been reached by short nitridation or by room temperature oxidation for some hours.

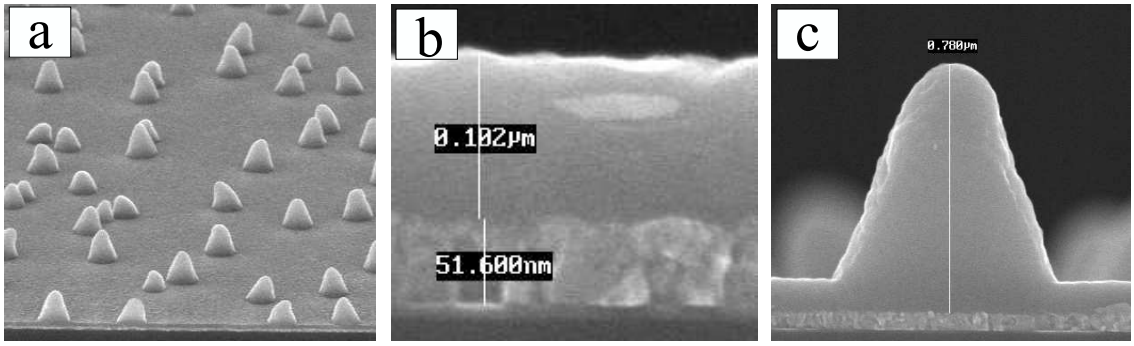

Fig. 4.27: SEM-images of abnormal polysilicon deposition on WSi_x. Deposition rate is enhanced only locally (a) while the process runs normally in large areas in between (b). Locally, layer thicknesses are increased by as much as a factor of seven (c).

4.3 Summary

In this chapter, single processes as well as process integration have been developed which facilitate the electrical characterization of metal electrodes described in Chapter 6. Main focus has been directed at the fabrication of metal gate capacitors and transistors on the one hand and on the processing of metal-fill deep trench capacitors on the other.

Chapter 5

Characterization of RTP Tunnel-Oxides

In this chapter, RTP-oxides with thicknesses scaled down into the tunnel regime are examined. In silicon process technology only furnace oxides are used as gate dielectric, since those have superior interface properties. RTP-oxides are usually employed as STI-passivation and as side wall oxide of the gate stack. For material characterization of electrodes, these oxides have the great advantage that splits in oxide thickness or gas composition during growth can be conducted much more easily during processing. However, the characteristics of these oxides are not well known. Therefore, interface properties as well as breakdown behavior are studied in this chapter before characteristics of the metal electrodes are investigated in Chapter 6.

5.1 CV-Measurements on Tunnel-Oxides

If the thickness of tunnel oxides is larger than 20 Å, it can be extracted from CV-curves using the method described in Chapter 2. First, measurements on fully integrated wafers have been analyzed and compared with results from planar capacitors. Fig. 5.1 shows CV-curves of diffusion-limited capacitors with different oxide thicknesses.

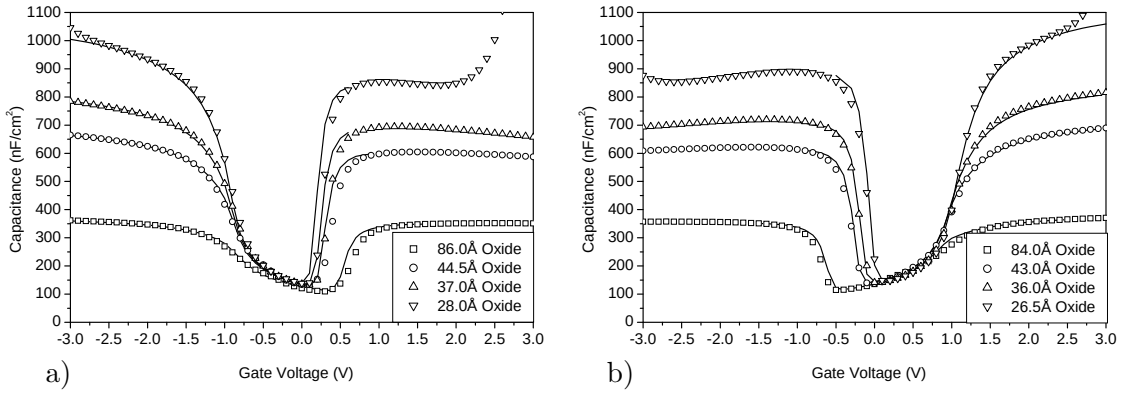

Fig. 5.1: CV-curves of fully integrated diffusion-limited capacitors with RTP-oxides of different thickness. Open symbols are measurement data, while straight lines represent simulated capacitances. Parameters extracted from CV-curves are summarized in the legend. Oxide thicknesses are slightly higher for NMOS-structures (a) as compared to PMOS-capacitors (b).

To obtain best agreements, simulated curves had to be shifted by 60-80 mV. This indicates a polysilicon work function which is 60-80 meV lower than that of crystalline

silicon. The slightly higher oxide thicknesses in NMOS-structures (Panel a) as compared to PMOS-capacitors (Panel b) are due to the faster oxide growth rates on boron-doped substrates. Fig. 5.2 shows CV-data of STI-limited capacitors of the same wafers.

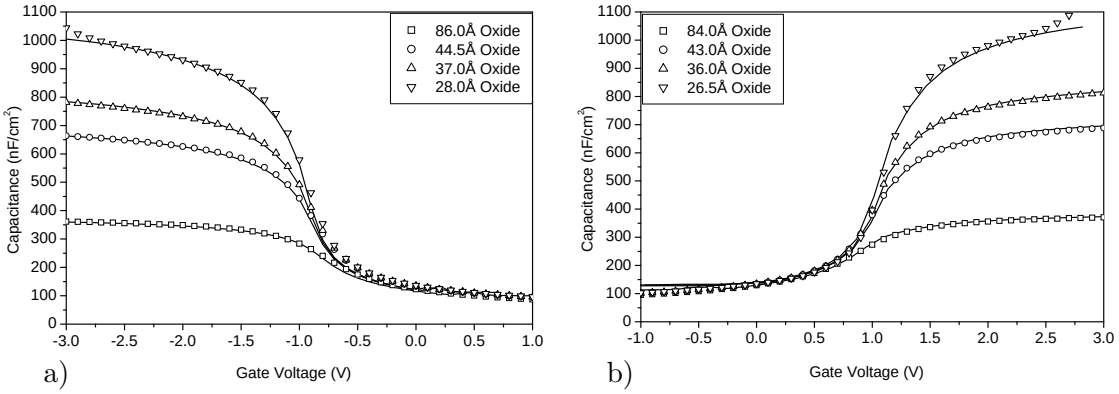

Fig. 5.2: CV-curves of fully integrated STI-limited capacitors with RTP-oxides of different thickness. Open symbols are measurement data, while straight lines represent simulated capacitances. Parameters extracted from CV-curves are summarized in the legend. Oxide thicknesses are slightly higher for NMOS-structures (a) as compared to PMOS-capacitors (b).

Substrate-doping can be extracted from CV-data in inversion, while the shape in accumulation allows for an estimate of the gate-doping. Flatband potential and physical oxide thickness have been extracted with the method described in Chapter 2 and are summarized in Table 5.1.

Parameter	NMOS (Diff)	PMOS (Diff)	NMOS (STI)	PMOS (STI)
Substrate Doping (cm^{-3})			$1.7 \cdot 10^{17}$	$2.4 \cdot 10^{17}$
Gate Doping (cm^{-3})	$2 \cdot 10^{19}$	$5 \cdot 10^{19}$	$2 \cdot 10^{19}$	$5 \cdot 10^{19}$
$V_{FB,1}$ (mV)	-1106	1199	-1029	1198
$V_{FB,2}$ (mV)	-1040	1153	-993	1164
$V_{FB,3}$ (mV)	-991	1163	-991	1164
$V_{FB,4}$ (mV)	-982	1187	-983	1188
$t_{ox,1}$ (Å)	86.9	84.6	87.1	84.7
$t_{ox,2}$ (Å)	45.7	43.7	45.8	43.7
$t_{ox,3}$ (Å)	38.3	36.3	38.5	36.5
$t_{ox,4}$ (Å)	28.8	25.8	28.9	25.9

Table 5.1: Data of 4 different RTP-oxides as extracted from CV-curves. Flatband potential V_{FB}, and physical oxide thickness, t_{ox}, of diffusion-limited (Diff) and STI-limited (STI) capacitors have been examined.

Apart from the thickest oxide of the NMOS-structure, both sets of data correlate well with each other, so that both structures are suitable for extraction of oxide thickness, t_{ox}, and flatband potential, V_{FB}. Automatically extracted oxide thicknesses are up to 1 Å thicker than those values giving the best fit. This difference results from the influence of the gate-doping which has been neglected in Chapter 2. By extrapolation to 0 nm thickness, interface charge density and work function difference are extracted from the flatband potential. The dependence of the flatband potential on the oxide thickness is shown in Fig. 5.3.

A single value of a furnace oxide is shown for comparison. An interface charge density

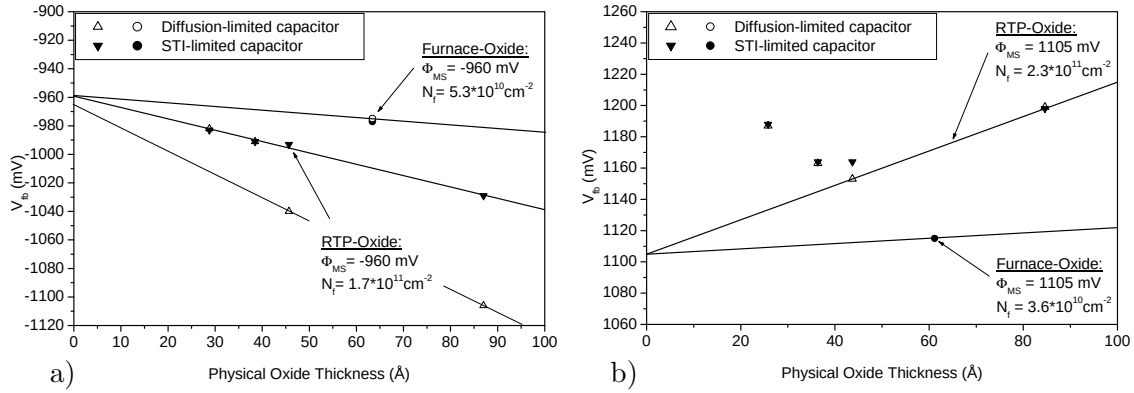

Fig. 5.3: Extracted flatband potential of RTP-oxides with different thickness of NMOS- (a) and PMOS- (b) capacitors. Closed symbols represent measurement data from STI-limited capacitors while open symbols show those of diffusion-limited structures. The straight line represents a linear fit. Given are extracted work function difference, ϕ_{MS}, and the interface charge density, N_{f}. A single value of a furnace oxide is shown for comparison.

extracted assuming an identical work function difference as for the RTP-oxide. The accuracy of this quantity, however, is limited since there was only one data point available. Comparison to the RTP-oxide shows an interface charge density which is lower by a factor of 5. The flatband potential of PMOS-structures as function of oxide thickness exhibits a deviation from the usual linear behavior for thin oxides. This deviation, however, is not understood so far.

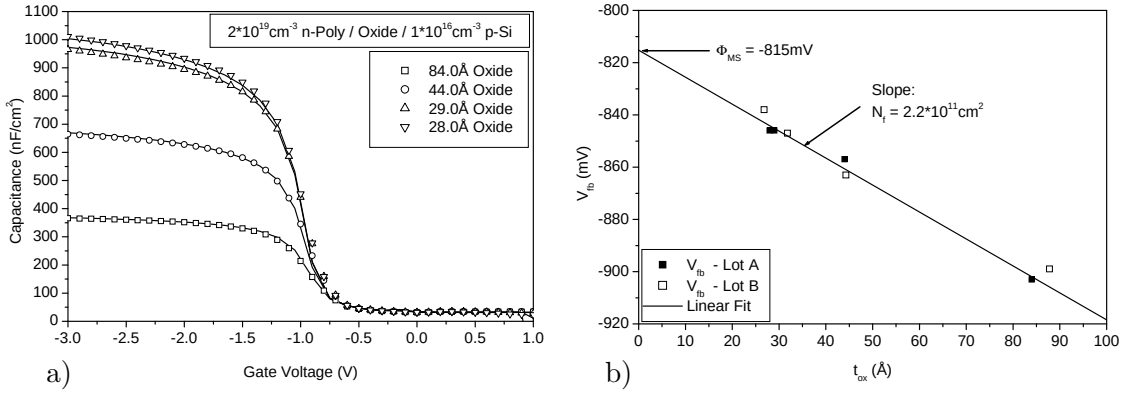

Fig. 5.4: CV-curves of simple planar capacitors with RTP-oxides of different thickness (a). Open symbols are measurement data while straight lines represent simulations. Panel (b) shows extracted values of the flatband potential for two lots. The straight line is a linear fit from which work function difference and interface charge density are extracted.

In the following, measurements on simple planar capacitors are analyzed to evaluate whether such samples are suitable for the extraction of parameters described so far. From a processing point of view, these structures are far easier and faster to fabricate, which shortens the learning cycle significantly. For even shorter processing times, only NMOS-capacitors were fabricated. Fig. 5.4 shows CV-data (a) and extracted values of the flatband potential (b) for two lots of planar capacitors with different oxide thicknesses. Straight lines in (Panel a) are simulated CV-curves of lot A. Both sets of data correlate well with each other and yield an interface charge density which is around 20% higher than that for integrated samples. The value of the work function difference is lower by 150 mV due to the reduced substrate-doping of $1 \cdot 10^{16}$ cm^{-3}. In summary, tunnel-oxides can well be characterized with the method shown in Chapter 2. Interface charge densities of

RTP-oxides are higher by a factor of 5 than those of furnace oxides, but are still sufficiently low to be used for the characterization of metal gate electrodes.

5.2 Characterization of Silicon Electrodes

Polysilicon electrodes are employed in almost all state-of-the-art circuits and are meanwhile well characterized. As described in Chapter 2, substrate doping can be extracted from CV-data or from the influence of the bulk potential on the threshold voltage of a transistor. The latter technique was applied to four different samples and the results are summarized in Fig. 5.5.

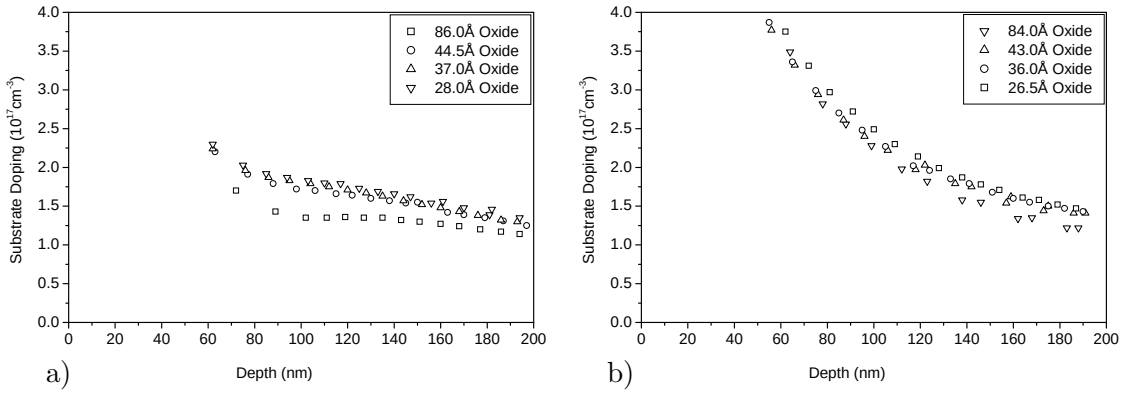

Fig. 5.5: Measured substrate doping of NMOS (a) and PMOS (b) transistors of four samples with different oxide thicknesses. Doping has been extracted from the influence of the substrate potential on the threshold voltage of a transistor.

Samples with thicker oxides show a slightly reduced substrate doping especially for the NFETs. This can be attributed to high temperatures during gate oxidation, whereby boron can diffuse out of the well. Together with extracted polysilicon-doping, band diagrams of the investigated NMOS- and PMOS-samples can be drawn as shown in Fig. 5.6.

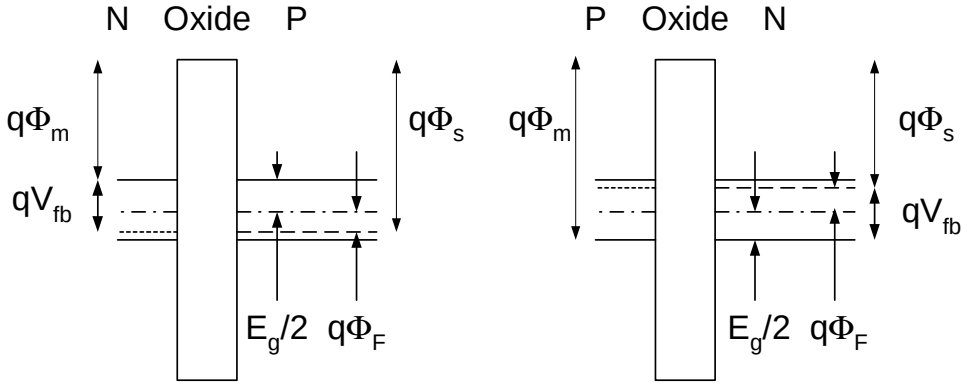

Fig. 5.6: Band diagram of a NMOS (left) and a PMOS (right) structure. Here, ϕ_{M} is the work function of the gate, ϕ_{S} the work function of the substrate, V_{fb} the flatband potential, E_g the band gap of the silicon and $q\phi_{\mathrm{F}}$ the Fermi-level.

Assuming the Fermi-level of the gate to coincide with the conduction band, the work function difference, ϕ_{MS}, can be expressed as follows [93]:

$$\phi_{\mathrm{MS,NMOS}} = -E_{\mathrm{g}}/2q - (kT/q)\ln(N_{\mathrm{A}}/n_{\mathrm{i}}) \quad (5.1)$$

$$\phi_{\mathrm{MS,PMOS}} = E_{\mathrm{g}}/2q + (kT/q)\ln(N_{\mathrm{D}}/n_{\mathrm{i}}) \quad (5.2)$$

Here, n_i is the intrinsic carrier density and N_A and N_D the substrate doping of NMOS and PMOS structures, respectively. Fig. 5.7 shows literature data of reference [116, 38, 43, 59], calculated values from equations 5.1 and 5.2 as well as measurement data from the previous section for polysilicon electrodes as a function of substrate-doping.

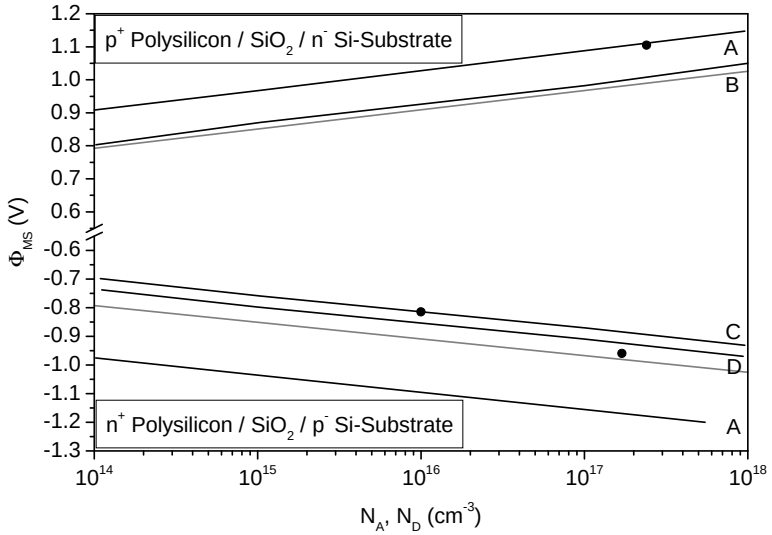

Fig. 5.7: Work function difference between polysilicon gate and silicon substrate as function of substrate doping. Shown are theoretical values of references [116] (A), [38] (B), [43] (C) and [59] (D), values calculated with equations 5.1 and 5.2 (grey), as well as values determined in this study (closed symbols).

NMOS data are described well by the dependence given in references [43] and [59], while PMOS data correlate well with values given in [116]. Values calculated with equations 5.1 and 5.2 might be shifted by 60-80 meV due to the difference in work function between crystalline and bulk silicon. These simple formulas would describe the data reasonably. In conclusion, CV-curves and substrate-threshold-voltage measurements give similar values for the substrate doping level. This fact was used to simulate the leakage current as described in the following section.

5.3 Leakage Current through Tunnel-Oxides

In the following, leakage currents through RTP-oxides are examined in order to determine the dominating tunneling mechanism and material parameters from the oxide. As shown in Chapter 2, only diffusion-limited capacitors are suitable to investigate leakage currents in inversion.

One aim of this section is to simulate leakage currents for thin and thick dielectrics with one set of parameters using doping densities and flatband potential extracted from CV-measurements. Silicon-oxides are close to ideal dielectrics, so that the tunneling current can be described accurately by quantum mechanical tunneling through a potential barrier. For high gate voltages, carriers have to tunnel through a triangular barrier which is well described by the Fowler-Nordheim expression stated in Chapter 2. On the other hand, the trapezoidal potential at low voltages can be dealt with using a method proposed by Ando et al. [2]. Considered are only motions perpendicular to the barrier so that the whole problem can be reduced to a one-dimensional case.

