Impact of 14/28nm FDSOI high-k metal gate stack processes on reliability and electrostatic control through combined electrical and physicochemical characterization techniques

Pushpendra Kumar

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Nano électronique et nano technologie**

Arrêté ministériel : 25 mai 2016

Présentée par

Pushpendra KUMAR

Thèse dirigée par **Gérard GHIBAUDO**
et codirigée par **Charles LEROUX et Florian DOMENGIE**

préparée au sein **de la société STMicroelectronics, du laboratoire CEA-LETI (Grenoble) et de l'IMEP-LAHC**
dans **l'École Doctorale EEATS**

Impact of 14/28nm FDSOI high-k metal gate stack processes on reliability and electrostatic control through combined electrical and physicochemical characterization techniques

Thèse soutenue publiquement le **19 décembre 2018**,
devant le jury composé de :

Mme. Mireille MOUIS
DR, CNRS-Grenoble, Président
Mme. Nathalie Labat
PR, laboratoire IMS, Rapporteur
M. Brice GAUTIER
PR, INSA-Lyon, Rapporteur
M. Charles Leroux
ING, CEA-LETI, Examinateur
M. Florian Domengie
ING, STMicroelectronics, Examinateur
Mr Gérard Ghibaudo
DR, CNRS Alpes, Directeur de thèse

Acknowledgements

I would like to thank all the people and the various teams who have contributed to the work presented in this thesis.

First, I would like to thank my supervisors Charles Leroux from CEA-LETI, Florian Domengie from STMicroelectronics, and my PHD director Gerard Ghibaudo from IMEP. Charles helped in building well-planned experiments and their timely completion, his valuable experience and knowledge of device physics and electrical characterization greatly helped in obtaining new and interesting results. Florian's knowledge and experience in process development helped greatly in device fabrication and building of new process for its performance improvements. Gerard's valuable advices and continuous support helped build a well-planned work and this thesis.

Secondly, I would like to thank the Metal R&D/Engineering/Hardware team of STMicroelectronics, with whom I spent half of my time during the PHD thesis. The team helped in device fabrication and in performing some non-standard process steps which were challenging. I would also like to thank the laboratory of electrical characterization (LCTE) of CEA-LETI, where I spent the other half of my thesis, for its help in electrical characterization and device reliability.

Next, I would like to thank the various process development, process integration and physical characterization teams of STMicroelectronics and CEA-LETI, who helped in some way in performing experiments and analysis for this work: Kristell Courouble (high-k), Fabien Deprat (SiGe), Côme de buttet (WET), Yann Escarabajal (Wet), Olivier Gourhant (DIA), Bilel saidi (TT), Francesco Abbate (XRR), Denis Combe (Metal), Delphine Doloy (XRF), Denis Guiheux (XPS), A.Roland, N.Miguel (TEM), Alexia Valery (ASTAR), Patrice Gergaud (XRD), Eugenie Martinez (XPS), Yves Morand, Jean-Michel pedini, Nils Rambal, Claude Tabone. I would like to specially thank Eugenie Martinez from CEA-LETI for her support and contribution in the development of XPS under bias technique.

Lastly, I would like to thank Carlos Suarez-Segovia and Sylvian Baudot, who worked on the development of the sacrificial gate process using which additives were introduced in the devices used in this work.

Contents

List of Abbreviations

MOSFET: Metal Oxide Semiconductor-Field Effect Transistor

CMOS: Complementary Metal Oxide Semiconductor

FDSOI: Fully Depleted Silicon On Insulator

WFeff : Effective workfunction

VFB: Flatband voltage

HKMG: High-k metal gate

STI: Shallow Trench Isolation

EOT: Equivalent Oxide Thickness

IL: Interfacial Layer

ALD: Atomic Layer Deposition

PVD: Physical Vapor Deposition

CMP: Chemical mechanical polishing

RF-PVD: Radio-Frequency Physical Vapor Deposition

S/D: Source-Drain

DIA: Drive in anneal

C-V: capacitance voltage

MGG: Metal gate granularity

FCC: Face centered cubic

NBTI: Negative Bias temperature instability

PBTI: Positive Bias temperature instability

TDDB: Time Dependent Oxide Breakdown

SILC: Stress Induced Leakage Current

HBD : Hard breakdown.

SBD : Soft breakdown

AHI : Anode hole injection

MVHR: multivibrational hydrogen release

XRD: X-Ray Diffraction

XPS: X-ray Photoelectron Spectroscopy

XRF: X-Ray Fluorescence

FWHM: Full width at half maximum

MPE: Multiphonon emission

TBD: time to breakdown at 63% devices

CL: Core level

VB: Valence band

BE: Binding energy

DUT: Device under test

General introduction

Context

Digital electronic devices have become a necessary part of our daily life, with applications in communications, entertainment, automotive, medical equipment and internet of things. This became possible, and is also described as the second industrial revolution, due to the continuous efforts in research and development to improve the performance of these electronic devices. These devices contain multiple semiconductor chips, each having a specific function. Each of these chips is made of many integrated circuits, containing billions of transistors, which is at the heart of these circuits. The first transistor, called the bipolar transistor (BJT), was developed in 1947 by Bell laboratory researchers John Bardeen, William Shockley and Walter Brattain, for which they received the Nobel prize in physics. Then in 1959, D. Kahng and Mr. Attala invented the first MOSFET (Metal Oxide Semiconductor-Field Effect Transistor). In the beginning, BJT devices were preferred due to their rapid switching speeds, but later MOSFET devices started to dominate the microelectronics industry due to their lower power consumption. Since then, MOSFET's became the key part of the CMOS (Complementary Metal Oxide Semiconductor) technology, used to fabricate logic gates of the integrated circuits.

In order to continually improve electrical performance, reduce power consumption and manufacturing costs of digital devices or circuits, the dimensions of the transistor needs to be reduced. This is known as downscaling, as predicted by the Moore's law in 1965, and involves reducing the transistor and gate oxide dimensions. The performance improvement is due to the increase of transistor channel current, which increases its switching speed. Thus, although the operating principle of the MOSFET is still the same, its size have decreased and its quantity placed on a chip has doubled every two years, as described by the International Semiconductor Technology Roadmap (ITRS). This led to massive technological progress in design and fabrication of integrated circuits, with an ever increasing transistor density, electrical performance and their applications.

However, these technological advances have faced some serious challenges since the 2000's, where the downscaling has resulted in major device physics and process integration issues. Indeed, due to miniaturization of the transistor dimensions and of the gate oxide to the nanometric scale, undesirable effects including some complex quantum phenomena have emerged. These include the high gate leakage current for very thin gate oxides, short channel effects for very small transistors, threshold voltage variability, etc. Increase in the gate leakage current affects transistor reliability and increases its power consumption. Thus, the simple downscaling was no longer sufficient to improve transistor performance.