Here, E is the electron energy and U_j the potential at the j-th element at the position

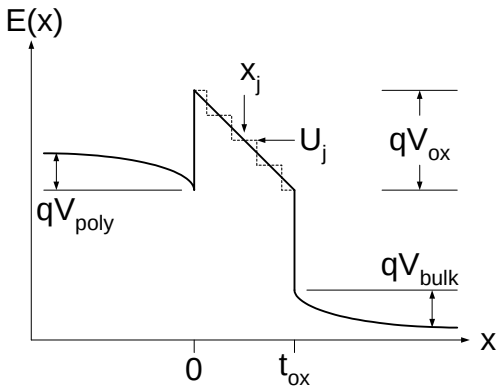

Fig. 5.8: Schematic of the conduction band as assumed for simulation. An arbitrary shape (solid line) is approximated by a staircase potential (dotted line). V_{Poly}, V_{ox} and V_{bulk} are the voltage drops across gate, oxide and substrate, respectively.

x_j. At each element, the wave function can be written generally as:

$$\Psi_j(x) = A_j \exp(ik_j x) + B_j \exp(-ik_j x) \tag{5.3}$$

where

$$k_j = \sqrt{2m_j^*(E - U_j)}/\hbar \tag{5.4}$$

and $\hbar$ is the reduced Planck's constant. The mass m_j^*, however, is the effective mass of the state the electron finally tunnels into and not that one at the j-th position inside the barrier. Assuming an electron travelling from left to right, the final wave function has the coefficients $B_{N+1} = 0$ whereas A_0 has been set to 1. The wave function at the l-th element can be calculated from the initial one by:

$$\begin{pmatrix} A_l \\ B_l \end{pmatrix} = \prod_{l=0}^{j-1} M_l \begin{pmatrix} A_0 \\ B_0 \end{pmatrix} \tag{5.5}$$

where

$$M_l = \frac{1}{2} \begin{bmatrix} (1+S_l)\exp[-i(k_{l+1}-k_l)x_l] & (1-S_l)\exp[-i(k_{l+1}+k_l)x_l] \\ (1-S_l)\exp[i(k_{l+1}-k_l)x_l] & (1+S_l)\exp[i(k_{l+1}+k_l)x_l] \end{bmatrix} \tag{5.6}$$

and

$$S_l = \frac{m_{l+1}^*}{m_l^*}\frac{k_l}{k_{l+1}} \tag{5.7}$$

Now the transmission coefficient, $D(E)$, is calculated as follows:

$$D(E) = \frac{m_0^*}{m_{N+1}^*}\frac{k_{N+1}}{k_0}|A_{N+1}|^2 \tag{5.8}$$

where

$$A_{N+1} = \frac{m_{N+1}^*}{m_0^*}\frac{k_0}{k_{N+1}}\frac{1}{M_{22}} \tag{5.9}$$

and

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \prod_{l=0}^{N} M_l \tag{5.10}$$

Assuming a transmission probability that depends only on the longitudinal electron energy, the current density is given as:

$$J(E_{\mathrm{x}})\,dE_{\mathrm{x}} = \frac{qm_0^*}{2\pi^2\hbar^3} D(E_{\mathrm{x}}, V_{\mathrm{ox}}) \times \int_{E_{\mathrm{x}}}^{\infty} [f_0(E) - f_{N+1}(E)]\,dE\,dE_{\mathrm{x}} \tag{5.11}$$

The basic Matlab-routine has been taken from the PhD-thesis of Thomas Pompl [77]. However, major modifications were necessary to apply this procedure to thin oxides at small voltages. As it was, only oxides thicker than 3 nm and gate voltages above 3 V were described well by the simulation. With the modifications developed here, tunneling currents for all voltages and all oxide thicknesses investigated could be simulated. In the following, approaches found in literature are discussed to understand the leakage currents at low voltages. The barrier height ϕ is the potential barrier between bulk silicon and SiO_2 has been determined to be 3.1 eV [7, 115], whereas for polysilicon somewhat smaller values were estimated. High applied biases lead to quantization of electronic states in the conduction band of the cathode. Assuming a triangular potential, the base level is given by Airy functions [115]:

$$E_0 = \frac{Z_0}{2^{1/3}} \left(\frac{q\hbar}{m_{\mathrm{Si}}^{1/2}} E_{\mathrm{Si}} \right)^{2/3} \tag{5.12}$$

where $Z_0 = 2.34$ is the first zero of the Airy function, m_{Si} the effective mass in silicon normal to the interface, and E_{Si} is the electric field inside the silicon. The latter does not have an exact triangular shape, so that equation 5.12 will overestimate the quantization effect. A high density of electrons will broaden the well significantly so that, as a first approximation, this effect was accounted for by only 50%. Quantum-mechanical simulations from Lo et al. [63, 62] showed that only the first subband contributes significantly to the tunnel current. In their work, leakage currents between 1 V and 3 V were well described, but for high and low voltages larger differences between simulation and measured data occurred. In the following, modifications to the simulations will be discussed which enable calculation of the current for all voltages.

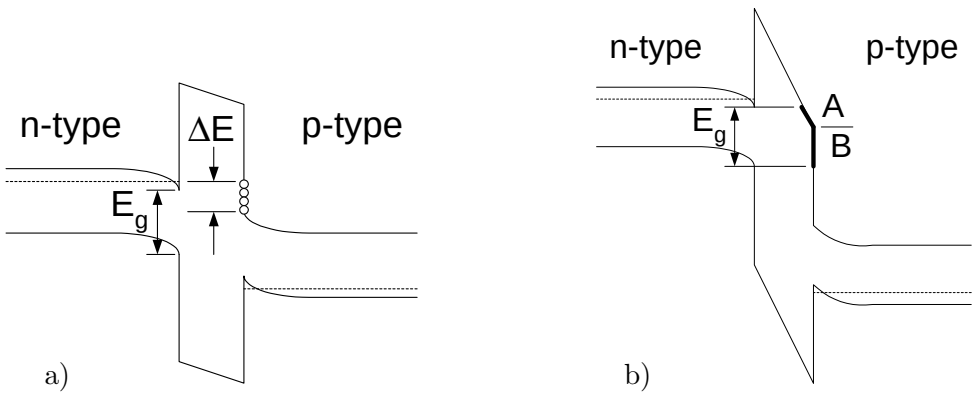

Fig. 5.9: a) Band diagram of a NMOS-structure in accumulation. It is assumed that only charge carriers from the energy band ΔE contribute to the tunnel current. Panel b) shows the band diagram of a NMOS-structure in strong accumulation. Electron tunnel from the cathode conduction band into the oxide conduction band (region A) and into the Si/SiO_2 interface states (region B).

To calculate the current, all energy levels within the band gap are considered at high voltages. For small voltages, the tunneling current in the opposite direction has to be calculated as well. Alternatively, only those energy levels within the band ΔE can be included, which energetically lie above the conduction band of the anode.

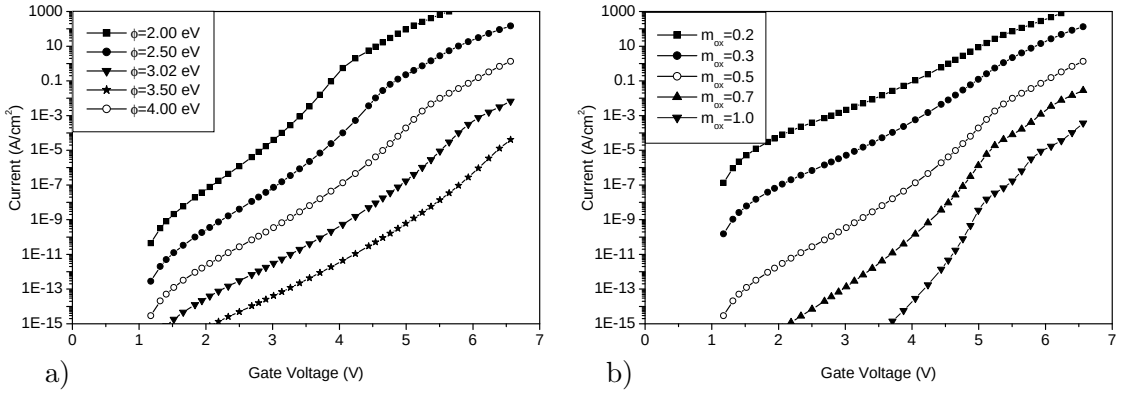

Fig. 5.10: Simulation of the tunnel current through a 4 nm silicon-oxide as a function of band offset, Φ, (Panel a) and the effective electron mass of the SiO_2 conduction band, m_{ox} (Panel b).

A drastic reduction in leakage current following the increase of barrier height and effective mass is observed when a 4 nm oxide is used (Fig. 5.10). In addition to the dielectric constant, these two quantities are crucial during the development of new dielectrics. Between 3 and 6 V there are oscillations which are not described by the Fowler-Nordheim formula. Electrons tunneling into the conduction band of the oxide are accelerated towards the anode and can be reflected at the transition to the electrode. Reflected and tunneling electrons can interfere and thus cause oscillations in the leakage current. A reduced effective mass in the anode leads to higher oscillations while the effective mass in the cathode influences the current only slightly. If electrons tunnel directly from one electrode to the other as is the case for oxide voltages smaller than 3 V, the process is referred to as direct tunneling.

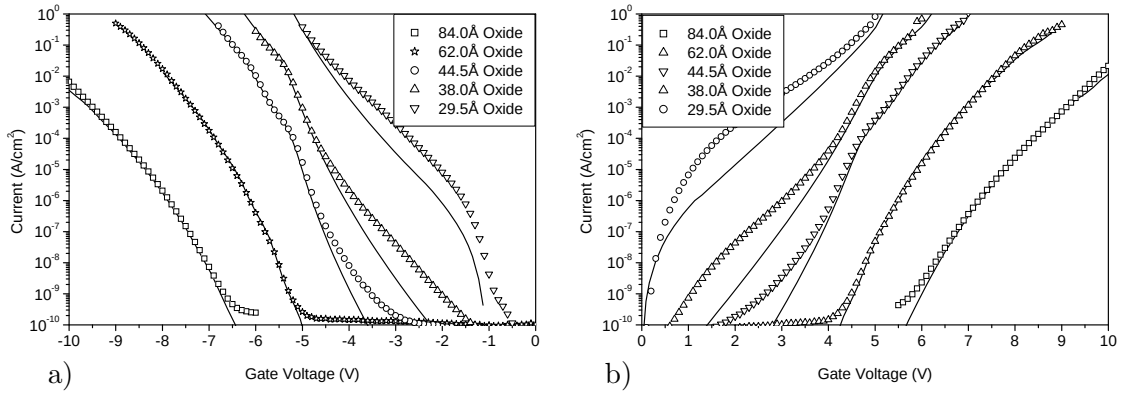

Fig. 5.11: Leakage currents through diffusion-limited NMOS-capacitors in accumulation (a) and inversion (b). Open symbols are measurement data, lines represent simulations. The legend shows oxide thicknesses extracted from simulations at voltages higher than 4 V.

Leakage currents through diffusion-limited NMOS-capacitors can be simulated well for voltages higher than 4 V (Fig. 5.11). Similarly, simulations for diffusion-limited PMOS-capacitors agree well with measurement data at voltages higher than 4.5 V (Fig. 5.12). Table 5.2 summarizes parameters used for simulation. Flatband potentials, $V_{FB,CV}$, have been extracted from CV-curves in the previous section. Oscillations were described most accurately when the effective mass of the polysilicon was assumed to be 1.3 m_0. This value is somewhat higher than the usual value of 1.08 m_0. The best description was reached by assuming a conduction band offset of 3.15 eV. For simplicity, a constant quantization energy of 150 meV had been assumed in accumulation, while variable quantization as described above had been used for NMOS in inversion. Interface state injection for PMOS

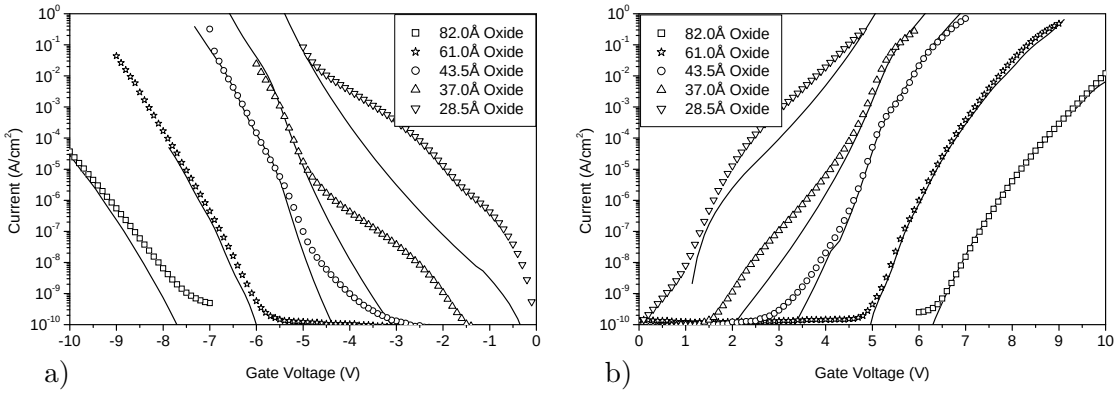

Fig. 5.12: Leakage currents through diffusion limited PMOS-capacitors in inversion (a) and accumulation (b). Open symbols are measurement data, while lines represent simulations. The legend shows data extracted from simulations.

in inversion does not lead to quantization.

	NMOS/ACC	NMOS/INV	PMOS/INV	PMOS/ACC
Φ_M polysilicon/SiO_2 (eV)	3.15		3.15	
Φ_S bulk-Si/SiO_2 (eV)		3.15		3.15
Gate Doping (cm^{-3})		$4.5 \cdot 10^{19}$		$7 \cdot 10^{19}$
Substrate Doping (cm^{-3})	$2 \cdot 10^{17}$	$2 \cdot 10^{17}$	$3 \cdot 10^{17}$	$3 \cdot 10^{17}$
Flatband Potential (V)	$V_{FB,CV}$	$V_{FB,CV}$	$V_{FB,CV}$	$V_{FB,CV}$
m_e within Substrate (m_0)	1.08	1.08	1.08	1.08
m_e within Gate (m_0)	1.30	1.30	1.30	1.30
m_e within SiO_2 (m_0)	0.50	0.50	0.50	0.50

Table 5.2: Parameters used for simulation of leakage currents through NMOS- and PMOS-structures in accumulation (ACC) and inversion (INV).

For small gate voltages, measured leakage currents are substantially higher than simulated ones. In former investigations this differences have been explained by oxide charges [77]. In the following, however, other approaches found in literature will be examined in order to describe the low-voltage case. Beside the Fowler-Nordheim formula for oxides thicker than 6 nm, there are accurate semi-physical descriptions for small voltages [57]. A uniform description for all voltages on the other hand can not be found. The parameter with the highest uncertainty is the effective electron mass in the oxide. At high electric fields electrons tunnel into the conduction band of the oxide which can be described very well by a parabolic effective mass of 0.5 m_0 as shown by Weinberg [115]. In this context parabolic means that a parabolic dispersion relation had been assumed. At low voltages electrons do not tunnel into the conduction band of the oxide but into the Si/SiO_2 interface. In recently published literature different effective oxide masses have been assumed to describe leakage currents at low bias. However, a constant effective mass has been assumed in all cases. Since the effective mass should be 0.5 m_0 at the conduction band of the oxide and no non-continuencies are observed, a continuously decreasing effective mass at the interface has been assumed here. Good agreement between measurement and simulation was obtained when assuming the effective mass shown in Fig. 5.13.

The effective mass is assumed to decrease from 0.5 m_0 at the oxide conduction band to

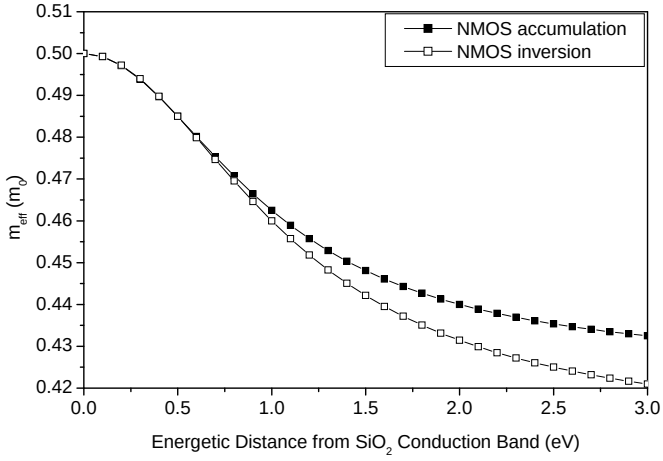

Fig. 5.13: Assumed variation in electron effective mass at the Si/SiO_2 interface as function of distance from the oxide conduction band. Slightly lower values were assumed for NMOS in inversion.

a lower value at the silicon conduction band. In case (A) shown in Fig. 5.9b for example, all tunneling matrices were calculated with 0.5 m_0 while all matrices in case (B) were calculated with a reduced effective mass. Slightly lower values were assumed for NMOS in inversion than for the other three conditions. These values do not agree with quantities expected inside the silicon conduction band which are around $1 * m_0$ or higher. This suggests that electrons tunnel either into interface-states or into bulk silicon where only one or two of the 6 valleys are available for electron tunneling.

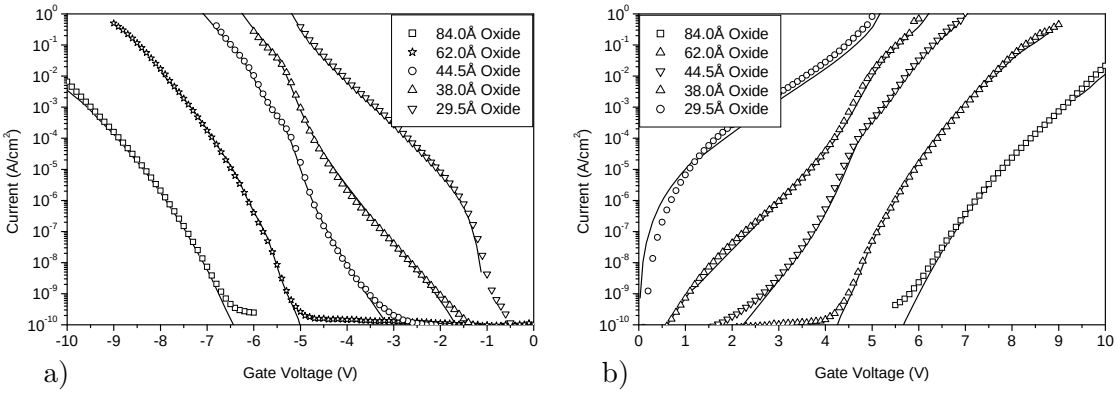

Fig. 5.14: Measured (open symbols) and simulated (solid lines) leakage current for NMOS-structures with RTP-oxides of different thicknesses in accumulation (a) and inversion (b). A variable effective electron mass at the Si/SiO_2 interface has been assumed.

A physical description of the effective mass at the interface is very difficult since electrons tunnel from a three-dimensional state of the cathode into a two-dimensional state of the anode. The effective mass given in Fig. 5.13 should therefore be considered as a means to describe the tunneling probability rather than be taken as a real physical quantity. Fig. 5.14 demonstrates good agreement between measurement and simulation which is reached when assuming a variable effective electron mass.

Using the method described above, all cases except PMOS in inversion can be described. In the latter case, no inversion carriers are supplied to the conduction band of the cathode and electrons have to tunnel from the valence band or interface states. Lee et al. have analyzed different tunneling mechanisms and concluded that the hole current

from the anode is dominant for low gate voltages [58].

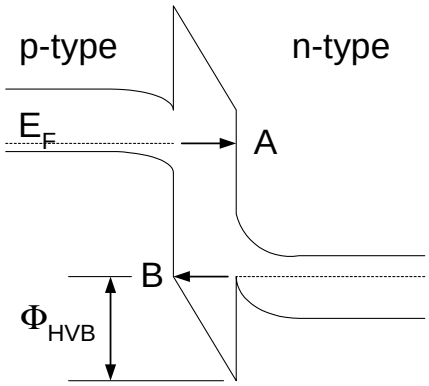

Fig. 5.15: Band diagram of a PMOS-structure in inversion. Schematically shown are the electron current from the conduction band of the cathode (A) and the hole current from the anode valence band (B).

The process is sketched in Fig. 5.15. A valence band offset of $\phi_{\mathrm{HVB}} = 4.5$ eV and an effective electron mass of 0.37 m_0 were assumed to simulate this effect [57]. Curves with (black solid line) and without (grey dashed line) including the hole tunneling are presented in Fig. 5.16.

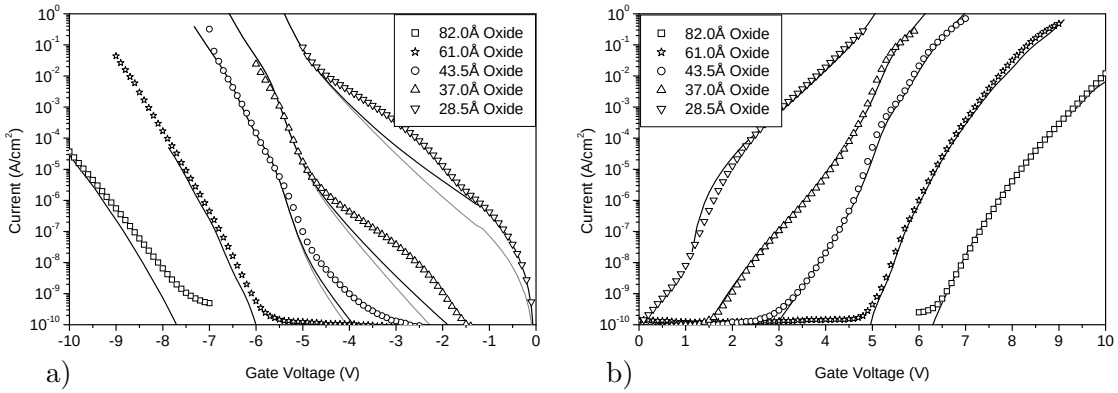

Fig. 5.16: Measured (open symbols) and simulated (solid lines) leakage current for PMOS-structures with RTP-oxides of different thickness in inversion (a) and accumulation (b). A variable effective electron mass at the Si/SiO_2-interface has been assumed. For inversion, curves with (black solid lines) and without (grey dashed lines) inclusion of the hole current are presented.

Measurements can be well described for low and high voltages. The cause of the increased leakage current between 2 and 3V, however, is not yet understood. In summary, for the first time leakage currents for a wide range of oxide thicknesses can be described for all gate voltages with a consistent set of parameters.

Finally, oxide thicknesses extracted by different techniques are compared (see Fig. 5.17). The straight line represents the ideal case, from which experimental data deviates only up to 1-2 Å. Ellipsometric measurements bring about a larger error since those measurements could not be done on the same wafers.

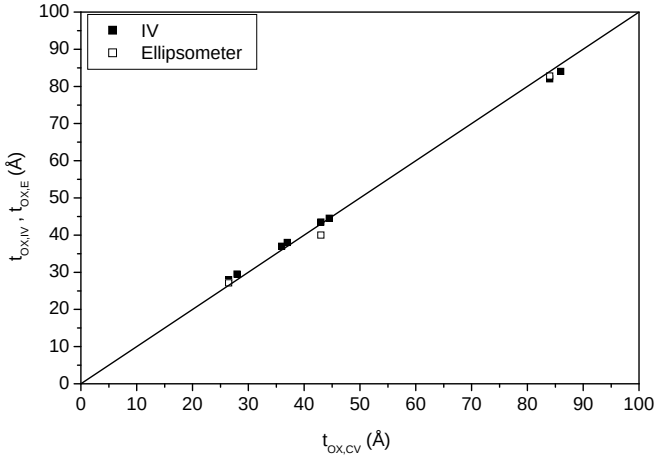

Fig. 5.17: Oxide thicknesses extracted from IV-data and ellipsometry as a function of oxide thicknesses as determined by CV-curves.

5.4 Interface Properties of RTP-Tunnel-Oxides

Interface traps of RTP-oxides were characterized with the charge-pumping technique described in Chapter 2. The results are summarized in Fig. 5.18a for NMOS and PMOS with oxide thicknesses between 25 Å and 85 Å. Interface trap densities shown in Panel a) are significantly higher for NMOS than for PMOS for all oxide thicknesses under study and are in the range of $2 - 5 \cdot 10^{10}$ cm^{-2} eV^{-1}. This trend is contrary to that usually observed in furnace oxides, where the NFET normally exhibits a lower trap density. An example of a 5.8 nm furnace oxide is shown for comparison.

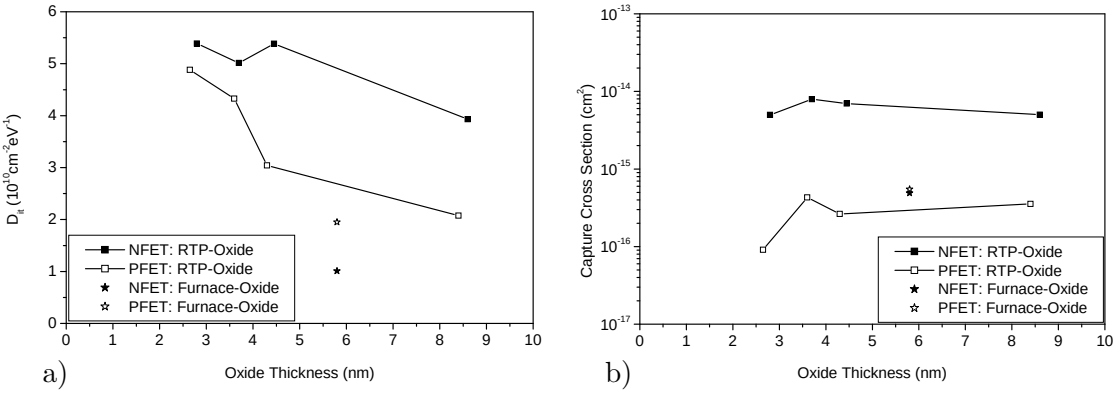

Fig. 5.18: Interface trap density (a) and capture cross section (b) of NMOS (closed symbols) and PMOS (open symbols) structures with RTP gate oxides.

Interface traps in RTP-oxides are a factor of 2-3 higher than those in furnace oxides and are generally lower for thicker oxides as compared to thinner ones. A second quantity that can be extracted from CP-measurements is the capture cross section of electrons and holes as shown in Fig. 5.18b. There is no significant dependence on the oxide thickness, and the values gained for PFETs are similar to those for furnace oxides lying in the range of 10^{-16} cm^2. NFETs with RTP-oxides, on the other hand, show a capture cross section one to two orders of magnitude higher than the PFETs. This might have a significant effect on the 1/f-noise and on the soft breakdown (SBD) of MOS-devices. A brief analysis of RTS after SBD in NFETs will be presented in the remainder of this section. The

comparison between NFETs and PFETs, however, will be left open for future studies. With continuous scaling of transistor size, the signal-to-noise ratio may degrade because of the decrease in signal amplitude. In submicron MOSFETs discrete switching events between two or more levels in the current at constant bias conditions are regularly observed. The fluctuations are known as random telegraph signals (RTS) and are attributed to the trapping and detrapping events of individual interface states at the silicon/silicon-oxide interface. Usually, RTS-signals increase significantly after a soft breakdown suggesting that RTS and SBD are correlated and that RTS might be a sensitive indicator for the quality of an oxide.

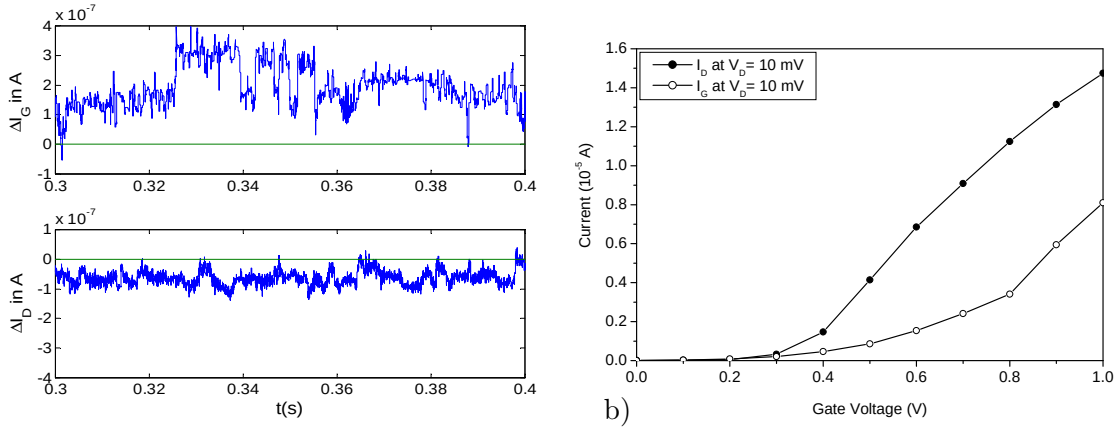

Fig. 5.19: RTS (a) and static gate and drain signals of an MOS-transistor after soft breakdown near the source. Graphs were taken from reference [5].

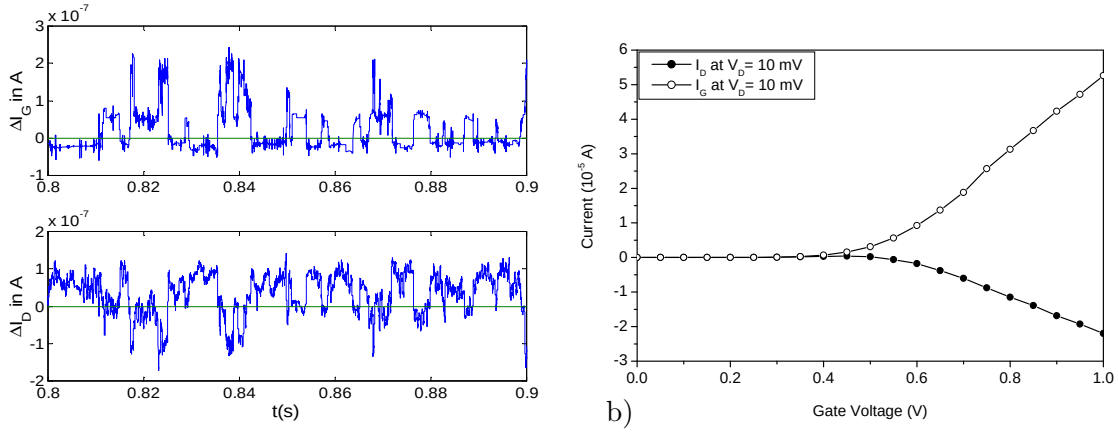

Fig. 5.20: RTS (a) and static gate and drain signals (b) of an MOS-transistor after soft breakdown near the drain. Graphs were taken from reference [5].

In the following, RTS and DC-signals recorded simultaneously for gate and drain currents will be presented [4]. All RTS-measurements were conducted by A. Avellán and published in reference [5]. NFETs with 27 Å RTP-oxides and gate lengths between 0.25 μm and 0.5 μm were subjected to a constant voltage stress of 5.1 V until the first breakdown occurred. After breakdown, constant gate (V_{G}) and drain (V_{D}) voltages were applied with bulk and source connected to ground. Gate (I_{G}) and drain (I_{D}) currents as well as their fluctuations were measured simultaneously at room temperature in separate amplifying circuits adopted from reference [40]. More details on the measurement can be found in reference [5].

Soft breakdown usually occurs where the local electric field across the oxide is highest or at places with a high density of hot carriers. Both conditions are met near the source- and the drain-region of a MOSFET. The defect can be localized by RTS and static currents as described in the following. RTS of gate and drain of sample A are shown in Fig. 5.19a. There is no correlation between the gate and drain RTS signal suggesting that the soft breakdown is at the source of the transistor. This is supported by the fact, that the static drain current shown in Fig. 5.19b shows a normal behavior.
RTS signals of gate and drain of a second sample (Fig. 5.20a), on the other hand, correlate very well, suggesting that the SBD is close to the drain. The opposite sign and very similar amplitudes between gate current fluctuations ΔI_{G} and drain current fluctuations ΔI_{D} suggest that the tunneling probability through the SBD-region fluctuates. This differs from the classical RTS which arises from modulations of the channel potential due to trapping and is probably negligible in this case. The results are also consistent with the static measurements shown in Fig. 5.20b, where the drain current is strongly effected by the gate leakage.

5.5 Modified CMOS-Process with RTP-Oxides

RTP-Oxides with the modified CMOS-process will be examined in this section in order to determine whether this process has an impact on the oxide properties. These samples can later on be used as a reference for metal gate transistors fabricated in the same way. During gate oxidation, source and drain are already doped which leads to a locally enhanced oxide growth. In Fig. 5.21 the CV-characteristics of capacitors with p-substrate (a) and n-substrate (b) are shown.

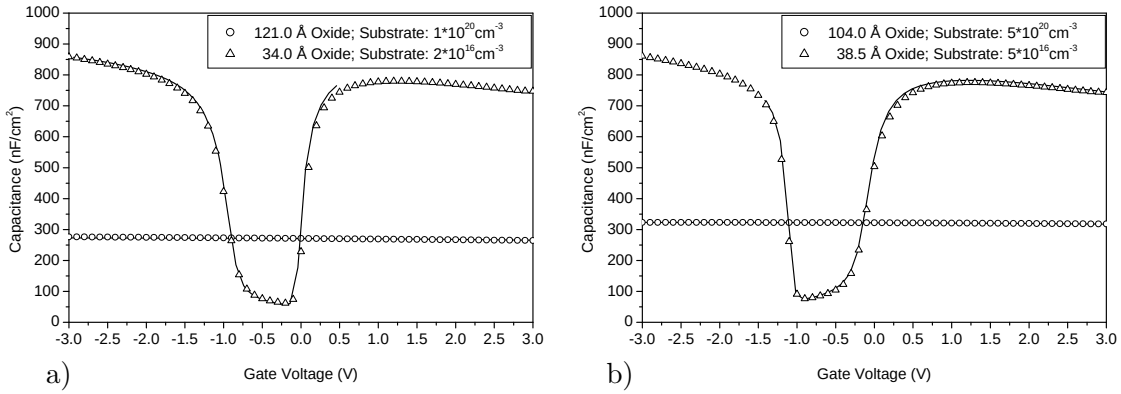

Fig. 5.21: CV-characteristics of capacitors fabricated using the modified CMOS-process for p-substrate (a) and n-substrate (b). Open symbols are measurement data while solid lines represent simulations.

Gate is n-doped with $8 \cdot 10^{19}$ cm^{-3} for all structures. Extracted oxide thicknesses and substrate doping values are summarized in the legend and in Table 5.3.

Substrate	p-	n-	p+	n+
Substrate Doping (cm^{-3})	$2 \cdot 10^{16}$	$5 \cdot 10^{16}$	$1 \cdot 10^{20}$	$5 \cdot 10^{20}$
Gate Doping (cm^{-3})	$8 \cdot 10^{19}$	$8 \cdot 10^{19}$	$8 \cdot 10^{19}$	$8 \cdot 10^{19}$
t_{ox} (Å)	34.0	38.5	121.0	104.0

Table 5.3: Parameters of modified CMOS-structures extracted from CV-curves.

Oxide grows significantly thicker on highly doped substrates. The expected transistor structure of the modified CMOS process is sketched in Fig. 5.22. Oxide thickness is

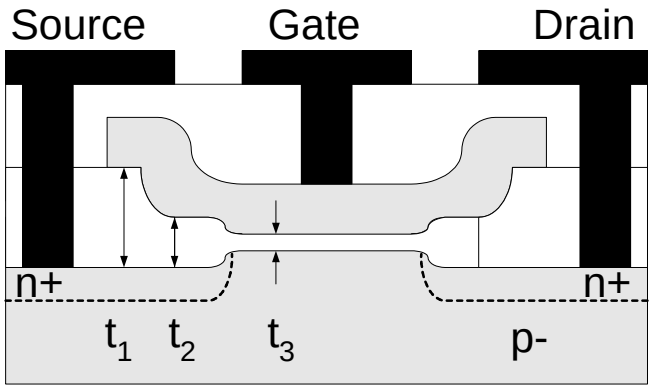

Fig. 5.22: Schematic of the modified CMOS structure as it is expected from the electrical data. Thicknesses are $t_1 = 60$ nm for the XP-oxide, $t_2 = 10$ nm above source/drain and $t_3 = 3.5$ nm for the gate oxide.

particularly large above source/drain, so that no breakdowns are expected in these regions during reliability analysis.

5.6 Summary

RTP-Tunnel-oxides that are used to study metal electrodes in Chapter 6 were characterized in this chapter. It has been possible to extract physical oxide thicknesses of MOS-capacitors down to 25 Å in good agreement with ellipsometric measurements and leakage current analysis. The same results were gained for fully integrated samples and for simple planar capacitors so that the latter can well be used for basic characterization. Modelling of the leakage current through thin oxides has been greatly improved by including some additional physical effects and by assuming a variable effective electron mass at the silicon/silicon-oxide interface. Simulations of the leakage current for all voltages and all oxide thicknesses studied were in excellent agreement with experimental data. A higher interface state density as compared to furnace oxides was observed for the RTP-oxides, but was still low enough to use these dielectrics to analyze metal electrodes. Finally, the modified CMOS-process has been characterized and showed no differences in MOS-properties compared to the standard CMOS-process. This technology can, therefore, be used to fabricate and study metal-gate transistors.

Chapter 6

Characterization of Metal Electrodes

This chapter describes electrical characterization of metal electrodes in MOS devices and their application to DRAM capacitors. Interface characteristics are studied on planar test structures with metal gate or substrate electrodes, respectively. Only metals that can be deposited in high aspect ratio trenches like CVD WSi_x and ALD TiN are suitable for DRAM application and have been investigated in this work. Limited thermal stability of metal gates on conventional dielectrics might require polysilicon/metal stacks to survive frontend temperatures. A polysilicon/TiN stack that can be integrated into a standard deep trench DRAM process was developed and is presented in the second half of this chapter.

6.1 Interface Characteristics of Metal Substrate Electrodes

Metal substrate electrodes for DRAM capacitors have the advantage that they lack a depletion region which further increases the overall capacitance. In addition, most metals under consideration have a higher work function than silicon and thus reduce the leakage current. For a specified maximum leakage current, the dielectric thickness can then be reduced which further increases the capacitance. For a successful integration, the substrate electrode has to support the defect healing process of the dielectric. Every dielectric has a certain density of defects right after deposition. Usually, plasma or thermal treatments are used to reduce the defect density before deposition of the top electrode. As described already in Section 4.1.1, silicon-rich WSi_x ($x > 2.0$) supports this process for standard NO [97]. In the following, test structures such as the ones mentioned in Section 3.2.1 with WSi_x as substrate electrode and NO as dielectric are characterized electrically. Fig. 6.1a shows capacitance-voltage curves of structures with three different WSi_x-compositions. The gate is phosphorus-doped to a level of $1 \cdot 10^{19}$ cm^{-3}.

Equivalent dielectric thicknesses have been estimated from the accumulation capacitance at -3 V. Extracted values of 4.1-4.3 nm are around 0.5 nm lower than typical values from polysilicon electrodes which is attributed to the lack of a depletion region in the substrate. The trend towards smaller capacitance values for higher silicon to tungsten ratios suggests that a pure silicon layer might be formed at the WSi_x/NO interface during high-temperature treatment after deposition. For highest capacitance, silicon-rich WSi_x with a composition close to the thermally stable phase $WSi_{2.0}$ will be most favorable. Also shown in Fig. 6.1a is the equivalent oxide thickness of a sample without surface passivation before NO-deposition as the one explained in Section 4.1.1. TEM-images show that the equiva-

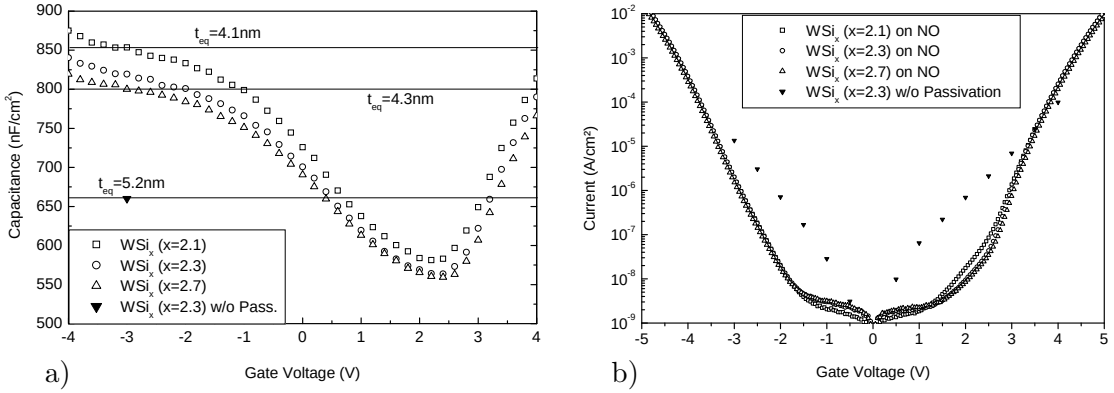

Fig. 6.1: Capacitance-voltage curves of MIS-capacitors with NO as dielectric and WSi_x of different composition as substrate electrode (a). Panel b) shows IV-characteristics of the same set of samples.

lent thickness increase is most likely due to oxide residuals at WSi_x grain-boundaries. Fig. 6.1b shows IV-data of the same samples as in Panel a). Leakage currents do not change with WSi_x-composition and are well below the typical DRAM specification of 10^{-8} A/cm^2 at $\pm$ 1 V gate voltage. A reason for the increased leakage current of samples without passivation could be the existence of metal oxides at the grain boundaries which would increase the electric field locally. In summary, a $WSi_{2.1}$-substrate electrode with passivation before NO-deposition satisfies the basic requirements for DRAM capacitors while increasing the capacitance by roughly 10%. Investigation of defect density and reliability values will be necessary to evaluate the process for further integration.

6.2 Electrical Properties of Metal Gate Electrodes

Metal gates deposited with processes such as the ones described in Chapter 4 are characterized in the following. WSi_x- and TiN-gates were investigated on planar structures and fully integrated wafers. While tungsten silicide gates investigated in this work exhibit insufficient leakage current behavior for thin gate oxides, MOS-structures employing TiN-gates show reproducible properties. Interface characteristics, leakage currents and thermal stability of these structures are, therefore, investigated in this section.

6.2.1 WSi_x-Gate Electrodes

Tungsten silicide is a well-known material in semiconductor industry and is regularly used together with polysilicon as polysilicon/metal-gate in DRAMs. The polysilicon serves as buffer layer that makes the whole gate stack stable up to high temperatures. Pure tungsten silicide gates have a work function close to midgap of bulk silicon making it potentially suitable for single-metal gate applications. There are, however, no data on the dependence of the work function on the composition of the silicide. Different MOS-capacitors were, therefore, fabricated by growing a nominal 9 nm thermal oxide on a highly n-doped substrate and depositing 55 nm of WSi_x on top of it. After annealing at 780 °C and structuring of the top electrode, CV- and IV-curves of structures with different WSi_x-compositions were measured as shown in Fig. 6.2 [97].