Therefore, new solutions were proposed to deal with these issues, while still being able to downscale. First was the introduction of the (high-k metal gate) HKMG stack, which replaced the conventional poly-Si/SiO_2 stack. The introduction of the high-k oxide decreases the gate leakage while still keeping the same capacitance density, because a higher oxide thickness can be used.

Moreover, introduction of a metal gate electrode eliminated the polysilicon depletion in gate electrode (thus increasing switching speed), eliminated the Fermi Level pinning with the high-k layer, decreased gate resistance and boron penetration effects. Second is the introduction of the FDSOI (Fully depleted silicon on insulator) substrate, containing an ultra-thin buried oxide. Compared to a bulk substrate, FDSOI reduced the threshold voltage (V_T) variability, short channel effects and gate leakage current.

In this context, optimization and tuning of electrical and process parameters, of the MOSFET devices with HKMG stack on FDSOI substrates, becomes very important. V_T of MOSFET's must be fine-tuned to satisfy certain requirements for specific FD-SOI devices. V_T is mainly dependent on the parameter known as the effective work function (WF_{eff}), which is related to the metal gate work function (WF_M) and the properties of the high-k layer. In order to fine-tune the WF_{eff} of MOSFET's, additives such as Lanthanum (La) and Aluminum (Al) have been introduced inside the HKMG stack by a drive in anneal (DIA) process. However, the impact of these additives on MOSFET reliability is not known. Moreover, even though a limited relationship between diffused dose of these additives and the shift in WF_{eff} has been investigated earlier, a clear understanding and modelling of the diffusion phenomena and effect on the WF_{eff} with the process conditions DIA have not been done. Although some contributions to the V_T variability are dramatically reduced, especially by the introduction of the FD-SOI substrate, studies have pointed out that the contribution due to the microstructure of the metal gate still remains important, and must be investigated. Lastly, in order to build a full understanding of the CMOS devices and for its further development, a comprehensive knowledge of band-energies of different layers in the gate stack is required, which cannot be extracted from simple electrical measurements. Thus, there is a need to develop a non-destructive characterization technique to measure the band energies of the HKMG stack.

Thesis objectives and outline

The main objectives of this thesis are to investigate : 1) the effects of La and Al additives on the BTI (bias temperature instability) and TDDB (time dependent oxide breakdown) reliability, 2) the impact of DIA conditions on the diffusion of these additives and device WF_{eff}, 3) the impact of TiN metal gate deposition conditions on its microstructure in the context of reducing V_T variability, mechanical properties and device WF_{eff}, and 4) development of XPS (X-ray Photoelectron Spectroscopy) under bias technique to determine the band energies of the MOSFET HKMG stack.

The first chapter introduces the gate stack technology for a MOSFET device of the 14 and 28 nm FDSOI technology. The working operation of a MOSFET device, its various performance parameters and how they can be tuned by the various charges, interface states and dipoles present in the gate stack are described. Then we discuss on the importance of the different layers consisting the HKMG stack and their processes, i.e. the interlayer, high-k and the metal gate. Special attention has been paid to the deposition of TiN, La and Al by the RF-PVD method, and their various process parameters have been presented. The introduction of La and Al additives for V_T engineering into the HKMG stack has also been described. Further, we describe the process flow used to manufacture the MOS devices used in this work, as well as its simplification for capacitance devices. An introduction on the various contributions to V_T variability has been given, and special

attention has been paid to the contribution of the metal gate microstructure. Finally, a brief introduction of the NBTI, PBTI and TDDB reliability has been given, which includes a review of state of the art and the various mechanisms and models that exists presently.

The second chapter introduces the various electrical and physicochemical techniques used in this work to characterize the gate stack. The capacitance voltage (C-V) measurement technique, analytical and automatic method to extract flatband voltage (V_{FB}) and equivalent oxide thickness (EOT) from experimental C-V measurements is presented. Moreover, a method to separately extract the charges and dipoles present in the gate stack is presented. Then, the four probe method used to extract sheet resistance of films is described. This chapter also presents the various non-electrical techniques that have been used in this thesis. This includes first the X-ray Photoelectron Spectroscopy (XPS) technique used to obtain the band energies of the gate stack layers, X-Ray Fluorescence (XRF) technique used to measure the additive dose and metal gate elemental composition, and finally the X-Ray Diffraction (XRD) technique used to obtain information on the microstructure of crystalline films by combining the in-plane and out of plane configurations. Lastly, the technique to measure the stress added by the deposition of a film on a Si substrate is presented.

The third chapter is based on the study of the impact of La and Al additives, used for threshold voltage adjustment, on MOSFET BTI and TDDB reliability. This chapter covers first the device fabrication and secondly the test methodology, analysis for NBTI and PBTI V_T shift, and the role of oxide field on the BTI reliability. Then, results on the role of La and Al additives on BTI life-time and time to failure of the device are studied, and the physical interpretation related to the impact of these additives on the HKMG stack is discussed. Next, the impact of these additives on the TDDB reliability is studied. This covers the electrical measurements for TDDB, time to breakdown detection and their representation on the Weibull distribution scale. Further, the methodology to select the gate stress voltage for TDDB tests is described. Lastly, results on the role of La and Al on NMOS and PMOS time to breakdown have been shown and their explanations by defect creation in the oxide are discussed.

The fourth chapter focuses on the effect of TiN metal gate deposition conditions, chamber pressure and RF power, on physical and electrical properties of MOS device and TiN film itself. First, the need for TiN microstructure modulation in the context of lowering the threshold voltage variability has been stated and then the possible RF-PVD process variations that can achieve this modulation in TiN microstructure are discussed. The HKMG stack process and the two different process flow (or devices) that have been used in this study, to measure physical and electrical properties are presented. Then, the different physical and electrical measurements and the results obtained for variations in RF-PVD chamber pressure and RF power are presented. This includes first the grain size calculated by in-plane XRD measurements and relative percentage of grain orientation calculated by out of plane XRD measurements, and then the mechanical stress and sheet resistance results. Further, the effective workfunction versus EOT and Ti/N ratio results are presented, and the correlation between metal workfunction, dipoles and Ti/N ratio is reported. The effect of TiN process conditions on the wafer non-uniformity of grain size and sheet resistance, and the impact of substrate temperature on the grain size and their relative orientations are also presented. Finally, results on the ASTAR technique for TiN microstructure analysis and comparison of its results with XRD are shown.

The fifth chapter investigates the diffusion of La and Al additives inside the HKMG stack, with DIA conditions (temperature and time) of the sacrificial gate process. First, the mechanisms and the fundamentals of the diffusion process are described. The sacrifical gate process and the use of two different process flow (or devices) are also introduced, that have been used to study additve diffusion and effective workfunction. Then, the results on diffused dose of La or Al in the gate stack with DIA time and temperature, measured by X-Ray Fluorescence (XRF) technique are presented. This section highlights the main differences between La and Al additive diffusion and different high-k. Finally, the modulation of effective workfunction or additive dipole by the DIA conditions is presented, followed by the correlation between the modulation of dipole and diffused dose by DIA conditions, for the two additives and high-k materials.