Due to the high substrate doping-level, only a minor variation in capacitance as a function of gate voltage is observed. For small absolute values of the gate voltage, silicon-rich WSi_x-gates ($x > 2.0$) show a capacitance similar to that of structures with polysilicon-electrodes which is equivalent to 9.0 nm oxide thickness. MOS-structures with tungsten-rich WSi_x, on the other hand, show an oxide thickness of 8.4 nm. The difference of 6 Å

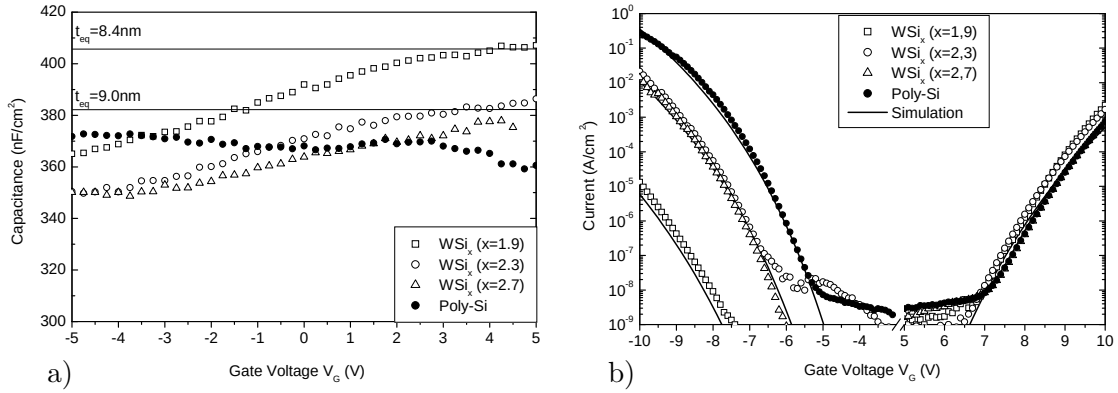

Fig. 6.2: Capacitance-voltage curves of MOS-capacitors with SiO_2 as dielectric and WSi_x of different compositions as gate electrode (a). Panel b) shows IV-characteristics of the same set of samples.

might indicate that the tungsten-rich layer lacks a depletion region while silicon-rich WSi_x depletes in inversion. This is supported by the IV-curves (Fig. 6.2b). For positive gate voltages, electrons tunneling through the oxide are emitted from the substrate conduction band. Since the substrate is the same for all samples, leakage currents are very similar. For negative gate voltages, electrons are emitted from the gate electrode. Large variations of the leakage current with electrode materials indicate differences in the work function. The parameters given in Table 6.1 were extracted by comparing experimental results with simulations.

Electrode material	WSi_x	WSi_x	WSi_x	Polysilicon
Composition x	1.9	2.3	2.7	-
$t_{ox}(CV)$	8.4	9.0	9.0	9.0
Band offset Φ (eV)	3.5	3.0	3.0	2.8

Table 6.1: Extracted parameters from MOS-capacitors with thermal oxide as dielectric and highly-doped substrates.

In order to study interface properties in more detail, samples with WSi_x-gate electrodes and RTP tunnel-oxides of different thickness ranging from 3-8 nm were prepared. However, for all samples leakage currents were very high which means that the gate oxide was already damaged severely by the WSi_x-layer. A deposition of this material using WF_6 and DCS as precursors directly on thin oxides is, therefore, not suitable. A barrier layer of any kind is required to protect the gate oxide from the WF_6-attack. As a consequence, no further investigation of WSi_x-gate electrodes is done in this work, whereas characterization of polysilicon/WSi_x-stacks will be presented later on in this chapter.

6.2.2 TiN-Gate Electrodes: Interface Properties

TiN gate electrodes have been analyzed in detail [121, 89, 68]. However, there is little information on TiN layers which have been deposited by an ALD-process directly on thin oxides [76]. The deposition process has a strong influence on the interface properties between oxide and gate electrodes and can also degrade the dielectric during deposition. Using the analysis described in Chapter 2, simple MOS-capacitors with TiN-gate electrodes, thin oxides and p-substrates are characterized at 58 points on each wafer. Substrate-doping was $1 \cdot 10^{16}$ cm^{-3} as determined from CV-measurements.

Fig. 6.3 shows the wafer maps of the extracted physical oxide thickness (Panel a) and

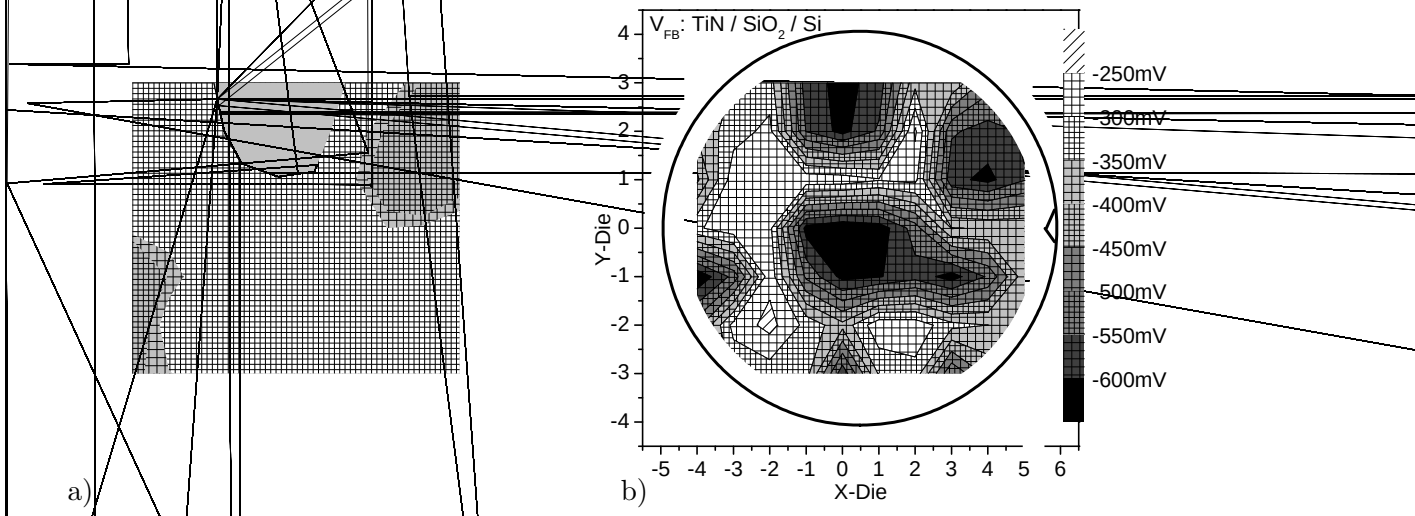

Fig. 6.3: Wafer maps of the physical oxide thickness (Panel a) and the flatband potential (Panel b) of a MOS-capacitor with a TiN-gate electrode deposited with the original ALD-TiN process. Nominal thickness of the oxide was 8.5 nm.

the flatband potential (Panel b). The standard ALD-process as supplied by the vendor was used to deposit TiN on an 8.5 nm RTP-oxide and annealed at 800 °C for 60 s in N_2. In comparison to the polysilicon electrodes described in Chapter 2, a strong variation in oxide thickness as well as in flatband potential is observed across one wafer. At the same time, leakage currents through thin oxides were very high. The patterns on both wafer maps are very similar suggesting that there is a correlation between flatband potential and oxide thickness. One possible explanation is a strong variation of the TiN composition near the interface. In regions of stoichiometric TiN an oxide thickness of 8.5 nm and a flatband potential of around -250 mV would be expected from literature [104, 23, 121, 54]. These values are indeed measured in some regions of the wafer. At the center and some sites at the wafer edge, however, flatband potentials below -650 mV and oxide thicknesses smaller than 8 nm indicate the existence of a titanium-rich TiN layer [64]. Such a layer is thermally unstable at 800 °C and reacts with the SiO_2 reducing its thickness. Pure Ti-gates would result in a flatband potential of -700 mV which correlated well to the measured values [111].

The wafer patterns have been compared to marks which were visible inside the TiN deposition chamber. These marks were created by non-uniformities of the plasma, that initiates the TiN-ALD process. To avoid these problems, an ALD-TiN process has been developed, that works completely without a plasma. Wafer maps for the same structure as shown in Fig. 6.3 but with the new TiN-ALD process are presented in Fig. 6.4. Non-uniformities of both quantities are being highly reduced, although there still are some variations across the wafer. In addition, leakage current through the oxide decreased substantially in particular for thin oxides as can be seen in Fig. 6.5a. The reduction in leakage current facilitates CV-measurements which are presented in Panel b) of the same figure.

Flatband voltages of the simulation were adjusted to described CV-data. A value of around -250 mV was assumed for all curves. It can be seen that the measured curves can be described well by the simulation, only the thinnest oxide shows major deviations for negative voltages. This is attributed to a direct leakage current through the oxide that makes the thinnest oxide unsuitable for automatic parameter extraction. The influence of the leakage current can be reduced by increasing the measurement frequency. With the setup used in this work, frequencies well above 100 kHz led to significant errors due to

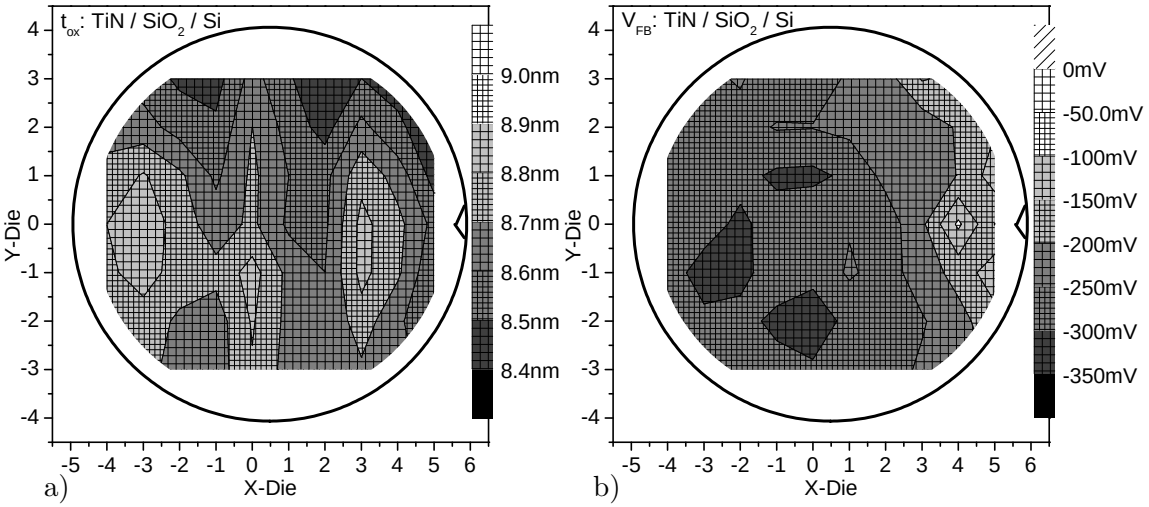

Fig. 6.4: Wafer maps of the physical oxide thickness (Panel a) and the flatband potential (Panel b) of a MOS-capacitor with a TiN-gate electrode deposited with the newly developed ALD-TiN process. Nominal thickness of the oxide was 8.5 nm.

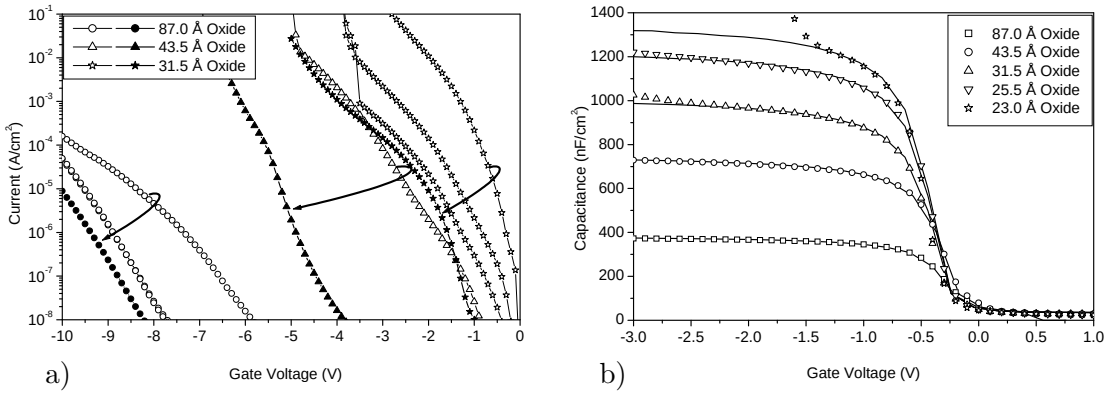

Fig. 6.5: Leakage currents through MOS-capacitors with TiN-gates and oxides of different thickness are presented in Panel a). Open symbols indicate samples that were prepared with the original TiN-ALD process while solid symbols show the same data for the improved TiN-process. Panel b) displays CV-curves of MOS-structures with TiN-gate electrodes and RTP-oxides of different thickness. Symbols represent measured data, while simulations are shown as solid lines. Extracted oxide thicknesses are given in the legend.

the inductance of the prober. Measurements were conducted at 100 kHz so that, from the data presented, it can be concluded that 25 Å is the minimum oxide thickness which can be evaluated with the method described in Chapter 2. Wafermaps of the remaining samples are presented in Fig. 6.6. Generally, there is a good oxide thickness uniformity across each wafer indicating that there is no severe reaction between the gate oxide and the TiN-electrode. The patterns observed for the flatband potential, on the other hand, are very similar for all oxide thicknesses. To distinguish between work function- and interface-charge-variation, the extracted data were plotted as a V_{fb} versus t_{ox} plot as shown in Fig. 6.7a. Connected are data points at the same position on the wafer. The slope of all lines is very similar while there is an offset of up to 100 mV between them. The similar slope suggests that there is a constant interface charge density $N_{\mathrm{f}} = 4.4 \cdot 10^{11}$ cm^{-2} across the whole wafer while the work function difference ϕ_{MS} has a range of 100 meV. N_{f} is roughly a factor of two higher than for polysilicon electrodes.

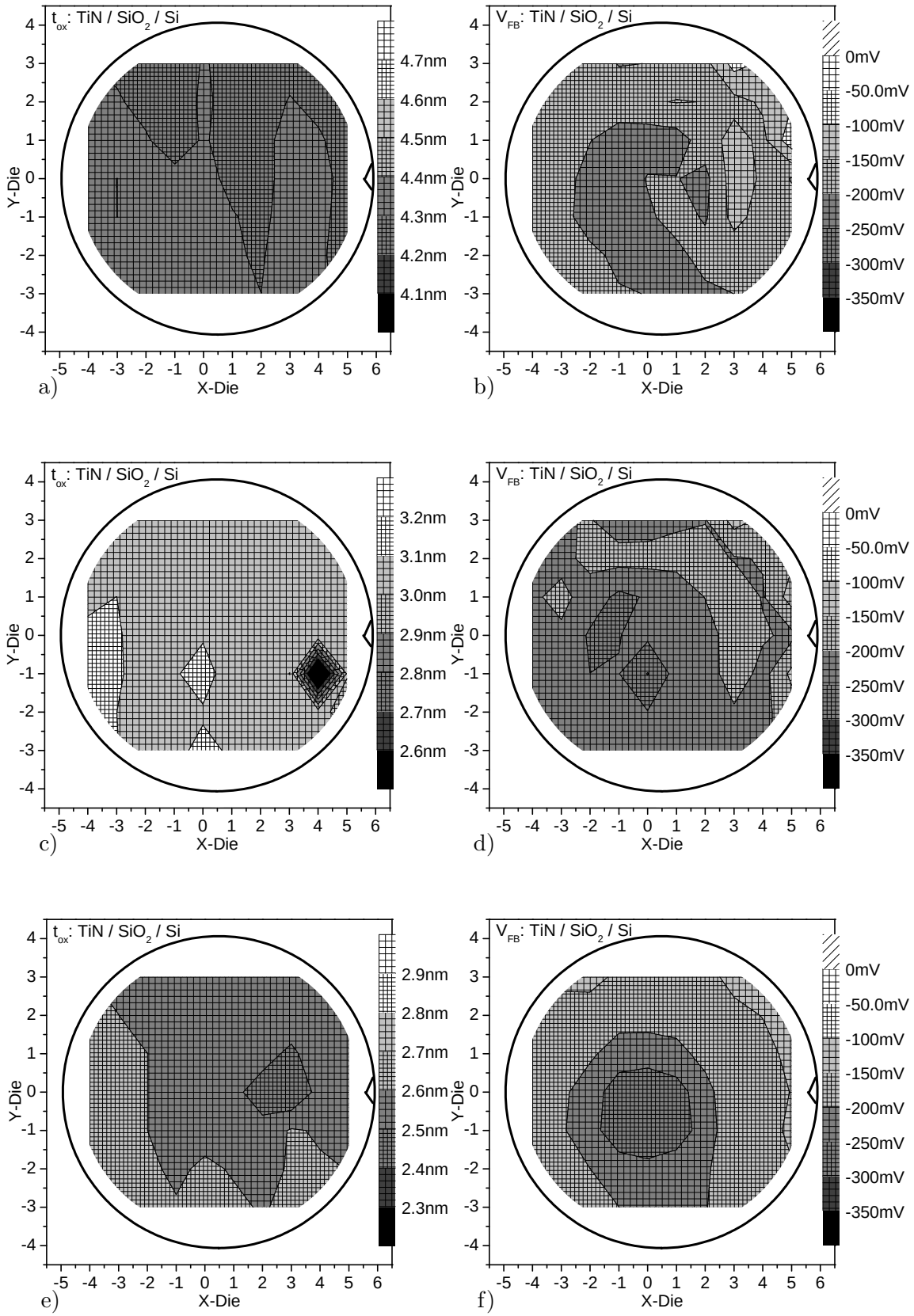

Fig. 6.6: Wafermaps of the physical oxide thickness and the flatband potential of MOS-structures with TiN-gates and RTP-oxides of different thickness on p-type substrate. An ALD-process without plasma initiation was used to deposit the TiN.

Fig. 6.7b shows a reference measurement with polysilicon electrodes. The range across the whole wafer is only 20 meV. It can, therefore, be concluded that the variation of ϕ_{MS} in Panel a) stems from non-uniformities in work function of the gate electrode. In addition, there seems to be an offset for the two samples with the thinnest oxides. These two oxides were grown with a 1000 °C RTP-oxidation, while the two other samples were oxidized at 1120 °C. The different oxidation temperatures might lead to different oxide-surfaces which in turn have an influence on the initiation of the TiN-ALD process.

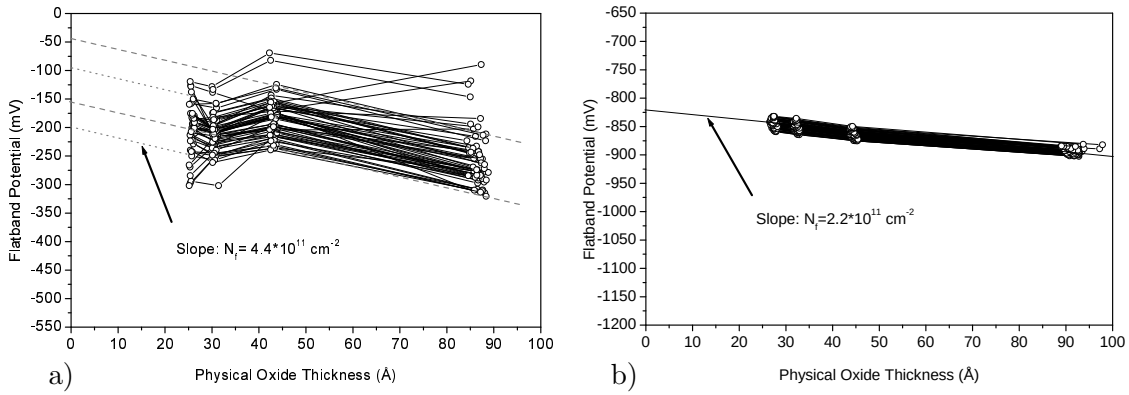

Fig. 6.7: Flatband potential as a function of physical oxide thickness of MOS-structures with RTP-oxides and TiN-gates (a) or polysilicon-gates (b).

A second set of samples with n-type substrates was processed to verify this observation. Measured and simulated CV-curves are shown in Fig. 6.8. A substrate doping level of $5 \cdot 10^{14}$ cm^{-3} was assumed for best description.

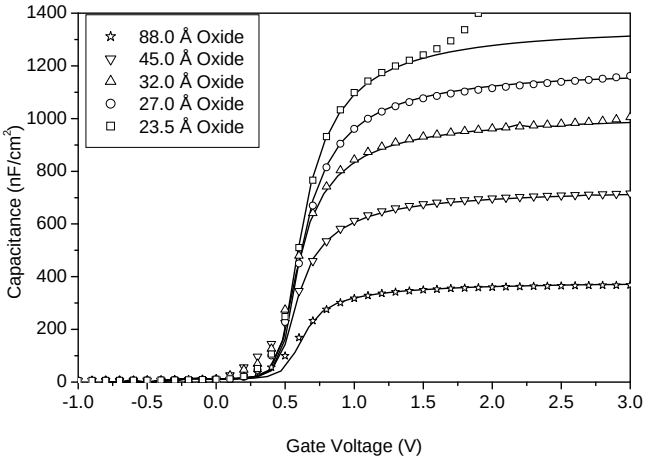

Fig. 6.8: CV-curves of MOS-structures with TiN-gate electrode and RTP-oxides of different thickness. Symbols present measured data, while simulations are shown as solid lines. Extracted oxide thicknesses are given in the legend.