The final chapter focuses on the development and validation of XPS under bias technique to analyze HKMG stack band energies, with the main aim to localize different dipoles. First, the need for this technique is presented and a comparison of its benefits with capacitance measurements is done. The state of the art and the major issue related to the conventional XPS and XPS under bias techniques are also presented. Then, the specific test structures, used to perform XPS under bias measurements, their process flow and the biasing issues occuring in them due to different parasitic serial resistances are discussed. The biasing solutions developed by us using a specific methodology that combines electrical measurements, XPS under bias measurements and modeling on different devices are presented. Further, the various experimental issues, related to XPS equipment and the samples, and their solutions are discussed. This includes the issue of device location inside the XPS equipment, the estimation of the exact size and the position error of the X-ray beam, the impact of the X-ray beam on the devices during XPS measurements and the experimental methods employed to decrease the parasitic bias drops in the Si substrate. Then, XPS under bias technique is validated by first developing a procedure to fit the XPS signals from the different layers of the HKMG gate stack, and then comparing the binding energy values obtained from fitting the XPS signals to the values obtained with our electrical modelling. Finally, XPS under bias technique will be used to localize and quantify dipoles related to the addition of La or Al, and dipoles occuring due to TiN gate thickness modulation. A methodology to use XPS measurements at zero bias for estimating the values at the flatband condition will also be presented.

This CIFRE Ph.D. work has been carried out in collaboration with STMicroelectronics in Crolles, the CEA-LETI and the IMEP-LAHC laboratories in Grenoble. Most of the steps in the devices fabrication was conducted at STMicroelectronics, and some at CEA-LETI. Electrical characterizations were performed in the Laboratoire de Caractérisation et Tests Electriques (LCTE) of the CEA-LETI. Physical characterizations were performed at both CEA-LETI and STMicroelectronics. For XPS under bias studies, the test structures and XPS measurements were done at CEA-LETI and at ST.

1. Gate stack technology for 14/28 nm FDSOI MOSFET devices

This chapter introduces the gate stack technology for a MOSFET device of the 14 and 28 nm FDSOI technology. Section 1.1 describes the working operation of a MOSFET device and its various performance parameters and how they can be tuned by the various charges, interface states and dipoles present in the gate stack. Section 1.2 describes the need for different layers consisting the HKMG stack and their processes, i.e. the interlayer, high-k and the metal gate. Special attention has been paid to the deposition of TiN, La and Al by the RF-PVD method, and their various process parameters have been discussed. Then, the introduction of La and Al for V_T engineering into the HKMG stack has been discussed.

Section 1.3 describes the process flow used to manufacture the MOS devices used in this work, and its simplification for capacitance devices. Section 1.4 presents the various contributions to V_T variability and special attention has been paid to the contribution of the metal gate microstructure to V_T variability. In section 1.5, a brief introduction of the NBTI, PBTI and TDDB reliability has been given. This describes the state of the art and the various mechanisms and models that exists presently.

1.1 MOSFET device

1.1.1 MOSFET operation

Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) is based on the principle of the field effect [1], which is the modulation of charge density in a material by an electric field applied perpendicularly to its expected flow. In a MOSFET the charge is controlled at the semiconductor surface, called the channel. A MOSFET is a four-terminal device consisting of: source, drain, gate and substrate as shown in Figure 1.1. Modulation of charge carriers in the semiconductor channel is governed by the polarization of gate (V_G) and substrate (V_B) through an insulating layer (the dielectric layer) by capacitive coupling. After their generation, carrier's transportation in the channel is controlled by potential difference (V_D) between the two charge reservoirs (source and drain) establishing a drain current I_D. The transistor thus behaves like a switch. It is in an 'OFF' state (V_G = 0 V) for which the current between source and drain is very low and in an 'ON' state for which V_G is equal to V_D, allowing the passage of current I_{ON}. The threshold voltage (V_T) is the minimum gate-to-substrate voltage difference that is needed to create a conductive channel between the source and drain terminals.

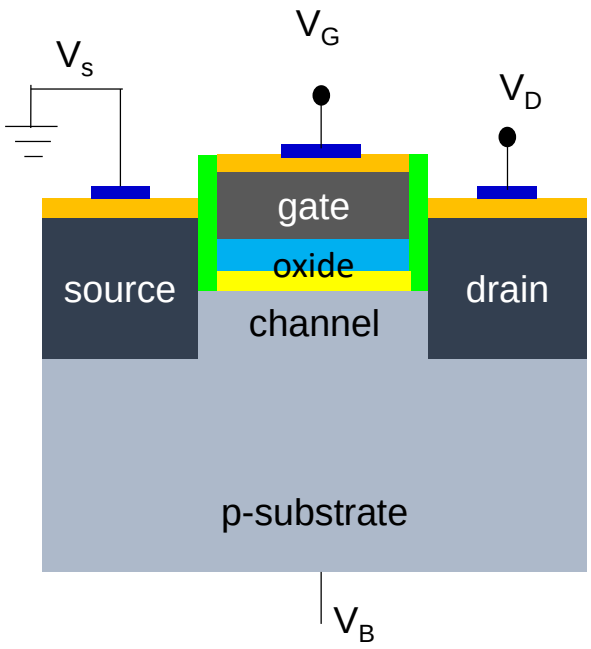


Figure 1.1: Schematic diagram of a four terminal n-MOSFET device

The MOS capacitor is the simplest MOS structure and is composed of a p-type (for NMOS) or n-type (for PMOS) doped silicon substrate, a gate oxide and a metal layer. When the different layers constituting the MOS structure are brought into contact, an alignment of Fermi levels of metal (EF_M) and semiconductor (EF_{Si}) appears. As a result, a unique thermodynamic system is formed which is characterized by a single Fermi level (Figure 1.2). Then, an electrostatic potential, called contact potential (V_C), is generated following the alignment of the Fermi levels. It is expressed by:

$$V_C = \Phi_M - \Phi_{Si} \qquad 1.1$$

where $q\Phi_M$ corresponds to the difference between the metal Fermi level and its vacuum level, $q\Phi_{Si}$ corresponds to the difference between the semiconductor Fermi level and its vacuum. $q\Phi_F$ is the difference between the intrinsic Fermi level of silicon and the resulting Fermi level. In addition, a bending of energy bands is directly driven by the contact potential in the semiconductor (V_{Si}) and in the oxide (V_{OX}) and at $V_G = 0$ V their relationship is given by:

$$V_C = \Phi_M - \Phi_{Si} = -(V_{OX} + V_{Si}) \qquad 1.2$$

When a gate bias V_G is applied, then equation 1.2 becomes:

$$V_G = \Phi_M - \Phi_{Si} + V_{OX} + V_{Si} \qquad 1.3$$

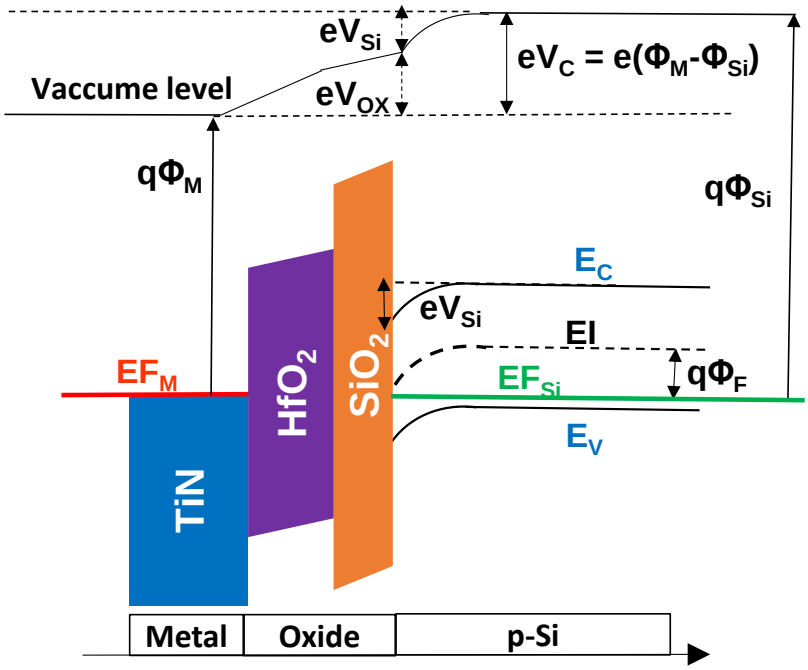


Figure 1.2: Energy band structure of the stack constituting the MOS capacitor at $V_G = 0$ (depletion condition)

Assuming that the oxide is perfect (absence of charges), the oxide potential drop V_{ox} will be linear. So, we can express it using Gauss theorem from the charge Q_{Si} created in the semiconductor:

$$V_{OX} / T_{OX} = -Q_{Si}/\varepsilon_{ox} \qquad 1.4$$

where Q_{Si} is the charge in the semiconductor and T_{Ox} is the oxide thickness. The relationship between Q_{Si} and V_{Si} can be described by the Poisson equation for electrostatics, according to classical electrodynamics [1]. However, for ultra-thin oxide thicknesses and low field strengths, such interactions are better described by quantum Poisson-Schrodinger simulations described in chapter 2.

The charges induced in the semiconductor may be of three types: majority carriers, minority carriers and depletion charge. These charges can be controlled by the voltage applied to the gate V_G, specifically by V_{Si}. For a p-type semiconductor substrate and with a NMOS transistor, the following regimes occurs depending on the applied V_G:

- $V_{Si} < 0$: *Accumulation regime*. Majority carriers are attracted to the surface of the semiconductor and their density becomes larger than that in the substrate volume. The transistor is in the "OFF" state (Figure 1.3a).

- $V_{Si} = 0$: *Flat band condition.* The flat band voltage (V_{FB}) is the bias condition of the gate V_G at which there is no charge in the substrate. At this condition the applied potential V_G equals the contact potential V_C. There is no bend bending at the Si interface, i.e $V_{Si} = 0$, and the concentrations of minority and majority carriers are constant throughout the substrate (Figure 1.3b).

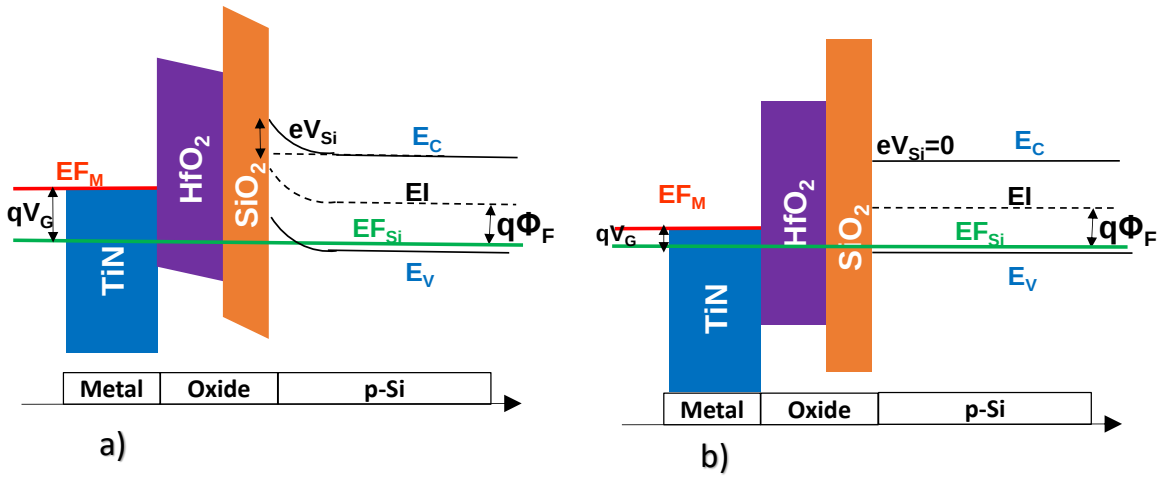


Figure 1.3: Energy band structure of the stack constituting the MOS capacitor in a) Accumulation condition and b) Flatband condition

- $0 < V_{Si} < \Phi_F$: *Depletion regime*. Majority carriers are reduced at the interface and minority carriers starts to increase at the semiconductor interface. An area depleted of carriers is created near the surface (Figure 1.2)

- $\Phi_F < V_{Si} < 2\Phi_F$: *Weak inversion or subthreshold regime*. Minority carrier density starts to increase and equals the majority carriers at the surface, but still remains low compared to that of the majority carriers in the substrate volume. Transistor is in the "OFF" state or weak inversion under threshold voltage (Figure 1.4a).

- $2\Phi_F < V_{Si}$: *Strong inversion regime*. The transistor is in the "ON" state. Minority carrier density becomes greater than the majority carrier concentration in the substrate volume. In this case, a minority carrier channel is formed between the source and the drain on the surface of the semiconductor and is called the inversion channel (Figure 1.4b).

The threshold voltage V_T of a MOS can be defined as the gate voltage V_G such that the condition $V_{Si} = 2\Phi_F$ is satisfied. Thus we obtain as follows:

$$V_T = \Phi_M - \Phi_{Si} + 2\Phi_F - \sqrt{(4\varepsilon_{Si} q N_{Si} \Phi_F)/C_{OX}} \qquad 1.5$$

where, ε_{Si} is the semiconductor permittivity, N_{Si} is the substrate doping concentration (cm^{-3}) and C_{OX} is the capacitance per unit area. In fact, if there is one unique V_{FB} condition corresponding to semiconductor flatband at the interface, V_T characterizing the onset of the transistor may find several definitions: such as specific surface potential V_{Si}, certain channel current density at a certain drain voltage V_D, maximum of transconductance, extrapolation to $I_D = 0$ etc.