Using this value, automatic extraction was conducted in the usual way and pertinent wafermaps are presented in Fig. 6.9. Similar patterns as observed for the p-doped samples support the assumption, that these variations stem from the TiN-initiation process. The region of reduced flatband potential close to the center of the wafer suggests that a titanium-rich layer is formed at this site of the wafer.

Analysis of the flatband potential as a function of physical oxide thickness shows an even higher spread as was observed for the p-type substrate (Fig. 6.10). The same interface charge density and a similar offset in flatband potential for the thinner oxides are

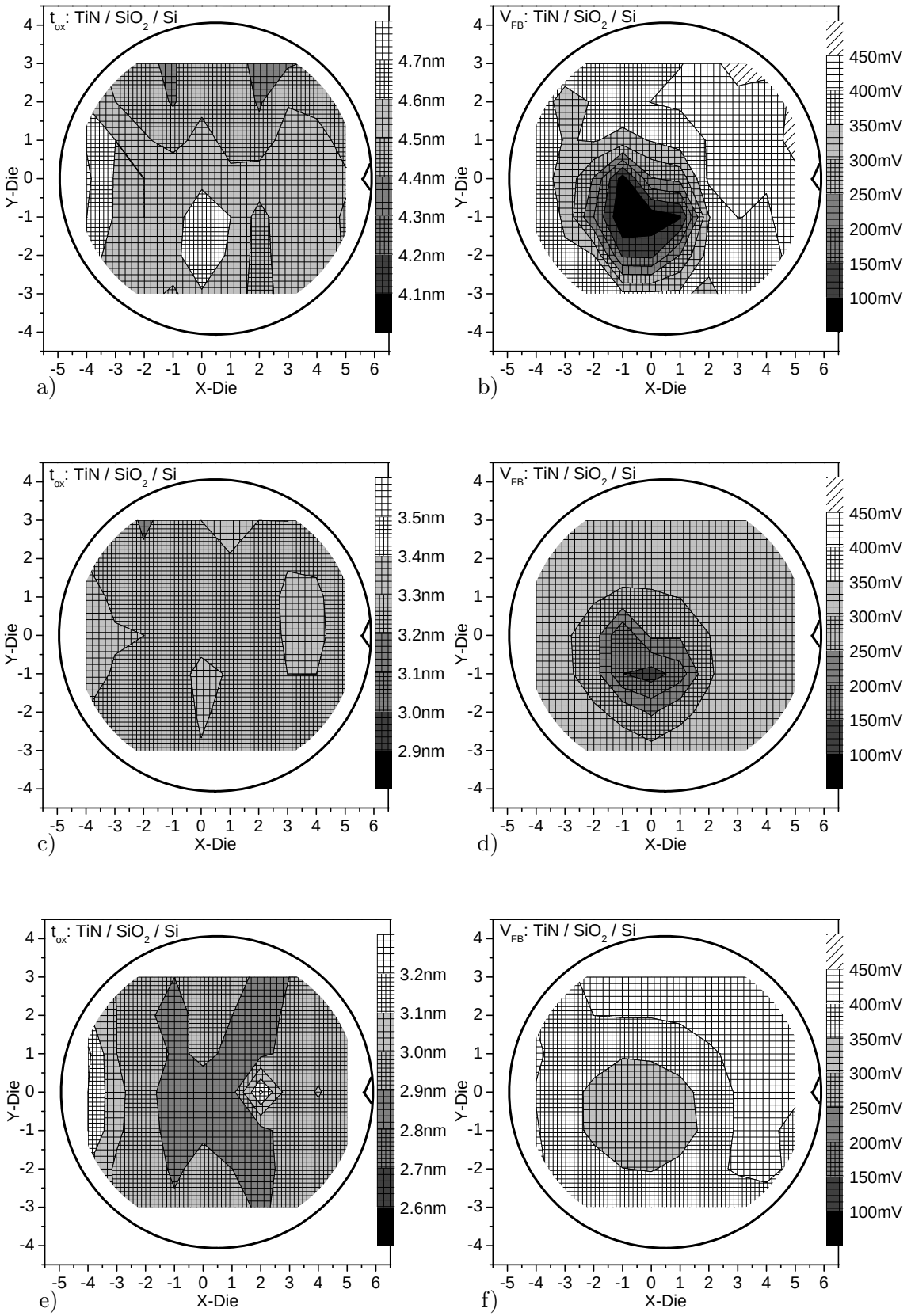

Fig. 6.9: Wafermaps of the physical oxide thickness and the flatband potential of MOS-structures with TiN-gates and RTP-oxides of different thickness on n-type substrate. An ALD-process without plasma initiation was used to deposit the TiN.

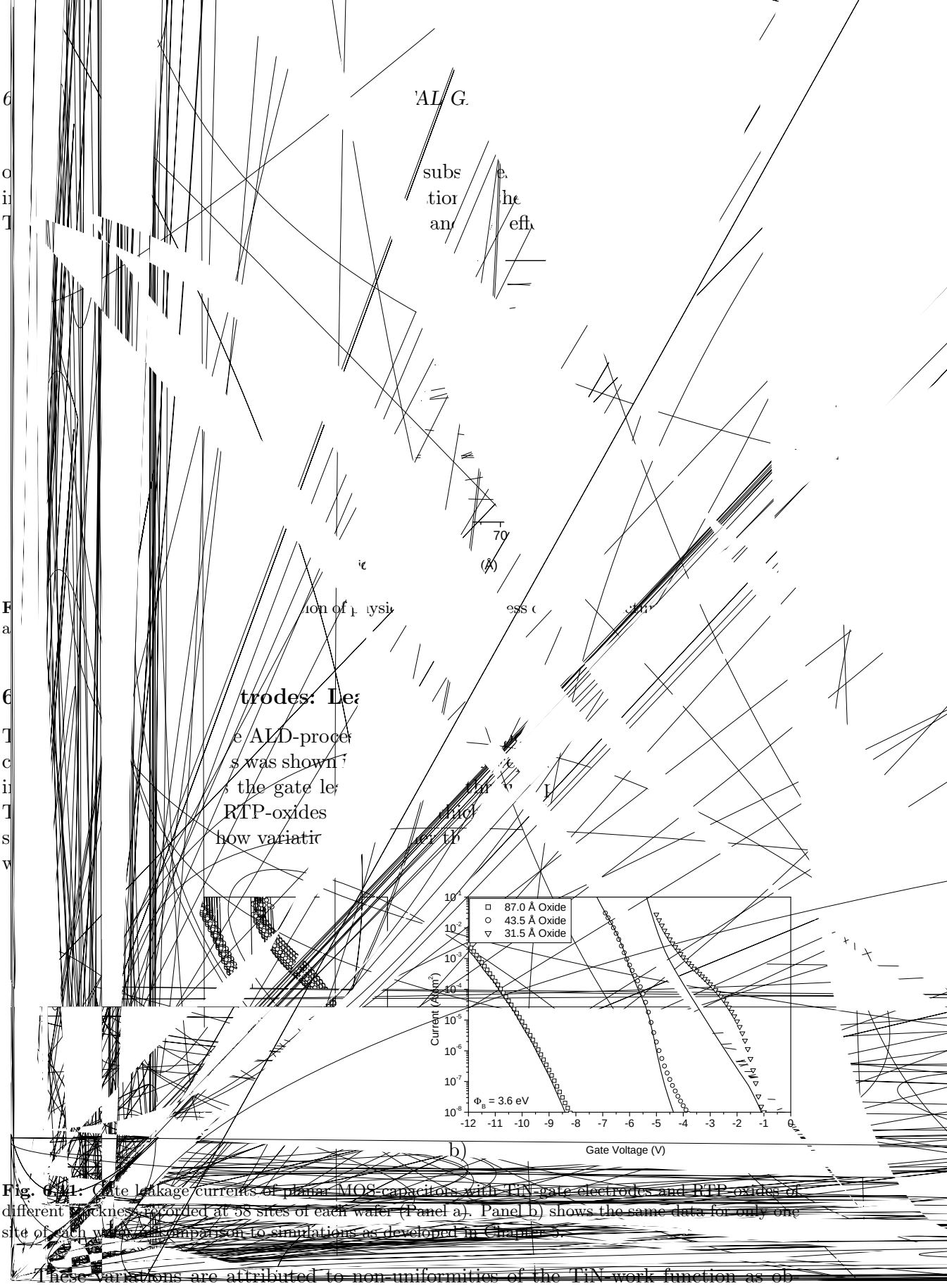

Fig. 6.11: Gate leakage currents of planar MOS capacitors with TiN gate electrodes and RTP oxides of different thickness, recorded at 38 sites of each wafer (Panel a). Panel b) shows the same data for only one site of each wafer in comparison to simulations as developed in Chapter 5.

These variations are attributed to non-uniformities of the TiN work function as observed by CV-measurements and to some variation in oxide thickness. When assuming oxide thicknesses as extracted from CV-measurements, good agreement is reached for thick oxides and a band offset $\Phi_B = 3.6$ eV. There is, however, a region between 1 and 5 Volt where the measured leakage current is significantly higher than the one predicted by simulation. It has been shown by Yang et al. that pure titanium gate electrodes can result in a large stress within the gate [121]. Measurements done in this work revealed an

as-deposited tensile stress of around 1000 MPa. As demonstrated by Matsuhashi et al. for tungsten gates, mechanical stress in the as-deposited film can strongly effect the interface charge density between the silicon substrate and the SiO_2 [70]. At the same time, traps might be generated at the interface and within the oxide that act in the same way as traps generated by electrical stress leading to a stress-induced leakage current (SILC). This current is proportional to the trap density and dominates the Fowler-Nordheim tunneling below 5 V gate potential [88].

6.2.4 TiN-Gate Electrodes: Thermal Stability

Planar capacitors with p-substrate, 45 Å SiO_2 and 20 nm TiN were fabricated and annealed at temperatures between 650 °C and 900 °C. CV-measurements presented in Fig. 6.12a revealed a flatband potential of -650 mV after annealing at 650 °C which changed to -200 mV after annealing at 800 °C or higher temperatures.

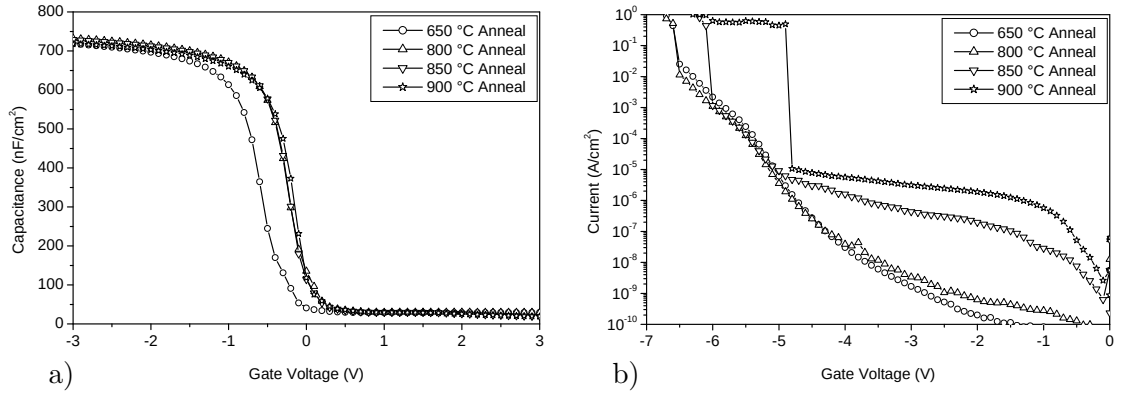

Fig. 6.12: Capacitance-voltage curves (a) and leakage current (b) of MIS-structures with a p-substrate, 45 Å SiO_2 and a TiN gate electrode. Samples were annealed at temperature ranging from 650 °C to 900 °C for 60 s in nitrogen.

A flatband potential of -200 mV indicates a near-midgap material with a work function of 4.8 eV which is in very good agreement to literature values for TiN [104]. The sample annealed at 650 °C suggests that the gate has a work function of roughly 4.4 eV after deposition which might be due to titanium-rich TiN at the interface. During an anneal above 650 °C excess nitrogen diffuses from the bulk TiN to the interface. This leads to a stoichiometric TiN that is thermally stable during high temperature anneals. Severe stress during these thermal treatments, however, led to an increase in tunneling currents as shown in Panel b) of Fig. 6.12. Already the sample annealed at 800 °C shows increased leakage currents for low voltages. An annealing temperature between 650 °C and 800 °C will be ideal to obtain a stoichiometric TiN-layer that does not degrade the gate oxide.

6.3 Thermal Stability of Polysilicon/Metal Gate Electrodes

Stacks of polysilicon and refractory metal silicides are widely used in production as gate electrodes. In the following, these stacks will be referred to as metal-silicon-insulator-silicon (MSIS) structures. While the polysilicon-layer leads to the well-known silicon-oxide-silicon structure, the metal serves to reduce the resistance. A similar stack might be useful for deep trench DRAM-capacitors, but due to the limited amount of space, the polysilicon-buffer should be as thin as possible. While gate stacks are typically fabricated with 80 nm polysilicon, a maximum of around 20-30 nm is acceptable for the deep trench fill of future generations. First, a polysilicon/WSi_x-stack is characterized on planar capacitors.

It will be shown, that there are a number of problems occurring from the tungsten silicide deposition process, so that an alternative stack employing TiN instead of WSi_x will be analyzed in the second half of this section.

6.3.1 Polysilicon/WSi_x-Gate Electrodes

Beside the challenges during deposition of WSi_x described in Chapter 4, there are some phenomena that occur when these layers are deposited on phosphorus-doped polysilicon. Polysilicon/metal-stacks are usually fabricated in a single-wafer cluster-tool in which wafers are transferred under vacuum from one chamber to another. This has the advantage that no thermal oxide layer can form between the polysilicon and the metal. The disadvantage is, that the initiation of the metal deposition process depends strongly on the properties of the polysilicon-layer. During the initial phase of a WSi_x deposition with DCS and WF_6, the WF_6 is mainly reduced by surface silicon instead of DCS which leads to a tungsten-rich layer near the interface.

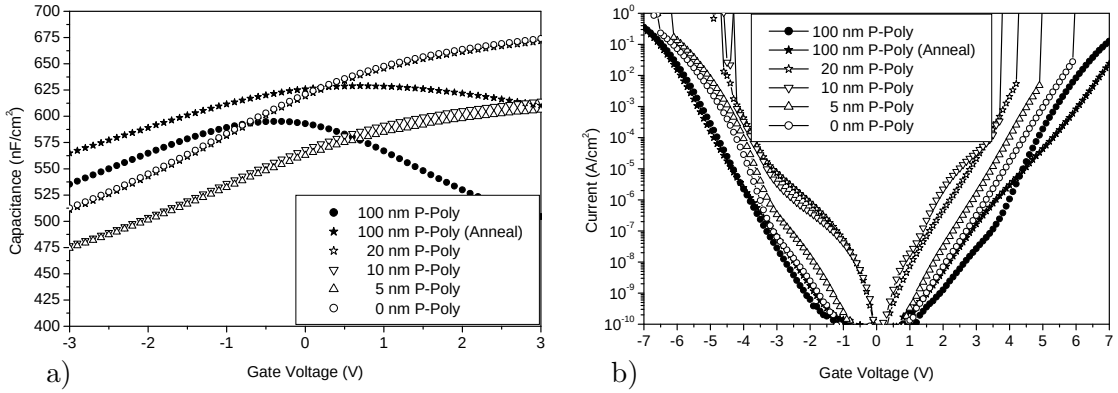

Fig. 6.13: Capacitance-voltage curves (a) and leakage current (b) of MSIS-structures with a polysilicon/WSi_x-stack for different polysilicon thicknesses. All samples were subjected to a 780 °C furnace-anneal and those stated *anneal* were additionally oxidized at 1050 °C.

To avoid this problem, deposition can be started with silane instead of DCS as is done regularly in production [47]. But even then, phosphorus atoms in the polysilicon will reduce WF_6 which at the same time increases the tungsten content in the film [13]. This problem is avoided by depositing a 20 nm undoped polysilicon-layer before the WSi_x. High-temperature anneals during further processing lead to sufficient dopant diffusion to guarantee an ohmic contact between these two layers. Fig. 6.13 compares electrical properties of planar capacitors with polysilicon/WSi_x-stacks, where the polysilicon thickness had been varied between 0 nm and 100 nm. Two sets of samples were prepared. One was subjected to a small thermal budget up to temperatures of only 780 °C. The second set stated as *anneal* was oxidized at 1050 °C for 85 s. Only samples with the thickest polysilicon film (100 nm) survived this treatment. Capacitance measurements and leakage currents of the samples before annealing, however, do not show a consistent picture as can be seen in Fig. 6.13. This might be due to the lowly doped polysilicon buffer that distorts electrical data substantially.

It is, therefore, more interesting to look only at samples that have been fully activated by a high-temperature treatment. Fig. 6.14 shows capacitance-voltage measurements (Panel a) and leakage currents (Panel b) of MSIS-structures with polysilicon/WSi_x-stacks, NO as dielectric and p^--substrates. Data for 40 nm and 100 nm polysilicon thickness, respectively, are comparable showing good leakage current characteristics. Additional data of samples with a 20 nm polysilicon buffer have been excluded from the graph because high

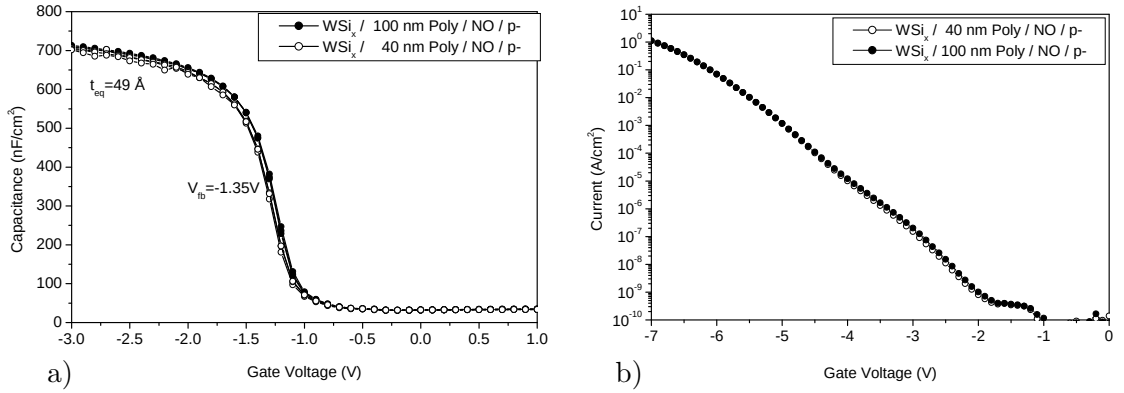

Fig. 6.14: Capacitance-voltage curves (a) and leakage current (b) of MSIS-structures with polysilicon/WSi_x-stacks of different polysilicon thickness. All samples were oxidized at 1050 °C.

leakage currents were observed. The doping level of the polysilicon has a strong impact on the whole stack. In the following, two polysilicon/metal-stacks are compared that have a nominal buffer thickness of 20 nm. The TEM-image in Fig. 6.15 shows a cross section of a stack consisting of a lowly doped polysilicon and a silicon-rich WSi_x after anneal.

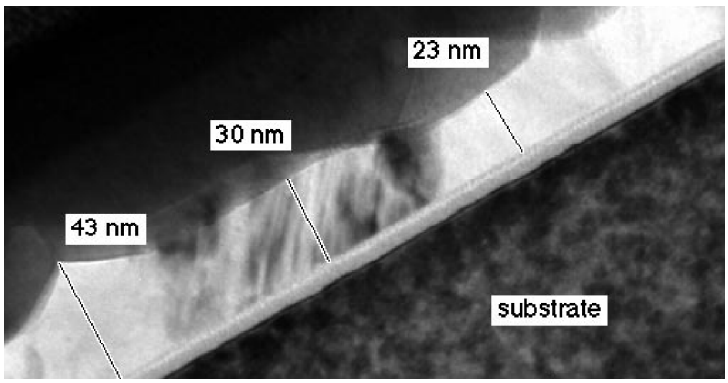

Fig. 6.15: TEM-image of the cross section of a MSIS-structure with silicon-rich WSi_x and a lowly doped polysilicon buffer.

The buffer thickness varies strongly, but is always larger than 20 nm. Preparing the same samples with highly doped silicon leads to structures as shown in Fig. 6.16. Panel a) presents the cross section revealing regions where the buffer had been fully consumed during the anneal. This might indicate that the WSi_x was initially deposited tungsten-rich which led to a diffusion of silicon into this layer. In addition, voids were formed in the WSi_x-layer as can be seen in the SEM-image in Fig. 6.16b.