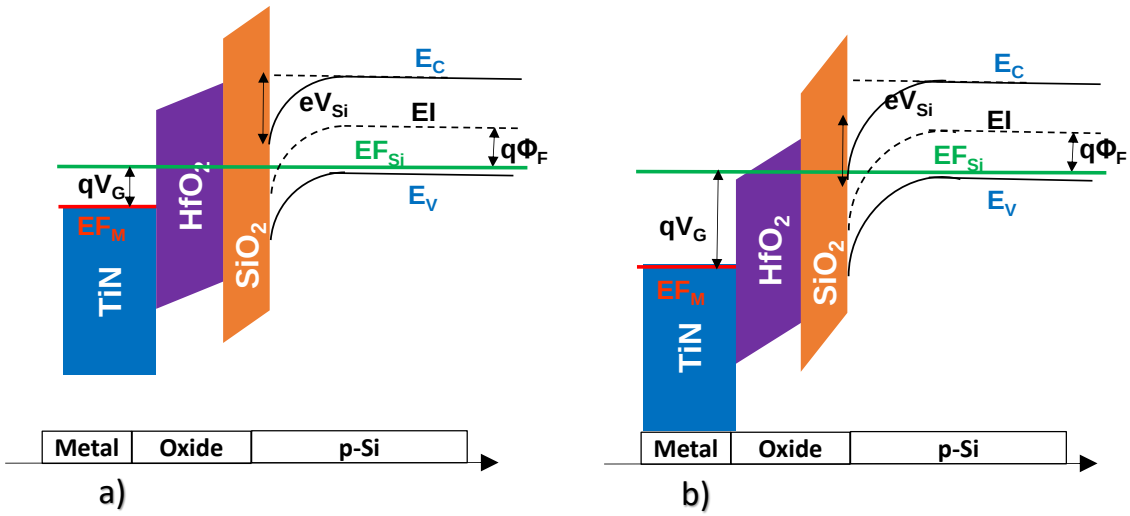


Figure 1.4: Energy band structure of the stack constituting the MOS capacitor in a) Weak inversion and b) Strong inversion condition

The heavily doped (n+ for NFETs and p+ for PFETs) source/drain regions, are used to make an ohmic contact with the conductive channel for $|V_G| > |V_T|$, so that a voltage difference between the source and the drain (V_{DS}) will result in a current flow (I_{DS}) of minority carriers (electrons for NFETs and holes for PFETs) from the positive voltage at the drain terminal to the negative voltage at the source (the polarity and current flow is the contrary for a PMOS). This current flow of minority carriers is also known as drive current and is one of the main MOSFET device performance parameters. In the "OFF" state, the drive current is very small (ideally zero) and in the "ON" state, it is a function of both V_G and V_D.

1.1.2 Metal-Oxide-Semiconductor gate stack properties

Gate dielectric capacitance and equivalent oxide thickness

Capacitance is the ability of a body to store an electrical charge on two electric conductors separated by a dielectric layer. When a potential difference V is applied to the conductors, an electric field develops across the dielectric, causing positive charge (+Q) on one plate and negative charge (-Q) on the other plate. The capacitance C is defined as C = Q/V. In a Metal-Oxide-Semiconductor capacitor, one of the plates is the metal gate and the other is the silicon substrate. The dielectric consists of the gate oxide with a relative permittivity ε_{ox} (commonly known as the dielectric constant) and thickness T_{OX}. Consequently, the gate dielectric capacitance per unit area (C_{OX}) can be expressed as follow:

$$C_{OX}=\varepsilon_0\varepsilon_{ox}/T_{OX} \qquad 1.6$$

where $\varepsilon_0 = 8.854.10^{-12}$ F/m is the vacuum permittivity.

SiO_2 was the industry standard oxide material for MOS capacitance measurements. In order to compare different dielectric materials, and with respect to SiO_2, the notion of EOT (Equivalent Oxide Thickness) has been introduced. EOT is the equivalent thickness of SiO_2 oxide giving the same capacitance for any dielectric (usually a high-k dielectric) of physical thickness T_{OX} and dielectric constant ε_{ox}. The EOT is given by:

$$EOT = T_{OX}\, \varepsilon_{sio2}/\varepsilon_{ox} \qquad 1.7$$

The capacitance is thus given by:

$$C_{ox} = \varepsilon_0 \varepsilon_{sio2}/EOT \qquad 1.8$$

The physical measurement of the nanoscale oxide thicknesses is difficult because it evolves during the manufacturing process and physical measurements reach their limit when thickness become close to the nanometer. The extraction of EOT constitutes a more accurate and relevant measurement to exploit the V_{FB} variations with that of the oxide thickness. We will therefore express various equations in terms of EOT rather than the physical semiconductor oxide thickness.

Flat band voltage V_{FB}

The flat band voltage (V_{FB}) is the bias condition of the gate V_G at which there is no charge in the substrate, leading to flat energy bands at the interface (V_{Si} =0, Q_{Si} =0) (Figure 1.3 b). Thus the voltage associated with the applied potential is called the flat band voltage and is expressed as follows:

$$V_{FB} = \Phi_M - \Phi_{Si} + V_{OX} \qquad 1.9$$

The term $V_{OX,}$ at flatband, is 0 for an ideal device (when no oxide charges or interfacial drops are present and Eq. 1.4 applies). For non-ideal cases, V_{OX} is not equal to 0 and so V_{FB} is influenced by oxide charges density and interfacial voltage drops such as dipoles. So for real devices it becomes a process dependent parameter, especially in bilayer high-k/SiO_2 oxide structures that are currently being used in advanced MOSFETs.

There are generally 4 types of charges possible in a MOSFET device that can affect the V_{FB} (Figure 1.5):

- The interface fixed charges Q_{fi}, that can be located near the Si/SiO_2 interface, high-k/SiO_2 interface, metal/high-k layer [2]. Fixed oxide charges do not move or exchange charge with the underlying silicon and also do not change with the applied voltage.

- Bulk oxide charges Q_{bulk} are the charges in the volume of the oxides and consists of the oxide trapped charges (Q_{ot}) and the bulk fixed charges (Q_{fb}). Like interface fixed charges, these do not move or exchange charge with the underlying silicon [3].

- The interface trapped charges (Q_{it}) are positive or negative charges located at the Si/SiOx interface. They are due to structural defects, oxidation-induced defects or dangling bonds at this interface. Unlike fixed charges or trapped charges, interface trapped charges are in electrical communication with the underlying channel and can thus be charged or discharged, depending on the Si surface potential V_{Si} [3].

- Mobile ionic charges (Q_m) are primarily due to positive alkali ions in the oxide such as Na+, K+ and Li+ [4], incorporated during device processing steps.

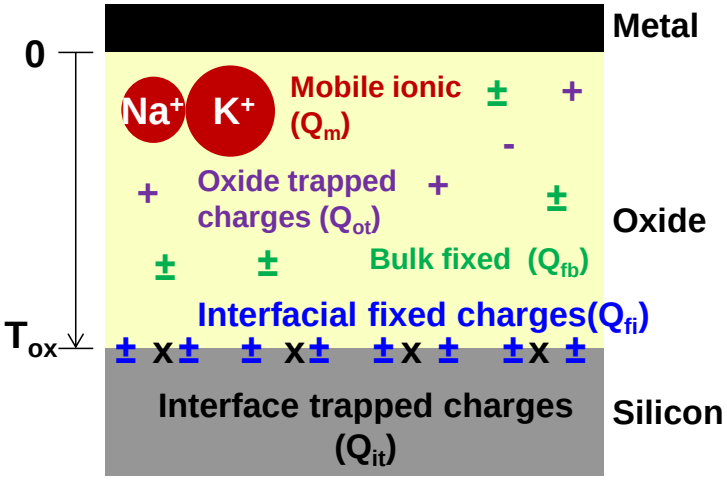


Figure 1.5: Location of charges in thermally oxidized silicon [4]

The effect of each charge on the V_{FB} condition depends on its distance from the oxide/silicon interface and can be calculated from Gauss law. Assuming a uniformly distributed charge per unit volume, V_{FB} shift induced by bulk charges will then vary with the square of its thickness (or EOT) [2]. V_{FB} shift induced by interfacial fixed charges results in a linear modification of the V_{FB} vs EOT plot for a given oxide.

It has been experimentally demonstrated that V_{FB} vs EOT plot is described by a straight line for SiO_2 and HfO_2 oxides [2] [5] [6]. Therefore, the impact of bulk charges in the dielectric layers is negligible. Moreover, recent studies have evidenced the independence of V_{FB} shift with HfO_2 thickness, implying that no isolated fixed charges are present at the HfO_2/SiO_2 interface or in HfO_2 bulk. Indeed, any charge at HfO_2/SiO_2 interface or in HfO_2 bulk would lead to increasing V_{FB} shift when HfO_2 thickness increases [6]. Based on these results, the V_{FB} equation including the charges at the Si/SiO_2 interface ($Q_{Si/SiO2}$) can be written as [5]:

$$V_{FB} = \Phi_M - \Phi_{Si} - Q_{Si/SiO2}\,(EOT/\,\varepsilon_{sio2}) \qquad 1.10$$

As the threshold voltage is related to the V_{FB}, charges also directly influence the threshold voltage and most of the performance parameters of MOSFET devices, such as effective mobility, junction leakage, noise, reliability and breakdown voltage in discrete transistors and digital integrated circuits.

Dipole effect

Interface dipoles can be present in the gate stack at the various interfaces, where they generate potential drops (δ) and so their effect on V_{FB} is independent of the thickness of the oxides. Several models have therefore been proposed to explain dipoles at various locations in the gate stack, in particular at metal/high-k interface and at high-k/SiO_2interface. Dipole formation at the interfaces have been mainly explained by these two phenomena

- Electronegativity differences: Two materials in contact with each other will have atoms with different electronegativities at their interface. For instance at the interface between SiO_2 and high-k, an atom of oxygen is bonded to an atom of high-k on one side and to an atom of Si on the other. Each bond has a dipole moment μ, which depends on the charge Q carried by the bond and its length d. In bulk high-k or SiO_2, the dipole moments compensate each other, but at the interface a net dipole moment is created, which is the sum of individual dipole moments. HfO_2 or La_2O_3 thus creates a dipole in contact with SiO_2 (Figure 1.6). This mechanism has been utilized to explain V_{FB} shifts due to dipole at many interfaces. Such as the dipole at the interface between metal gate and high-k dielectrics [7] [8], and also for the case of capping the top surface of the high-k layer with additives such as aluminum [9] or lanthanum [10].

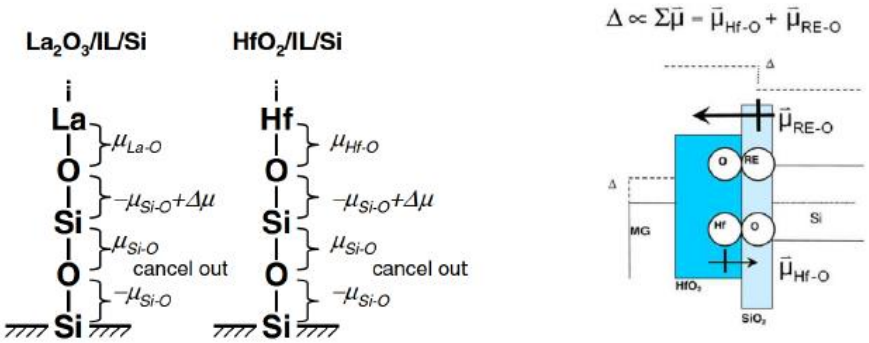


Figure 1.6 Dipole moments associated with HKMG stack interfaces [11][12]

- Oxygen vacancy formation: According to this model, difference of oxygen atomic density (σ) at the interface causes deformation and strain at the interface, leading to an increase of free energy of the interface. The free energy of the interface should be therefore minimized by the movement of oxygen from larger σ side to the smaller [10]. As a result, a charge imbalance at the interface is induced by the movement of oxygen as it is negatively charged. This mechanism has been used to explain V_{FB} shifts due to dipole at the metal/high-k interface [13] [14], at high-k/SiO_2 interface [15] or at the Si/SiO2 interface [16].

The final equation of the V_{FB} combining the effects of all the charges inside the gate stack and dipoles at the interfaces becomes:

$$V_{FB} = \Phi_M - \Phi_{Si} - Q_{Si/SiO2}\,(EOT/\,\varepsilon_{sio2}) + \Sigma\delta \qquad 1.11$$

V_{FB} and EOT can be calculated by fitting experimental capacitance voltage curves. The above relationship between V_{FB} and EOT can be used to extract gate stack charges and dipoles. $QS_{i/SiO2}$ can be calculated from the slope of the V_{FB} vs EOT curve and $WF_M + \delta$ can be estimated from the extrapolated V_{FB} value at EOT=0 [17]. Practically EOT is modulated along the wafer radius by varying the SiO_2 IL, called as the bevel process (section 1.3.3) which has been developed at STMicroelectronics's 300 mm wafer fab.