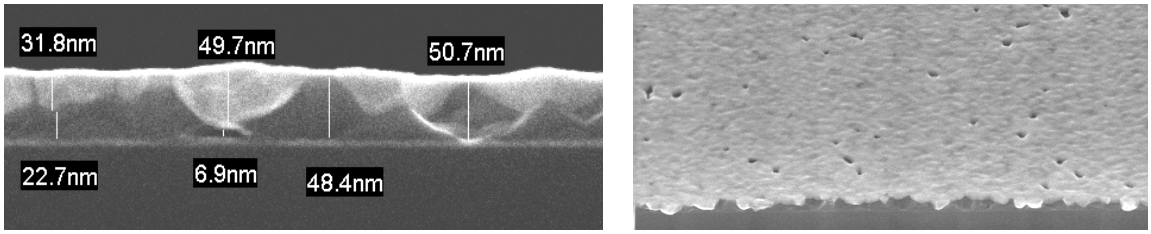

Fig. 6.16: SEM-images of an MSIS-structure with WSi_x, doped polysilicon, NO and p-substrate. Panel a) shows a cross section while Panel b) presents an angled view of the WSi_x-layer.

In summary, the WSi_x deposition process used incorporates many drawbacks that make an integration into deep trench capacitors very difficult. While it might be possible to solve these problems, other polysilicon/metal-stacks are more promising. In particular, a deposition technique is required that enables good step coverage in high-aspect ratio trenches and at the same time deposits stable layers on highly doped polysilicon. For this purpose, an ALD-TiN layer is studied in the following.

6.3.2 Polysilicon/TiN-Gate Electrodes

Planar MSIS-capacitors were fabricated on p-substrates, NO as dielectric and a stack of nominal 40 nm polysilicon and 30 nm ALD-TiN. The TiN was deposited with the original process that includes plasma initiation. Thereafter, the samples were annealed in pure nitrogen for 60 s at temperatures ranging from 800 °C to 1050 °C. CV- as well as IV-curves are presented in Fig. 6.17.

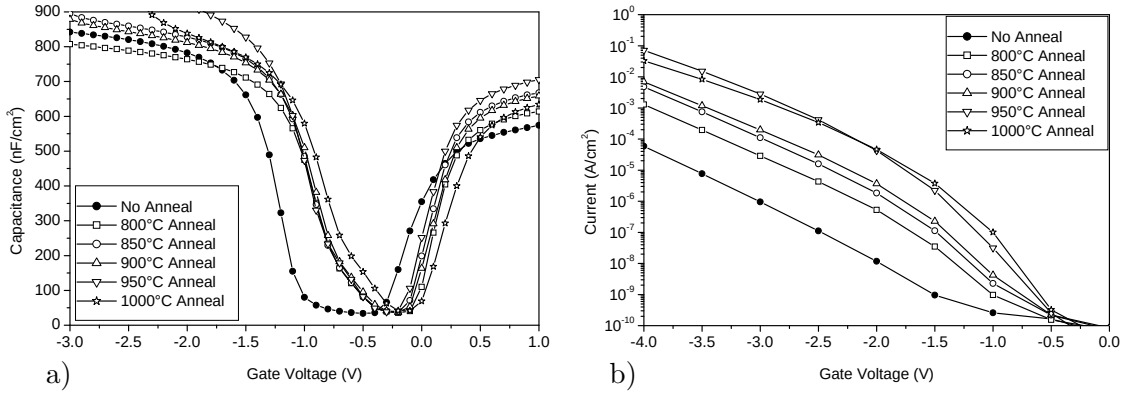

Fig. 6.17: CV- (a) and IV- (b) data of MOS-capacitors with p-substrate, NO as dielectric and a stack of 40 nm polysilicon and 30 nm TiN that were annealed at various temperatures.

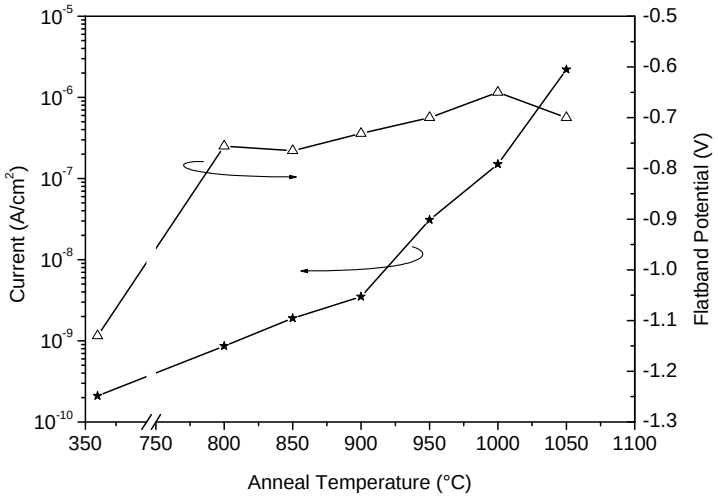

Fig. 6.18: Flatband potential and leakage current at -1 V of planar MSIS-capacitors with p-substrate, NO as dielectric and a stack of 40 nm polysilicon and 30 nm TiN. The sample that was not annealed is plotted at the deposition temperature of 360 °C.

Generally, it is observed that CV-curves as well as the gate leakage are substantially modified by high-temperature treatments. The shift in the CV-curves is best described by the extracted flatband potential while for DRAM-applications the leakage current at -1 V

is of most interest. Both quantities are summarized as functions of anneal temperature in Fig. 6.18.

Due to the TiN-ALD process, the flatband potential varies strongly across each wafer, so that the plotted values should be considered as a trend only. The shift between 360 °C and 800 °C anneal indicates that some interfacial layer existed either between the NO and the polysilicon or between the latter and the TiN. In addition, the leakage current increased by four orders of magnitude after annealing at 1050 °C. Note that a strong variation of the work function across a whole wafer was observed on MOS-capacitors with TiN-gates which is most likely due to variations in the TiN-composition as shown before. When polysilicon/TiN-stacks are fabricated, titanium-rich regions on the wafer lead to a formation of titanium-silicide (TiSi) that consumes the polysilicon-buffer and increases the stress. A careful investigation of the metal deposition process is, therefore, indispensable when the thermal stability of polysilicon/metal-stacks is to be evaluated.

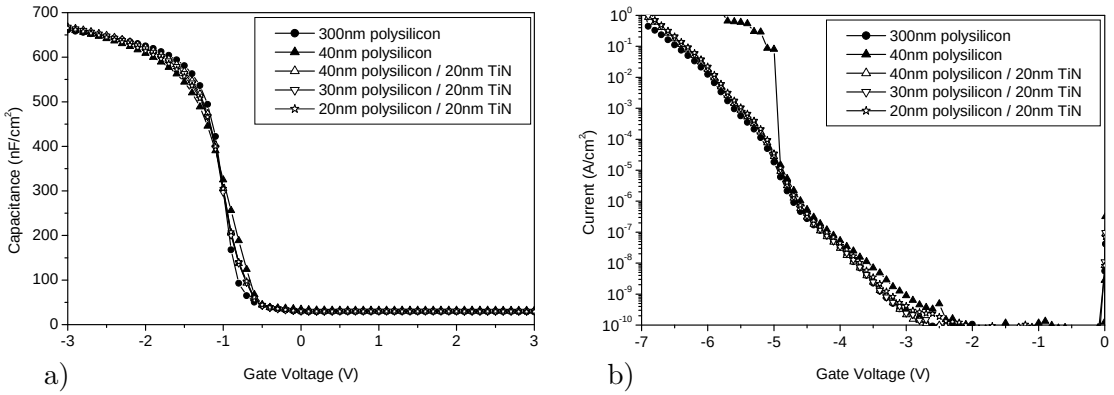

Fig. 6.19: CV- (a) and IV- (b) data of MOS-capacitors with p-substrate, 45 Å SiO_2 as dielectric and a stack of polysilicon and TiN for different polysilicon thicknesses. All samples were annealed at 1000 °C for 60 s in pure nitrogen.

The following characterizes polysilicon/TiN-stacks which include the newly developed TiN-ALD process without plasma and polysilicon-thicknesses ranging from 20 nm to 40 nm. All samples incorporate 45 Å SiO_2 as gate dielectric and p-type substrates. For comparison pure polysilicon gates with 40 nm and 300 nm thickness were prepared and all samples annealed at 1000 °C for 60 s in pure nitrogen.

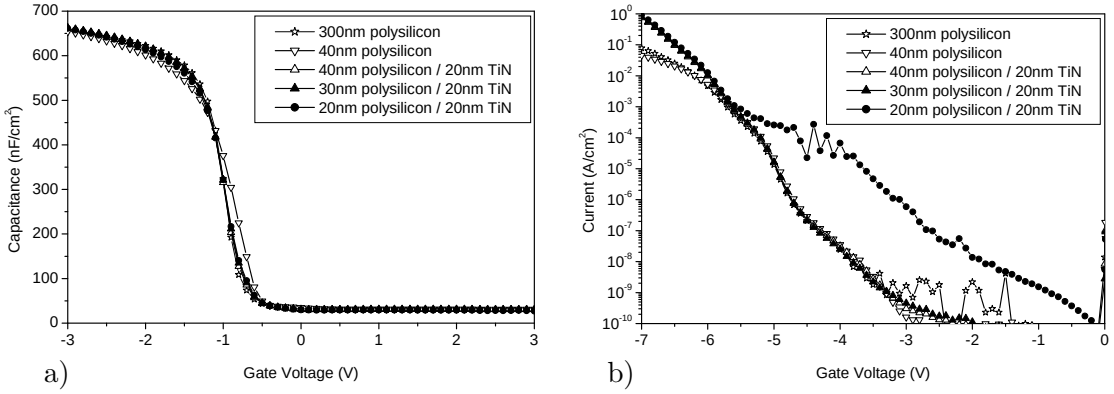

Fig. 6.20: CV- (a) and IV- (b) data of MOS-capacitors with p-substrate, 45 Å SiO_2 as dielectric and a stack of polysilicon and TiN for different polysilicon thicknesses. All samples were annealed at 1050 °C for 60 s in pure nitrogen.

Fig. 6.19 demonstrates a flatband potential of -900 mV for all capacitors which is the expected value for highly n-doped silicon gates on lowly doped p-type substrates.

Leakage currents are very low for all structures indicating that a 20 nm polysilicon buffer is sufficient to avoid any degradation of the gate oxide during a 1000 °C anneal. The early breakdown of the sample with 40 nm polysilicon and no TiN is attributed to the mechanical stress applied during the electrical measurements. Electrical data of a similar set of samples annealed at 1050 °C for 60 s in nitrogen is presented in Fig. 6.20. While there is no change in CV-data, a significant increase in leakage current is observed for capacitors with a 20 nm polysilicon buffer. In conclusion, a minimum buffer thickness between 20 nm and 30 nm is required if the gate dielectric has to withstand a 1050 °C anneal during further processing.

6.4 Metal Electrodes for DRAM Capacitors

This section describes the direct application of thin metal layers in deep trench DRAM capacitors. The results that were presented so far will be used to fabricate deep-trench capacitors with a polysilicon/TiN-stack as inner electrode that withstand typical frontend temperatures.

6.4.1 Material Properties of Thin Metal Layers

Three different metals, namely WSi_x, WN and TiN, were studied in this work. Basic properties of these materials were examined on blanket wafers and will be described in the following.

Tungsten Silicide

The resistivity of tungsten silicide as a function of composition has been demonstrated in detail in Chapter 4. In the following, an example of the resistance of $WSi_{2.3}$ as a function of thickness is presented.

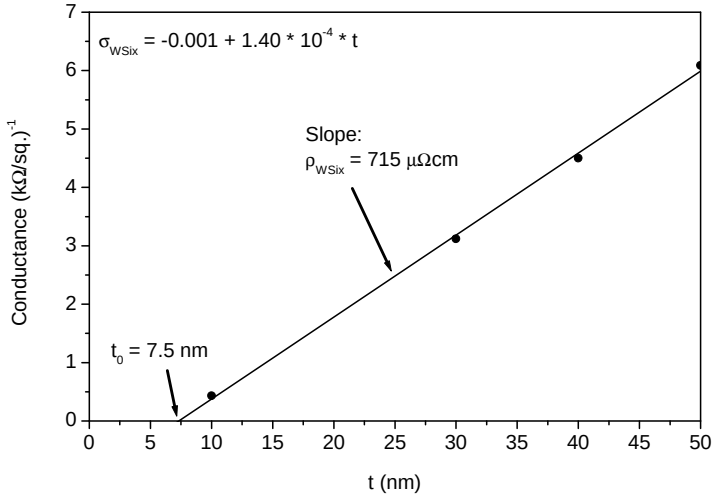

Fig. 6.21: Measured conductance of tungsten-silicide layers (Si:W = 2.3:1) as a function of layer thickness.

Fig. 6.21 shows the measured conductance of tungsten silicide layers with thicknesses ranging from 10 nm to 50 nm. From the slope, a bulk resistivity of 715 $\mu\Omega$cm is extracted while the intercept at zero conductance yields a thickness of 7.5 nm of an interfacial layer that does not contribute significantly to the overall conductance. This layer has to be taken into account when small-geometry devices are to be built with thin metal layers. The thickness, however, might be a strong function of deposition conditions and annealing

temperature. As stated in Chapter 4, the resistivity reduces to 60 $\mu\Omega$cm after a high-temperature anneal.

Tungsten Nitride

WN-layers fabricated by two different deposition techniques have been investigated in this work. Tungsten nitride CVD with ammonia and WF_6 as precursors leads to smooth layers with a good as-deposited resistivity of 180 $\mu\Omega$cm and an interface layer of 5 nm (Fig. 6.22a). However, a similar process development as the one for the tungsten silicide (see Chapter 4) would be necessary to deposit this material into deep trenches.

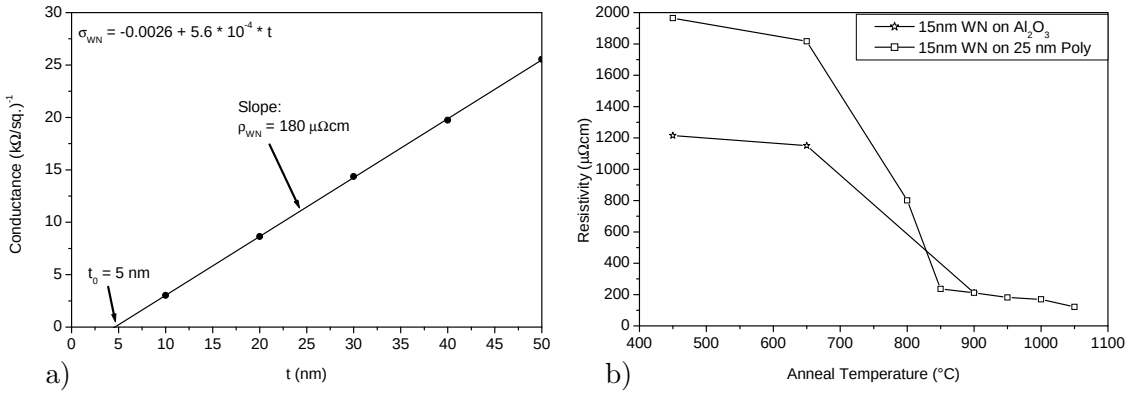

Fig. 6.22: Conductance of CVD-WN layers as a function of thickness (Panel a). Panel b) shows the specific resistivity of a 15 nm ALD-WN layer as a function of annealing temperature.

Tungsten nitride layers deposited by ALD showed good step coverage and are characterized in the following. Fig. 6.22b presents the specific resistivity of a 15 nm ALD-WN layer as a function of annealing temperature. The as-deposited resistivity is very high while layers annealed above 850 °C exhibit similar values than the CVD-WN. The high resistivity after deposition can be explained by the composition shown in Table 6.2 as determined by RBS measurements.

Quantity	Wafer Center	Wafer Edge
Tungsten (cm^{-2})	$7.58 \cdot 10^{16}$	$8.03 \cdot 10^{16}$
N:W Ratio	1.65:1	1.60:1
Titanium (cm^{-2})	$0.042 \cdot 10^{16}$	$0.170 \cdot 10^{16}$

Table 6.2: Composition of as-deposited ALD-WN as determined by RBS measurements at the wafer center and the wafer edge.

The measured N:W ratio of 1.6:1 is significantly higher than the thermally stable phase with N:W = 0.5:1. During an anneal, the excess nitrogen diffuses out of the layer which leads to a severe shrinking of the material as seen in the SEM-images of Fig. 6.23. Massive crack formation is observed which can only be avoided by changing the deposition process in a way in which less nitrogen is incorporated into the layer. Due to the nature of the ALD-process this is not easily achieved by changing the deposition parameters during

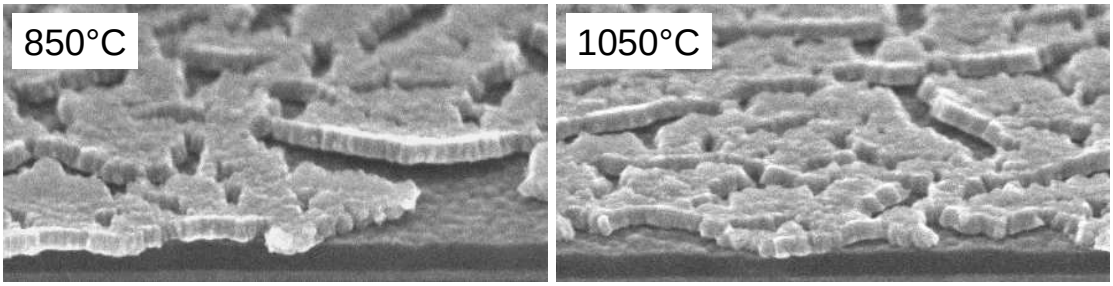

Fig. 6.23: SEM-images of an ALD WN-layer annealed at 850 °C (left Panel) and 1050 °C (right Panel).

processing. Rather, a modification of the precursors used will be necessary. This, however, could not be done in this work. For this reason, ALD-WN has not been investigated further for deep-trench applications. The remainder of this section will, therefore, focus on ALD-TiN layers.

Titanium Nitride

ALD-TiN layers of different thickness were deposited on thermal oxide and the conductance measured (Fig. 6.24). A specific resistivity of 95 $\mu\Omega$cm and an interface-layer thickness of 5 nm have been determined for the as-deposited layers. The resistivity improves even further after high-temperature anneals (Fig. 6.24b).

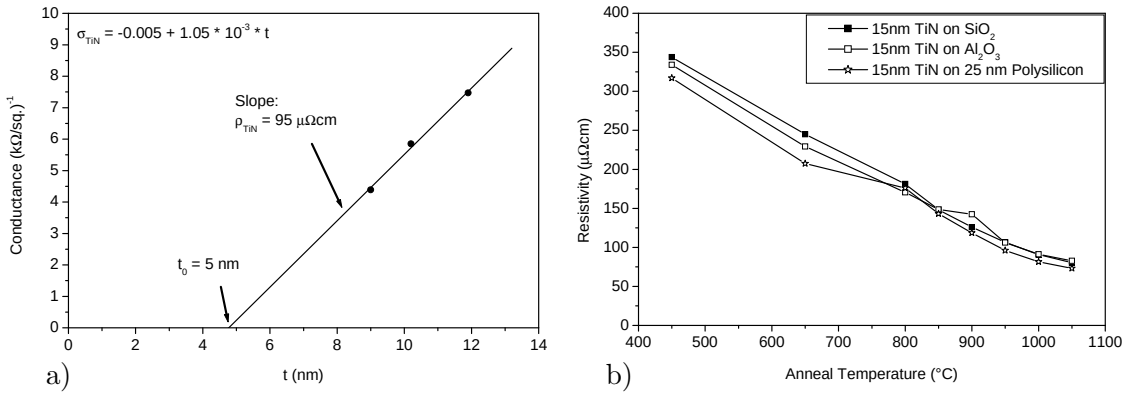

Fig. 6.24: Conductance of ALD-TiN layers as a function of thickness (Panel a). Specific resistivity of a 15 nm ALD-TiN layer as a function of annealing temperature for different substrates (Panel b).

Quantity	Wafer Center	Wafer Edge
Titanium (cm^{-2})	$2.14 \cdot 10^{17}$	$2.12 \cdot 10^{17}$
N:Ti Ratio	1.1:1	1.1:1
Cl:Ti Ratio	0.008:1	0.016:1

Table 6.3: Composition of as-deposited ALD-TiN as determined by RBS-measurements at the wafer center and the wafer edge.

No significant differences are observed for cases where SiO_2, Al_2O_3 or polysilicon have been used as substrates. At the same time, the layer remains intact even after a 1050 °C anneal for 60 s in nitrogen. Degradation is not observed because the as-deposited composition measured by RBS is very close to the thermal stable nitrogen to tungsten ratio of 1:1 (Table 6.3). These good properties of the ALD-TiN facilitate the fabrication of polysilicon/metal electrodes for deep trench capacitors as described in the next section.

6.4.2 Metal Electrodes in Deep Trench Short Loops

The simplest approach to evaluate electrodes for deep trench applications are so-called **D**eep **T**rench **S**hort **L**oops (DTSLs - compare Chapter 3). One main advantage of metal deep trench fills compared to polysilicon is the low resistance. On the other hand it eases the integration process if the well-known properties of polysilicon as electrode could be used further. Hence, a stack of polysilicon and metal inside the DT seems to be most promising. Fig. 6.25 shows SEM-images of such a stack deposited into deep trenches after a 850 °C anneal for 60 s in nitrogen. Similar images are gained for structures annealed at 1050 °C.