Metal effective work function

The work function of the metal WF_M is defined as the minimum energy needed to remove an electron from the metal to a point in the vacuum immediately outside the metal surface. For an ideal device without any charges and dipoles, it is given as follows:

$$WF_M = q\Phi_M = qV_{FB} + q\Phi_{Si} \qquad 1.12$$

We have seen that V_{FB} of real devices can be affected by various process steps that can introduce fixed charges at the interface or in the bulk of oxides and interface dipoles. In this case, equation 1.12 does not hold any longer and the concept of effective work function WF_{eff} has to be introduced. WF_{eff} takes into account not only $q\Phi_M$ but also the V_{FB} contribution of the charges and dipoles in the gate stack. So the WF_{eff} is described as:

$$WF_{eff}/q = V_{FB} + \Phi_{Si} = \Phi_M + \Delta V_{FB}, \quad \Delta V_{FB} = - Q_{Si/SiO2}\,(EOT/\,\varepsilon_{sio2}) + \Sigma\delta \qquad 1.13$$

1.2 Gate stack fabrication process

1.2.1 Gate dielectrics

Gate dielectrics are responsible for the capacitive effect produced in a MOSFET device. Traditionally, SiO_2 has been used as the gate dielectric, but continuous device scaling for future technology nodes required reduction in the thickness (or EOT) of gate dielectrics in order to maintain high drive current and adequate gate capacitance. This scaling of SiO_2 was challenged due to exponential increase in gate leakage currents as the thickness decreased [18].

In order to replace SiO_2 and to overcome these problems, SiO_2 incorporated with nitrogen was proposed to form oxynitrides (SiO_xN_y). At the beginning this became an appropriate solution to increase MOSFET performance[19]. Compared to non-nitrided films, the oxynitrides films containing nitrogen atoms increased its diffusion barrier properties, leading them to be more resistant to further oxidation and diffusion of dopants from the gate to the channel and vice versa. Moreover, nitrogen atoms at the Si/SiO_2 interface resulted in the reduction of defect generation and traps in the oxide [20]. In addition, nitrogen increases the dielectric constant of the oxynitride. This

increase is linear with the percentage of nitrogen, from "SiO_2 = 3.9 to "Si_3N_4 = 7.8 [21] and thus reduces gate leakage. SiON however reached its limits from the 45nm node because of the weak permittivity and degradation of reliability due to the presence of nitrogen atoms. Gate leakage on SiON dielectrics became superior to specifications recommended by the ITRS.

As an alternative to SiON gate dielectrics, a high permittivity (high-k) material was introduced to continue the aggressive scaling of advanced MOSFETs. Due to its high permittivity a larger thickness can be used, compared to SiON, while keeping the same EOT and thus decreasing the leakage current. The direct deposition of the high-k dielectric on the silicon substrate forms a layer of unintentional silica of poor quality and is detrimental to transistor electrical properties, such as threshold voltage, channel carrier mobility, interface traps and charges in the dielectric [22]. On the other hand, a well-controlled silicon oxide or oxynitride (SiO_2, SiO_x, SiON) (interfacial Layer (IL) deposited on Si substrate has excellent surface quality and therefore can reduce or even eliminate the above problems associated with a High-k/Si interface [23].

Interlayer dielectric growth

As mentioned above, the electrical properties of MOSFET devices are strongly correlated with structure and defects near and at the Si/SiO_2 interface. High quality ultra-thin oxides films are not easily fabricated with conventional thermal oxidation methods. Therefore, enhanced techniques to grow ultra-thin SiO_2 films with superior quality and performance have been designed and implemented in recent years. As discussed before, SiON has superior electrical properties compared to SiO_2 and so we will focus on the deposition techniques for SiON.

The main fabrication techniques, used for SiO_2 film growth, are the rapid thermal oxidation (RTO), in-situ steam generation (ISSG), rapid thermal chemical vapor deposition (RTCVD) and remote plasma enhanced chemical vapor deposition (RPECVD) [24]. Nitridation of SiO_2 films, formed by RTO or ISSG, by DPN (decoupled plasma nitridation) or RPN (remote plasma nitridation) have been investigated earlier [24]. Nevertheless, nitridation in amonia (NH_3) is a simple way to introduce relatively high concentrations of nitrogen into SiO_2. The ultra-thin oxynitride (SiON) films (between 8Å and 12 Å), used in this work as the interlayer dielectric, is fabricated at STMicroelectronics by performing first an enhanced ISSG oxidation of silicon at 800°C, followed by NH_3 nitridation and RTP anneal at 700°C.

High-k deposition

The SiO_2 or SiON dielectrics by themselves are not able to satisfy the gate leakage requirements at the 45 nm node and beyond. The high-k material has to meet several criteria to be considered as a candidate to replace SiO_2, as listed below. The high-k must:

- Have a high enough dielectric constant
- Have a large conduction and valence band offsets relative to metal Fermi level and to the conduction and valence bands of silicon, as gate leakage decreases with band offsets.
- Form a good quality interface with the channel so as not to degrade mobility
- Have good stability in the different stages of transistor fabrication, meaning retention of the amorphous phase at high temperatures. Indeed, a polycrystalline material could result in grain boundaries acting as current leakage paths.
- Have minimum traps and fixed charges

- Meet the feasibility criteria in industrial conditions (thin layers, short deposit times, low thermal budgets and cost)
- Meet the requirements in terms of reliability

Among the various potential candidates, Hafnium-based dielectrics (HfO_2 and HfSiOx silicates, HfSiON) best meet these criteria [24]. Within the International Semiconductor Development Alliance (ISDA), which included STMicroelectronics and IBM, HfSiON material was selected to succeed SiON in the technological nodes 32 / 28nm. Indeed, HfSiO silicates have better results than HfO_2, because of better thermal stability, a gain in mobility and a reduction in charge trapping. However, it has been reported that HfO_2 has a higher dielectric constant, about 18-25 [25], compared to about 15 for HfSiO [26]. Moreover, Hf-based gate dielectrics have high band offsets, Ec = 1.5 eV and Ev = 3 eV [25], leading to leakage current densities at least two orders of magnitude lower than SiO_2. The presence of nitrogen in the film (HfON or HfSiON) enhances the thermal stability relative to HfSiO, leading to amorphous films up to annealing temperatures of 1100°C [27]. This high thermal stability makes these dielectrics suitable for integration into the CMOS process flow, being able to handle the subsequent source and drain high temperature anneals.

Initially, physical vapor deposition (PVD) was used to deposit hafnium-based gate dielectrics [27]. Now, chemical vapor deposition (CVD) or Metal organic CVD (MO-CVD) and atomic layer deposition (ALD) are being used to deposit these layers. This has been done to meet the compositional control and conformality requirements of the films. The CVD or MO-CVD processes, used to deposit HfSiO films, uses metal organic precursors. ALD processes are used to deposit HfO_2 uses hafnium tetrachloride ($HfCl_4$) as precursor. Finally, in order to incorporate nitrogen into the films to further increase the dielectric constant, nitridation methods such as Decoupled Plasma Nitridation are currently used. In 14 or 28 nm MOSFET devices fabricated at STMicroelectronics, HfON and HfSiON are deposited over the interlayer dielectric (SiON) by ALD or by MO-CVD respectively in order to keep a high mobility interface, followed by nitridation using DPN and Post-Nitridation Anneal before metal gate deposition. The resulting high-k layer has a thickness between 19Å and 21 Å.