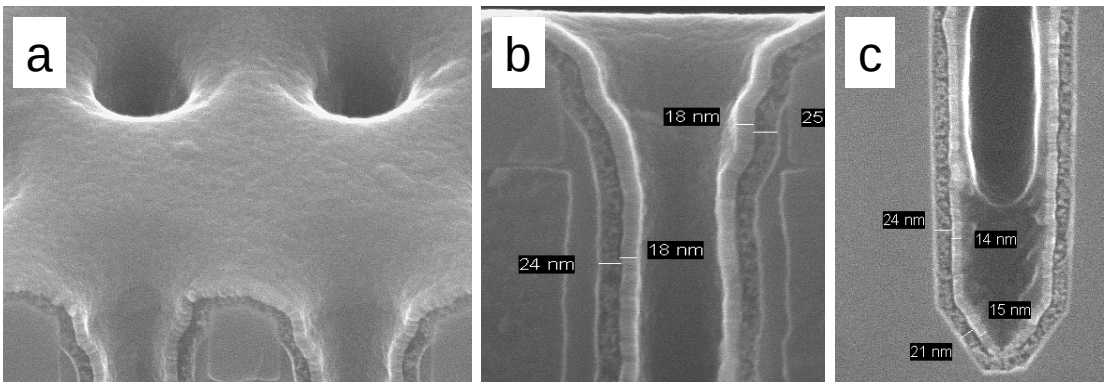

Fig. 6.25: SEM-images of a polysilicon/metal-stack deposited into deep trenches after a 850 °C anneal for 60 s in nitrogen. Panel a) shows an angled view of the wafer surface, Panel b) a cross section at the upper part of the trench and Panel c) a cross section at the bottom.

A smooth and thermally stable layer is observed in all parts of the trench. For electrical characterization, a special DTSL has been developed that enables to do all high-temperature processes before the deposition of the TiN. By this, capacitance and leakage currents can be studied for a large range of applied thermal budgets. Capacitors with a highly doped substrate electrodes were prepared as described in Fig. 6.26. Subsequently, the top electrode was structured in a way in which roughly 10^6 trenches were switched in parallel. A sample with a conventional pure polysilicon-electrode has been processed for comparison. No significant changes in capacitance and leakage current were observed even at very high temperatures. In a second set of sample, temperatures were raised to 1050 °C still leading to no degradation. This thermal budget is sufficient to fully integrate this material stack into a DRAM cell.

6.5 Summary

Metal substrate- and gate-electrodes have been investigated in this chapter. A tungsten-silicide substrate-electrode has been developed which supports the defect healing mechanism of the standard NO and at the same time increases the capacitance and reduces the leakage current. Different gate-electrode materials have been tested and ALD-TiN has

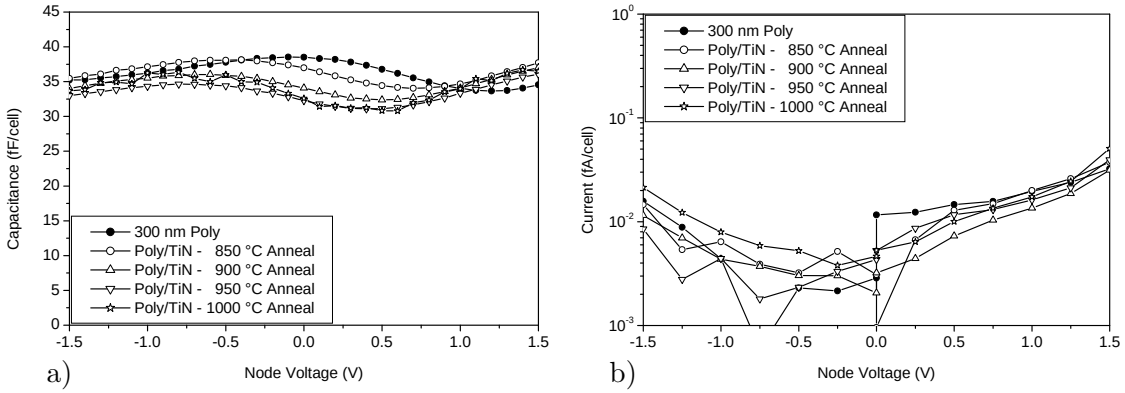

Fig. 6.26: CV- and IV characteristics of deep trench short loop capacitors where roughly 10^6 trenches are switched in parallel. Capacitors have a highly doped substrate-electrode, NO as dielectric and a stack of 25 nm polysilicon and 15 nm TiN as top-electrode and were annealed at temperatures ranging from 850 °C to 1000 °C for 60 s in nitrogen.

been identified as a stable metal on thin oxides as well as on polysilicon. The characterization procedure developed in Chapter 2 was used on TiN-gate MOS-structures and revealed severe problems with the TiN-deposition process which were not seen directly by other means. Wafer signatures of extracted data were compared to marks on the deposition tool and enabled the identification of the core problem. This could finally be solved by modifying the deposition process enabling the fabrication of deep-trench short loops with polysilicon/TiN-electrodes suitable for deep-trench DRAMs.

Chapter 7

Conclusion

The main result of this study is the development of a full analysis procedure for metal-gate MOS-structures that allows to identify problems during processing and to extract physical parameters of metal electrodes. With knowledge gained from this analysis a polysilicon/TiN-stack has been developed and successfully integrated as low-resistance electrode into state-of-the-art deep trench DRAM-capacitors. These electrodes will be required to fabricate sub-100 nm deep trench DRAMs. The most important results can be summarized as follows:

7.1 Metal-Gate MOS-Structures

Parameters were gained from device simulations of MOS-capacitors that allow for an automatic extraction of flatband potential and physical oxide thickness. Physical oxide thicknesses were extracted from NMOS and PMOS devices with polysilicon and metal-gate electrodes within an accuracy of 1-2 Å when compared to ellipsometric measurements and IV-analysis. An improved model has been presented for the gate-leakage current describing the measurement data accurately for all voltages and for all gate oxide thicknesses under study.

To investigate these effects, a new test chip has been developed which enables the fabrication of metal-gate MOS-structures on short loops as well as on fully-integrated wafers. This test chip is now regularly employed to study extrinsic reliability properties of new gate dielectrics and electrodes on large-area capacitors. Intrinsic properties are investigated best on small-area structures. Therefore, a new measurement technique has been invented which enables capacitance-voltage measurements at very low capacitance levels.

A process technology has been developed which facilitates the fabrication of metal-gate MOS-structures even in the case that no dry etch process selective to the gate-oxide is available. Encountered problems like XP-oxide-thinning or gate-erosion during hard-mask removal were systematically solved and finally a method found out to analyze different metals as gate electrodes. While the deposition process of tungsten silicide leads to severe gate-oxide degradation, ALD-TiN was stable up to 800 °C on SiO_2. The TiN-gate structures were then successfully used to characterize the TiN-ALD process, to identify process problems and to eliminate them.

7.2 Metal Electrodes for DRAM-Capacitors

Thermal stability, a clean interface and the ability to support the defect-healing process have been identified as the basic requirements for metal substrate-electrodes of DRAM-capacitors. Using this knowledge, a WSi_x substrate-electrode has been developed that enables reoxidation of the standard NO ensuring at the same time a higher capacitance and a lower leakage current.

The thermal stability of pure metal gate-electrodes has been found to be insufficient for deep-trench DRAM-applications. A stack of 25 nm polysilicon and 20 nm TiN has been identified as a suitable low-resistance deep trench fill that is thermally stable up to 1050 °C. To integrate this stack into DRAMs, the metal has to be deposited into deep trenches with a good step coverage. A tungsten silicide chemical vapor deposition process has been investigated with the result that the major challenge was identified as the simultaneous reaching of good step-coverage and a thermally stable composition. A regime for such a deposition process has been proposed which, however, was beyond the parameter space available with the tool employed in this study. On the other hand, a TiN atomic layer deposition process was successfully modified to give 70% step coverage in trenches with an aspect ratio of 40:1. Using this process, deep trench short loops were fabricated that withstood a 60 s anneal at 1050 °C while reaching the target values for the maximum allowed leakage current and the minimum capacitance.

7.3 Future Aspects

Results presented here enable the fabrication of low-resistance deep trench fills for deep trench DRAM-capacitors. According to the ITRS roadmap, these polysilicon/metal-electrodes will be required from the 70 nm generation onwards which is anticipated to be produced starting in the year 2005. Integration into a full process flow is under development and even now that the feasibility has been shown, one major challenge will be to find a metal deposition process with good step coverage and high throughput.

Bibliography

[1] M. O. Andersson, A. Lundgren and P. Lundgren. Surface-potential dependence of interface-state passivation in metal-tunnel-oxide-silicon diodes. *Phys. Rev. B 50*, 16 (1994), p. 11666–11676.

[2] Y. Ando and T. Itoh. Calculation of transmission tunneling current across arbitrary potential barriers. *J. Appl. Phys. 61*, 4 (1987), p. 1497–1502.

[3] V. Anitha, A. Bhattacharya, N. Patil and S. Major. Study of sputtered molybdenum nitride as a diffusion barrier. *Thin Solid Films 236* (1993), p. 306–310.

[4] A. Avellan, W. Krautschneider and S. Schwantes. Observation and modeling of random telegraph signals in gate and drain current of tunneling metal-oxide-semiconductor field effect transistors. *Appl. Phys. Lett. 78* (2001), p. 2790–2792.

[5] A. Avellan, W. Krautschneider and B. Sell. Softbreakdown and RTS in gate and drain currents of MOS transistors with ultrathin oxides. *16th International Conference on Noise in Physical Systems and 1/f Fluctuations* (Gainesville, USA, 2001), p. 375 ff.

[6] A. Avellan and B. Sell. Signal analysis during charge pumping measurements of MOSFETs. Tech. rep., TUHH, 2001.

[7] R. Baunach. *Elektrische und morphologische Untersuchungen an sehr dünnen Dielektrika für dreidimensionale Si-Strukturen.* Dissertation, Bayerische Julius-Maximilian-Universität, 1988.

[8] G. V. d. bosch, G. V. Groeseneken, P. Heremans and H. E. Maes. Spectroscopic charge pumping: A new procedure for measuring interface trap distributions on MOS transistors. *IEEE Trans. Electron. Dev. 38* (1991), p. 1820–1831.

[9] C. Bowen, C. L. Fernando, G. Klimeck, A. Chatterjee, D. Blanks, R. Lake, J. Hu, J. Davis, M. Kulkarni, S. Hattangady and I. Chen. Physical oxide thickness extraction and verification using quantum mechanical simulation. *Technical Digest of the 1997 International Electron Device Meeting (IEDM)* (1997), p. 869–872.

[10] W. L. Brown. n-type surface conductivity on p-type germanium. *Phys. Rev. 91* (1953), p. 518–537.

[11] J. S. Brugler and P. G. A. Jespers. Charge pumping in MOS devices. *IEEE Trans. Electron. Dev. ED-16*, 3 (1969), p. 297–302.

[12] D. A. Buchanan, F. R. McFeely and J. J. Yurkas. Fabrication of midgap metal gates comparable with ultrathin dielectrics. *Appl. Phys. Lett. 73*, 12 (1998), p. 1676–1678.

[13] J. Byun, B. Lee, J. Park, D. Sohn, S. Choi and J. Kim. Formation of high conductivity WSi_x layer and its characterization as a gate electrode. *J. Electrochem. Soc. 145*, 9 (1998), p. 3228–3235.

[14] T. S. Cale. EVOLVE, developed at Arizona State University and Motorola Inc.

[15] T. S. Cale, T. H. Gandy, M. K. Jain, G. B. Raupp, M. Govil and A. Hasper. Comparison of calculated and experimental film profiles for blanket tungsten LPCVD. *Advanced Metallization for ULSI Applications* (Murray Hill, NJ, USA and Tokyo, Japan, 1992), Materials Research Society, p. 101–104.

[16] T. S. Cale, J. H. Park, G. B. Raupp, M. K. Jain and B. R. Rogers. The inherently transient nature of deposition processes. *Advanced Metallization for ULSI Applications* (Murray Hill, NJ, USA and Tokyo, Japan, 1992), Materials Research Society, p. 93–99.

[17] Y. K. Chae, Y. Egashira, Y. Shimogaki, K. Sugawara and H. Komiyama. Experimental and numerical analysis of rapid reaction to initiate the radical chain reactions in WSix CVD. *Thin Solid Films 320* (1998), p. 151–158.

[18] K.-M. Chang, I.-C. Deng, C.-W. Shih, K.-D. Lain and C.-M. Fu. The characteristics of chemical vapor deposited amorphous-like tungsten film as a gate electrode. *Jpn. J. Appl. Phys. 37*, I 9A (1998), p. 4933–4937.

[19] J. C. Chen, D. Sylvester and C. Hu. An on-chip, interconnect capacitance characterization method with sub-femto-farad resolution. *IEEE Trans. Semicond. Manufact. 11*, 2 (1998), p. 204–210.

[20] J. R. Chen, Y. K. Fang and S. L. Hsu. Growth of tungsten silicide films by low pressure CVD method. *Vacuum 37*, 3/4 (1987), p. 357–361.

[21] W. L. Chen and T. P. Ma. A new technique for measuring lateral distribution of oxide charge and interface traps near MOSFET junctions. *IEEE Electron. Dev. Lett. 12* (1991), p. 393–395.

[22] C.-H. Choi, J.-S. Goo, T.-Y. Oh, Z. Yu, R. W. Dutton, A. Bayoumi, M. Cao, P. V. Voorde and D. Vook. C-V and gate tunneling current characterization of ultra-thin gate oxide MOS ($t_{ox} = 1.3 - 1.8nm$). *Proceedings of the 1999 Symposium on VLSI Technology* (1999), p. 63–64.

[23] B. Claflin, M. Binger and G. Lucovsky. Interface studies of tungsten nitride and titanium nitride composite metal gate electrodes with thin dielectric layers. *J. Vac. Sci. Technol. A 16*, 3 (1998), p. 1757–1761.

[24] T. E. Clark. Characterization and optimization of low-pressure chemical vapor deposited tungsten silicide using screening and response surface experimental designs. *J. Vac. Sci. Technol. B 6*, 6 (1988), p. 1678–1687.

[25] M. Day, M. Delfino, J. Fair and W. Tsai. Correlation of electrical resistivity and grain size in sputtered titanium films. *Thin Solid Films 254* (1995), p. 285–290.

[26] M. v. Dort, P. Woerlee and A. Walker. A simple model for quantization effects in heavily-doped silicon MOSFETs at inversion condition. *Solid-State Electronics 37*, 3 (1994), p. 411–414.

[27] Y. Egashira, H. Aita, T. Saito, Y. Shimogaki, H. Komiyama and K. Sugawara. Modeling and simulation of blanket chemical vapor deposition of WSi_x from WF_6/Si_2H_6. *Electron. Communic. Jap., Part 2 79*, 1 (1996), p. 83–92.

[28] A. B. Elliot. The use of charge pumping currents to measure surface state densities in MOS transistors. *Sol.-Stat. Electron. 19* (1976), p. 241–247.

[29] W. Fikry, M. Ragheb and H. Haddara. Adaptation of drain current charge pumping technique for interface trap characterization in short channel MOS transistors. *Proceedings of the 2000 European Solid State Devices Research Conference (ESSDERC)* (2000), p. 316–319.

[30] D. M. Fleetwood. Fast and slow border traps in MOS devices. *IEEE Trans. Nucl. Sci. 43* (1996), p. 779–786.

[31] R. H. Fowler and L. Nordheim. Electron emission in intense electric fields. *Proceedings of the Royal Society Series A* (London, 1928), Volume 119, p. 173–181.

[32] J. Frenkel. *Phys. Rev. 54* (1938), p. 647.

[33] G. Groeseneken, H. E. Maes, N. Beltran and R. F. D. Keersmaecker. A reliable approach to charge-pumping measurements in MOS transistors. *IEEE Trans. Electron. Dev. ED-31*, 1 (1984), p. 42–53.

[34] M. Gutsche, H. Seidl, J. Lützen, A. Birner, T. Hecht, S. Jakschik, M. Kerber, M. Leonhardt, P. Moll, T. Pompl, H. Reisinger, S. Rongen, A. Sänger, U. Schröder, B. Sell, A. Wahl and D. Schumann. Capacitance enhancement techniques for sub-100nm trench DRAMs. *International Electron Device Meeting (IEDM)* (2001), p. .

[35] H. Haddara. Drain current charge pumping: A new tequnique for interface trap characterization in VLSI MOS transistors. *National Radio Science Conference, NRSC'99* (Cairo, Egypt, 1999), p. D5/1–D5/7.

[36] F. Hamelmann, S. Petri, A. Klipp, G. Haindl, J. Hartwich, L. Dreeskornfeld, U. Kleineberg, P. Jutzi and U. Heinzmann. W/Si multilayers deposited by hot-filament MOCVD. *Thin Solid Films 338* (1999), p. 70–74.

[37] T. Hara, T. Miyamoto, H. Hagiwara, E. I. Bromley and W. R. Harshbarger. Composition of tungsten silicide films deposited by dichlorosilane reduction of tungsten hexafluoride. *J. Electrochem. Soc. 137*, 9 (1990), p. 2955–2958.

[38] T. W. Hickmott and R. D. Isaac. Barrier heights and polycrystalline silicon-SiO_2 interface. *J. Appl. Phys. 52* (1981), p. 3464–3475.

[39] T. Hori. *Gate Dielectrics and MOS ULSIs.* Electronics and Photonics 34. Springer, Berlin, 1997.

[40] K. Hung, K. Ko, C. Hu and Y. Cheng. An automated system for measurement of random telegraph noise in metal-oxide-semiconductor field-effect transistors. *IEEE Trans. Electron. Dev. 36* (1989), p. 1217–1219.

[41] P. Ireland. High aspect ratio contacts: A review of the current tungsten plug process. *Thin Solid Films 304* (1997), p. 1–12.

[42] S. Iwata, N. Yamamoto, N. Kobayashi, T. Terada and T. Mizutani. A new tungsten gate process for VLSI appications. *IEEE Trans. Electron. Dev. ED-31*, 9 (1984), p. 1174–1179.

[43] D. B. Kao, K. C. Saraswat and J. P. McVittel. Annealing of oxide fixed charges in scaled polysilicon gate MOS structures. *IEEE Trans. Electron. Dev. ED-32* (1985), p. 918–925.

[44] L. H. Kaplan and F. M. D'Heurle. The deposition of molybdenum and tungsten films from vapor decomposition of carbonyls. *J. Electrochem. Soc. 117* (1970), p. 693–700.

[45] N. Kasai, N. Endo and A. Ishitani. Deep-submicron tungsten gate CMOS technology. *International Electron Device Meeting (IEDM)* (San Francisco, CA, USA, 1998), IEEE, p. 242–245.

[46] J. Kelsey, C. Goldberg, G. Nuesca, G. Peterson and A. Kaloyeros. Low temperatur metal-organic chemical vapor deposition of tungsten nitride as diffusion barrier for copper metallization. *J. Vac. Sci. Technol. B 17*, 3 (1999), p. 1101–1104.

[47] H. Kim, S. Lee, S. Lee, I. Yeo and S. Lee. A gate electrode fabrication technique using dichlorosilane-based W-polycide with monosilane-based WSi_x nucleation layer. *Electrochem. Sol. Stat. Lett. 2*, 2 (1999), p. 88–90.

[48] K. Kim, J. Sone, S. Kim, J. Park and H. Kim. Deposition mechanism and electrical properties of low pressure chemically vapor deposited W as a gate electrode. *J. Vac. Sci. Technol. A 14*, 3 (1996), p. 919–923.

[49] K. Kimbara. *Gaseous combustion theory.* Shokabo Press, 1985, p. 120 ff.

[50] N. Kobayashi, N. Hara, S. Iwata and N. Yamamoto. Non-selective tungsten CVD technology for gate electrodes and interconnects. *Third International Proceedings of IEEE, VLSI Multilevel Interconnection* (Santa Clara, USA, 1986), p. 436–442.

[51] K. S. Krisch, J. D. Bude and L. Manchanda. Gate capacitance attenuation in MOS devices with thin gate dielectris. *IEEE Electron. Dev. Lett. 17*, 11 (1996), p. 521–524.

[52] F. A. Kuznetsov, V. A. Titov, A. N. Golubenko and A. A. Titov. Thermodynamic simulation of deposition of molybdenum and tungsten disilicide in MOCVD processes. *Proceedings of SPIE* (Warsaw, Poland, 1992), Volume 1783, The International Society for Optical Engineering, p. 541–550.

[53] K. Lai and H. Lamb. Tungsten chemical vapor deposition using tungsten hexacarbonyl: microstructure of as-deposited and annealed films. *Thin Solid Films 370* (2000), p. 114–121.

[54] D. Lee, S. Joo, G. Lee, J. Moon, T. Shim and J. Lee. Characteristics of CMOSFETs with sputter-deposited W/TiN stack gates. *Proceedings of the 1995 International Symposium on VLSI Technology* (Kyoto, Japan, 1995), p. 119–120.

[55] D. Lee, K. Yeom, M. Cho, N. Kang and T. Shim. Gate oxide integrity (GOI) of MOS transiostors with W/TiN stacked gate. *Proceedings of the 1996 International Symposium on VLSI Technology* (1996), p. 208–209.

[56] J.-G. Lee, J.-Y. Kim, J.-Y. Lee, J.-S. Roh and J.-S. Huh. P and As implantation enhanced formation of metal-free oxide on WSi_2. *Jap. J. Appl. Phys. 36* (1997), p. 7140–7145.

[57] W.-C. Lee and C. Hu. Modeling gate and substrate currents due to conduction- and valence-band electron and hole tunneling. *Symposium on VLSI Technology* (2000), p. 198–199.

[58] W.-C. Lee, T.-J. King and C. Hu. Evidence of hole direct tunneling through ultrathin gate oxide using p+ poly-sige gate. *IEEE El. Dev. Lett. 20*, 6 (1999), p. 268–270.

[59] N. Lifshitz. Dependence of the work-function difference between the polysilicon gate and silicon substrate and the doping level of polysilicon. *IEEE Trans. Electron. Dev. ED-32* (1985), p. 617–621.

[60] J. Lin, A. Tsukune, T. Suzuki and M. Yamada. Different effect of annealing temperature on resistivity for stoichiometric, W rich, and N rich tungsten nitride films. *J. Vac. Sci. Technol. A 17*, 3 (1999), p. 936–938.

[61] W. Liu, X. Jin, Y. King and C. Hu. An efficient and accurate compact model for thin SiO_2 on silicon. *Sol.-Stat. Electron. 35*, 6 (1999), p. 797.

[62] S.-H. Lo, D. A. Buchanan and Y. Taur. Modeling and characterization of quantization, polysilicon depletion, and direct tunneling effects in MOSFETs with ultrathin oxides. *IBM J. Res. Develop. 43*, 3 (1999), p. 327–337.

[63] S.-H. Lo, D. A. Buchanan, Y. Taur and W. Wang. Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFETs. *IEEE Electron. Dev. Lett. 18*, 5 (1997), p. 209–211.

[64] Q. Lu, Y. Yeo, P. Ranade, H. Takeuchi, T.-J. King, C. Hu, S. Song, H. Luan and D.-L. Kwong. Dual-metal gate technology for deep-submicron CMOS transistors. *Symposium on VLSI Technology* (2000), p. 72–73.

[65] Y. Lu, R. Weng, W. Hwang and Y. Yang. Study of phase transition and electrical resistivity of tantalum nitride films prepared by DC magnetron sputtering with OES detection system. *Thin Solid Films 389-399* (2001), p. 356–360.

[66] P. Lundgren. Impact of the gate material on the interface state density of metal-oxide-silicon devices with an ultrathin oxide layer. *J. Appl. Phys. 85*, 4 (1999), p. 2229–2232.

[67] J. Lützen, A. Birner, M. Goldbach, T. Hecht, S. Jakschik, A. Orth, A. Sänger, U. Schröder, H. Seidl, B. Sell and D. Schumann. Integration of capacitor for sub-100-nm DRAM trench technology. *Proceedings of the 2002 Symposium on VLSI Technology* (2002), p. .

[68] B. Maiti, P. J. Tobin, C. Hobbs, R. I. Hedge, F. Huang, D. L. O'Meara, D. Jovanovic, M. Mendicino, J. Chen, D. Connelly, O. Adetutu, J. Mogab, J. Candelaria and L. B. La. PVD TiN metal gate MOSFETs on bulk silicon and fully depleted silicon-on-insulator (fdsoi) substrates for deep sub-quarter micron CMOS technology. *International Electron Device Meeting (IEDM)* (San Francisco, CA, 1998), p. .

[69] P. Masson, J.-L. Autran and J. Brini. On the tunneling component of charge pumping current in ultrathin gate oxide MOSFETs. *IEEE Electron. Dev. Lett. 20*, 2 (1999), p. 92–94.

[70] H. Matsuhashi, S. Nishikawa and S. Ohno. Effect of w stress on w-gate mos characteristics. *Extrended Abstracts of the 21st Conference on Solid State Devices and Materials* (Tokio, Japan, 1989), p. 17–20.

[71] C. C. McAndrew and P. A. Layman. MOSFET effective channel length, threshold voltage, and series resistance determination by robust optimization. *IEEE Trans. Electron. Dev. 39*, 10 (1992), p. 2298–2311.

[72] J. Molarius, S. Franssila, G. Drozdy and J. Saarilahti. Tungsten silicide formation from sequentially sputtered tungsten and silicon films. *Appl. Surf. Sci. 53* (1991), p. 383–390.

[73] K. K. Ng and J. R. Brews. Measuring the effective channel length of MOSFETs. *IEEE Circ. Dev. 6* (1990), p. 33–83.

[74] E. H. Nicollian and J. R. Brews. *MOS (Metal Oxide Semiconductor) Physics and Technology.* John Wiley & Sons, New York, 1982.

[75] A. Pacelli, A. Spinelli and L. Perron. Carrier quantization at flat bands on MOS devices. *IEEE Trans. El. Dev. 46*, 2 (1999), p. 383–387.

[76] D.-G. Park, K.-Y. Lim, H.-J. Cho, T.-H. Cha, I.-S. Yeo, J.-S. Roh and J. Park. Impact of atomic-layer -deposited TiN on the gate oxide quality of W/TiN/SiO2/Si metal-oxide-semiconductor structures. *Appl. Phys. Lett. 80*, 14 (2002), p. 2514–2516.

[77] T. Pompl. *Gateisolatoren für MOS-Feldeffekttransistoren.* Dissertation, Universität der Bundeswehr München, 2000.

[78] H. H. Poole. *Phil. Mag. S. 6.32*, 187 (1916), p. 112.

[79] L.-A. Ragnarsson, P. Lundgren, Z. Ovuka and M. O. Andersson. Oxide thickness- and bias-dependence of postmetallization annealing of interface states in metal-oxide-silicon diodes. *J. Electrochem. Soc. 144*, 5 (1997), p. 1866–1869.

[80] G. B. Raupp, T. S. Cale, M. K. Jain, B. Rogers and D. Srinivas. Step coverage of tungsten silicide films deposited by low pressure dichlorosilane reduction of tungsten hexafluoride. *Thin Solid Films 193/194* (1990), p. 234–243.

[81] H. Reisinger, H. Oppolzer and W. Hönlein. Thickness determination of thin SiO_2 on silicon. *Sol.-Stat. Electron. 35*, 6 (1992), p. 797–803.

[82] H. Reisinger, H. Wendt, G. Beitel and E. Fritsch. Dielectric breakdown, reliability and defect density of $(Ba_{0.7}Sr_{0.3})TiO_3$ (BST). *Proceedings of the 1998 Symposium on VLSI Technology* (1998), p. 58–59.

[83] B. Ricco, P. Olivo, T. N. Nguyen, T.-S. Kuan and G. Ferriani. Oxide-thickness determination in thin-insulator MOS structures. *IEEE Trans. Electron. Dev. 35*, 4 (1988), p. 432–438.

[84] R. Rios and N. D. Arora. Determination of ultra-thin gate oxide thickness for CMOS structures using quantum effects. *Technical Digest of the 1994 International Electron Device Meeting (IEDM)* (1994), p. 613–616.

[85] J. Robertson and M. J. Powell. *Appl. Phys. Lett. 44*, 4 (1984), p. 415.

[86] E. J. Rode and W. R. Harshbarger. Thermodynamic analysis of the W–F–Si–H and W–F–Si–H–Cl systems. *J. Vac. Sci. Technol. B 8*, 1 (1990), p. 91–94.

[87] K. Roh, S. Youn, S. Yang and Y. Roh. Tungsten silicide for the alternate gate metal in metal-oxide-semiconductor devices. *J. Vac. Sci. Technol. A 19*, 4 (2001), p. 1562–1565.

[88] E. Rosenbaum and L. F. Register. Mechanism of stress-induced leakage current in MOS capacitors. *IEEE Trans. El. Dev. 44*, 2 (1997), p. 317–323.

[89] A. L. P. Rotondaro, H. Yang, J. C. Hu, R. Kraft, G. A. Brown and C. Appel. Dry process for the definition of sub-0.1μm W/TiN gates. *Electrochem. Soc. Proc.* (1997), Volume 97, p. 344–349.

[90] T. Saito, Y. Shimogaki, Y. Egashira, H. Komiyama, K. Sugawara, K. Takahiro, S. Nagata and S. Yamaguchi. Kinetic study of WSi_x-CVD processes - a comparison of WF_6/SiH_4 and WF_6/Si_2H_6 reaction systems. *Electron. Communic. Jap., Part 2 78*, 10 (1995), p. 73–84.

[91] K. C. Saraswat, D. L. Brors, J. A. Fair, K. A. Monnig and R. Beyers. Properties of low-pressure CVD tungsten silicide for MOS VLSI interconnections. *IEEE Trans. Electron. Dev. ED-30* (1983), p. 1497–1505.

[92] J. E. J. Schmitz, W. L. N. v. d. Sluys and A. H. Montree. Comparison of calculated and experimental step coverage of the H_2 / WF_6 and SiH_4 / WF_6 chemistries used in the blanket deposition of tungsten. *Tungsten and Other Advanced Metals for VLSI/ULSI Applications V* (1990), MRS, p. 117–126.

[93] D. K. Schroder. *Semiconductor Material and Device Characterization.* John Wiley & Sons, New York, 1998.

[94] B. Sell, A. Avellan and W. Krautschneider. Charge based capacitance measurements on mos devices. *Accepted for publication in IEEE Trans. Dev. Mat. Reliab.* (2001), p. .

[95] B. Sell, A. Sänger, G. Schulze-Icking, K. Pomplun and W. Krautschneider. Chemical vapor deposition (CVD) of tungsten silicide (WSi_x) in high aspect ratio trenches. *Submitted to Thin Solid Films* (2001), p. .

[96] B. Sell, D. Schumann and W. Krautschneider. Fast interface characterization of tunnel oxide MOS structures. *Submitted to IEEE Trans. Nanotechnology* (2001), p. .

[97] B. Sell, J. Willer, K. Pomplun, A. Sänger, D. Schumann and W. Krautschneider. Interface characteristics between tungsten silicide electrodes and thin dielectrics. *Microeletronic Engineering 55* (2001), p. 197–203.

[98] J. A. Sethian. Level set methods and fast marching methods. *Cambridge Monographs on Applied and Computational Mathematics.* ISBN 0-521-64204-3, p. .

[99] H. Shimada, I. Ohshima, T. Ushiki, S. Sugawa and T. Ohmi. Tantalum nitride metal gate FD-SOI CMOS FETs using low resistivity self-grown bcc-tantalum layer. *IEEE Trans. El. Dev. 48*, 8 (2001), p. 1619–1626.

[100] Y. Shimogaki, T. Saito, R. Ikawa, Y. Egashira, K. Sugawara and H. Komiyama. How to control the silicon contents in the chemically vapor deposited WSi_x films. *Eleventh International VLSI Multilevel Conference (VMIC)* (1994), p. 496–498.

[101] Y. Shimogaki, T. Saito, F. Tadokoro and H. Komiyama. The reactivity and molecular size of film precursors during chemical vapor deposition of WSi_x. *J. de Phys. IV Colloq. C2, suppl. au J. de Phys. II 1* (1991), p. 95–102.

[102] C. Spee, F. Verbeek, J. Kraaijkamp, J. Linden, T. Rutten, H. Delhaye, E. v. d. Zouwen and H. Meinema. Tungsten deposition by organometallic chemical vapour deposition with organotungsten precursors. *Mat. Sci. Eng. B 17* (1993), p. 108–111.

[103] J. Y.-C. Sun, M. R. Wordeman and S. E. Laux. On the accuracy of channel length characterization of LDD MOSFETs. *IEEE Trans. Electron. Dev. ED-33*, 10 (1986), p. 1556–1562.

[104] S. C. Sun and T. F. Chen. Effects of electrode materials and annealing ambients on the electrical properties of TiO_2 thin films by metalorganic chemical vapor deposition. *Jap. J. Appl. Phys. 36* (1997), p. 1346–1350.

[105] D. Sylvester, J. C. Chen and C. Hu. Investigation of interconnect capacitance characterization using charge-based capacitance measurement (cbcm) technique and three-dimensional simulation. *IEEE J. Sol. Stat. Circuits 33* (1998), p. 449–453.

[106] S. M. Sze. *Physics of Semiconductor Devices*, 2. Ed. John Wiley & Sons, New York, 1981.

[107] Y. Taur and T. H. Ning. *Modern VLSI Devices*, 1. Ed. Cambridge University Press, Cambridge, 1998.

[108] S. G. Telford, M. Eizenberg, M. Chang, A. K. Sinha and T. R. Gow. Chemically vapor deposited tungsten silicide films using dichlorsilane in a single-wafer reactor. *J. Electrochem. Soc. 140*, 12 (1993), p. 3689–3701.

[109] S. Tiwari, F. R. McFeely, D. A. Buchanan and E. Cartier. Tungsten gate using carbonyl chemistry: Electrical properties as MOSFET gates. p. .

[110] N. Volkenshteyn, V. Dyakina, P. Kumin, V. Startsev, R. Nasyrov and V. Manakov. Electrical resistivity of high-purity tantalum in the range 4.2-300 K. *Phys. Met. Metall. 67*, 4 (1989), p. 104–110.

[111] R. M. Wallace and G. D. Wilk. Exploring the limits of gate dielectric scaling. *Semiconductor International June* (2001), p. 153–158.

[112] R. M. Wallace and G. D. Wilk. High-k gate dielectrics. *Semiconductor International July* (2001), p. 227–236.

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

[114] J. Wang, C. Cao, H. Wang and S. Zhang. On the formation of defects and morphology during chemical vapor deposition of tungsten. *J. Electrochem. Soc. 141*, 8 (1994), p. 2192–2198.

[115] Z. A. Weinberg. On tunneling in metal-oxide-silicon structures. *J. Appl. Phys. 53*, 7 (1982), p. 5052–5056.

[116] W. W. Werner. The work function difference of the MOS system with aluminium field plates and polycrystalline silicon field plates. *Solid-State Electron. 17* (1974), p. 769–775.

[117] J. Willer. private communication, 2000.

[118] H. S. Wong, M. H. White, T. J. Krutsick and R. V. Booth. Modeling of transconductance degredation and extraction of threshold voltage in thin oxide MOSFETs. *Sol.-Stat. Electron. 30* (1987), p. 953–968.

[119] T. H. T. Wu, R. S. Rosler, B. C. Lamartine, R. B. Gregory and H. G. Tompkins. Properties of WSi_x using dichlorosilane in a single-wafer system. *J. Vac. Sci. Technol. B 6*, 6 (1988), p. 1707–1713.

[120] N. Yamamoto, H. Kume, S. Iwata, K. Yagi, N. Kobayashi, N. Mori and H. Miyazaki. Fabrication of highly reliable tungsten gate MOS VLSIs. *J. Electrochem. Soc. 133*, 2 (1986), p. 401–407.

[121] H. Yang, G. Brown, J. Hu, J. Lu, R. Kraft, A. Rotondaro, S. Hattangady, I. Chen, J. Luttmer, R. Chapman, P. Chen, H. Tsai, B. Amirhekmat and L. Magel. A comparison of tin processes for cvd w/tin gate electrode on 3nm gate oxide. *Proceedings of the Internation Electron Device Meeting (IEDM)* (Washington, USA, 1997), p. 459–62.

[122] K. Yang, Y.-C. King and C. Hu. Quantum effects in oxide thickness determination from capacitance measurements. *Proceedings of the 1999 Symposium on VLSI Technology* (Kyoto, Japan, 1999), p. 77–78.

[123] W.-K. Yeh, Y.-C. Shiau and M.-C. Chen. A new tungsten gate metal oxide semiconductor capacitor using a chemical vapor deposition process. *J. Electrochem. Soc. 144*, 1 (1997), p. 214–217.

Acknowledgements

First of all I would like to thank Professor Wolfgang Krautschneider for his supervision, his guidance throughout the whole work and for numerous helpful discussions. Many thanks also to Professor Wolfgang Albrecht for carefully prove reading the thesis. It is my great pleasure to thank Alejandro Avellán for fruitfully working together, for many exciting discussions which lead to nice joint publications and for enjoyable times after work.

There are a lot of people from Infineon Technologies I would like to acknowledge for their support. Many thanks to Dr. Josef Willer for creating the PhD project, for his supervision during the initial phase of the project and for interesting discussions. I would like to express my deep gratitude to Dr. Dirk Schumann for building up a fascinating innovation group in Dresden, for being a great team leader that transfers a lot of responsibility to his employees and for many serious squash matches.

I would like to thank all members of the innovation group for the nice and creative atmosphere. In particular I would like to thank Dr. Annette Sänger for carrying out the metal project together and for a lot of metal ALD tuning, Peter Moll for valuable dry etch developments, Dr. Martin Popp for jointly building up the characterization laboratory, Dr. Matthias Goldbach and Dr. Till Schlösser for many fruitful discussions, Steffen Breuer for some nice Matlab programming and all who contributed to this work but are not mentioned in this acknowledgement. During the creation of the test chip I very much appreciated the support from Dr. Martin Kerber, Dr. Thomas Pompl, Mira Rest, Walter Zatsch, Wolfram Ziegler and Dr. Jürgen Holz. In addition I would like to thank all people from the single process and characterization departments in Dresden for their work. In particular I would like to thank Kerstin Pomplun and Dieter Zeiler for many metal deposition experiments as well as Dr. Georg Schulze-Icking for simulations of the WSi_x-CVD process and for strong support during writing of the WSi_x deposition paper.

I would like to express my deep gratitude to my parents Wolfgang and Brigitte and my sister Anne for intense prove reading of the thesis and my whole family for their continuous support. Last but not least I would like to thank all friends for the joyful times after work.

Bernhard Sell

Office Address:
Infineon Technologies Dresden
Königsbrücker Straße 180
01099 Dresden
Germany
Tel.: ++ 49 351 886 7758

Home Address:
Prießnitzstraße 41
01099 Dresden
Germany
Tel.: ++ 49 351 802 36 12
E-mail: Ben.Sell@t-online.de

PERSONAL

Date of birth:	24 September 1972 in Hamburg, Germany
Citizenship:	German

EDUCATION

Ph.D.:	Electrical Engineering, Technical University Hamburg-Harburg; 6/02 Thesis Title: **"Interface Characterization of Metal-Gate MOS-Structures and the Application to DRAM-Capacitors"** Research Advisor: Professor Wolfgang Krautschneider
German Diploma:	Physics (Distinction), University of Hamburg (Germany) and Imperial College, London (UK); 03/98 Thesis Title: **"Electrochemical Capacitance-Voltage Profiling of Si/SiGe Heterostructures**"
Vordiplom:	Physics, University of Hamburg (Germany); 1994

WORK EXPERIENCE

12/98 - present	Infineon Technologies Dresden, DRAM Innovation Group: · Single process development and process integration for implementation of metal electrodes into a deep trench DRAM process. · Development of a testchip for the characterization of new gate dielectrics and gate electrode materials. · Development of electrical characterization methods and building up of a characterization laboratory. · Characterization of high-k materials and metal electrodes for DRAMs. · Training of new employees.
7/98 - 11/98	Siemens Microelectronics Center Dresden, Quality Management: · Reliability analysis of production processes. · Evaluation of productive DRAM lots.

PUBLICATIONS

B. Sell, D. Schumann and W. Krautschneider, **"Fast Interface Characterization of Tunnel Oxide MOS Structures"**, submitted to IEEE Transactions on Nanotechnology.

A. Avellán, E. Miranda, B. Sell, D. Schroeder and W. Krautschneider, **"Temperature dependence of the hard breakdown current of MOS capacitors"**, accepted for ESSDERC 2002.

B. Sell, A. Avellán and W. Krautschneider, **"Charge-Based Capacitance Measurements on MOS Devices"**, accepted for publication in IEEE Transactions on Device and Materials Reliability.

B. Sell, A. Sänger, G. Schulze-Icking, K. Pomplun and W. Krautschneider, **"Chemical Vapor Deposition (CVD) of Tungsten Silicide (WSi_x) in High Aspect Ratio Trenches"**, submitted to Thin Solid Films.

J. Lützen, A. Birner, M. Goldbach, M. Gutsche, T. Hecht, S. Jakschik, A. Orth, A. Sänger, U. Schröder, H. Seidl, B. Sell and D. Schumann, **"Integration of Capacitor for Sub-100 nm DRAM Trench Technology"**, accepted for Symposium on VLSI Technology 2002.

M. Gutsche, H. Seidl, J. Lützen, A. Birner, T. Hecht, S. Jakschik, M. Kerber, M. Leonhardt, P. Moll, T. Pompl, H. Reisinger, S. Rongen, A. Sänger, U. Schröder, B. Sell, A. Wahl and D. Schumann, **"Capacitance Enhancement Techniques for Sub-100nm Trench DRAMs"**, International Electron Device Meeting (IEDM), December 2001.

A. Avellán, W. Krautschneider and B. Sell, **"Softbreakdown and RTS in Gate and Drain Currents of MOS Transistors with Ultrathin Oxides"**, 16th International Conference on Noise in Physical Systems and 1/f Fluctuations (Gainesville, USA, 2001), p. 375-378.

B. Sell, J. Willer, K. Pomplun, A. Sänger, D. Schumann and W. Krautschneider, **"Interface Characteristics between Tungsten Silicide Electrodes and Thin Dielectrics"**, Microelectronic Engineering 55 (2001), p. 197-203.

B. Sell, C. Gatzke, J. M. Fernández, **"Electrochemical Capacitance-Voltage Profiling of Heterostructures using Small Contact Areas"**, Semicond. Sci. Technol. 13 (1998), p. 423-427.

PATENTS

3 granted patents with Infineon Technologies (DRAM storage cells).
25 pending patent applications with Infineon Technologies in the field of DRAM process technology.

REFERENCES

Dr. Dirk Schumann

Former Manager of DRAM Innovation Group
Infineon Technologies Dresden
Now with:
u2t photonics AG
Reuchlinstrasse 10/11
10553 Berlin
Germany

Tel.: 0049 30 726 113 531
E-mail: schumann@u2t.de

Prof. Wolfgang Krautschneider

Technical University of Hamburg-Harburg
Electrical Engineering, Microelectronics
Eißendorfer Straße 38
21073 Hamburg

Tel.: 0049 40 42878 3030
E-mail: krautschneider@tu-harburg.de