1.2.2 Metal gate electrode and deposition techniques

Choice of gate material

Poly-crystalline silicon (poly-Si) has been used for decades as the gate for MOSFET. It was used for its good electrical conduction properties when doped with impurities. However, poly-Si could not be used for advanced technological nodes and metal gates were introduced for several reasons mentioned below:

- Metal gates eliminate the depletion zone formed at the poly-Si and high-k interface that acts as a parasitic capacitance, and boron penetration effects. This thus reduces the EOT penalty and increases the ON current of the transistors.

- Poly-Si in contact with high-k causes the Fermi level pinning. This causes the modification of the Poly-Si Fermi level due the formation of dipole at the poly-Si and high-k interface [13].

- Surface phonon scattering degrades the electron mobility in the high-k, reducing the speed at which transistors can switch. Introduction of metal gate significantly increases the density of electrons in the gate electrode [28]. Therefore, the higher density of electrons in a metal screen out the vibrations and let current to flow more smoothly, compared to poly-Si gate with lower free electron concentration.

- Metal gates show less resistance compared to Poly-Si gate.

Choice of Metal gate

The first criteria for the choice of metal gate is its work function WF_m that determines the threshold voltage of the transistor. Advanced technology nodes requires low threshold voltage for both NFET and PFET devices. The effective work function target depends on the channel doping density [29] (Figure 1.7), as predicted by the term Nsc in Eq. 1.5. According to this, for high doping concentration, metals with workfunction close to the silicon conduction band (CB) or the silicon valence band (VB) are required, for NFETs and PFETs respectively [30]. In contrast, for low doping concentration, the target is almost midgap work functions, like TiN, at only 100 meV from either side of the intrinsic silicon Fermi level. Therefore, in fully depleted silicon on insulator (FD-SOI) devices with undoped Si channel, almost midgap metals will be required for low V_T, in contrast to bulk technologies [31] [32].

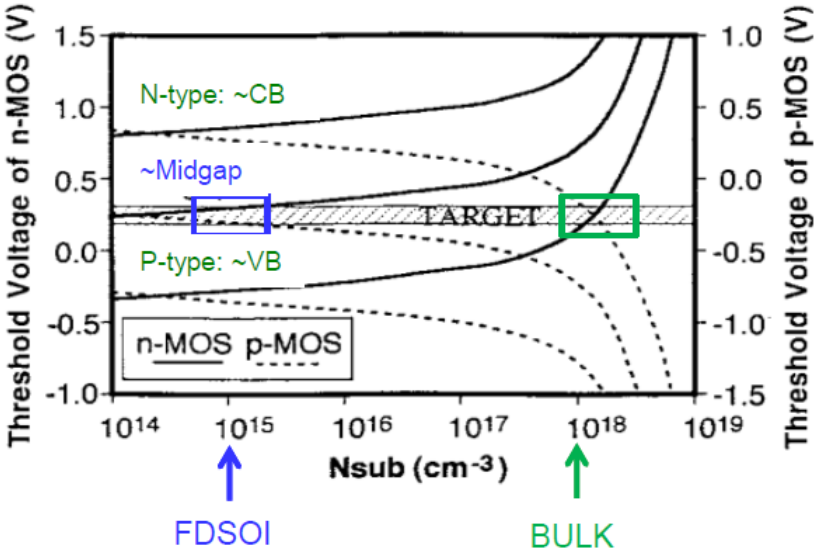


Figure 1.7: Metal work function requirements as a function of channel doping density, to achieve low V_T for both NFET and PFET devices [29]

The second criteria for the choice of the metal gate is its thermal stability. This means that their microstructure should remain stable and that they do not react with the other layers. During the process of MOSFETs devices, the development of HKMG stack is followed by a high temperature annealing (around 1050°C) for source and drain dopant activation (S-D anneal). This process method is termed as the Gate First approach. Figure 1.8 shows that only the mid-gap metals, including TiN, and high WF_m metals are thermally stable above 700°C [33]. In contrast, low WF_m metals, being less electronegative, will tend to form an oxide at high temperature.

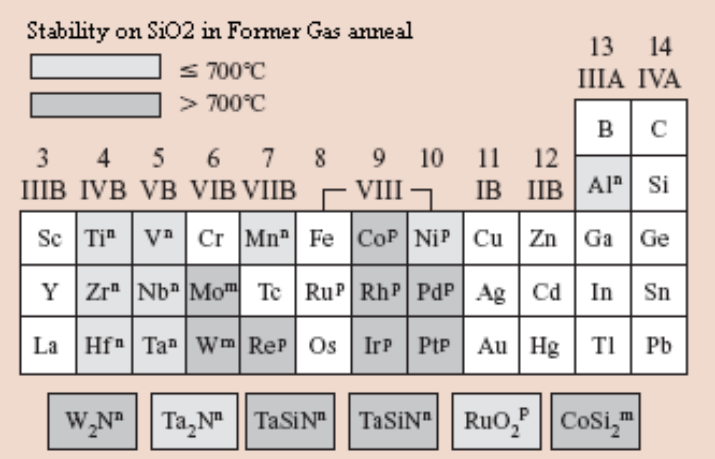


Figure 1.8: Characteristics of metals in terms of thermal stability and workfunction, indicated by the gray level for the stability and an indication of N-like or P-like behavior index of each element [33]

Another effect of the high temperature annealing is the diffusion of oxygen from the metal gate towards the pedestal oxide and thus degrading the EOT of the device. This has been observed in particular for P type metals such as tungsten, platinum or rhenium but solubility of oxygen in pure titanium (Ti) is high (10%). In case of a $Ti/HfO_2/SiO_2/Si$ stack, annealing even leads to the removal of some interfacial oxide [9]. This is due to the diffusion of oxygen from the pedestal SiO_2, through the HfO_2, toward the TiN gate due to the high solubility of oxygen in titanium nitride (TiN) and is called the oxygen scavenging effect. This strong affinity of TiN for oxygen makes it possible to reach EOTs of the order of 1 nm. All the advantages mentioned above (mid gap WF_M, excellent thermal stability and the scavenging effect), makes TiN almost indispensable in a Gate First integration strategy.

Metal gate deposition techniques

TiN metal films are mainly deposited by Atomic Layer Deposition (ALD) and Physical Vapor Deposition (PVD) techniques. ALD is a thin film deposition technique that is based on the sequential use of a gas phase chemical process. ALD makes atomic scale deposition control possible and provides extremely conformal, uniform thickness and low impurity level films. It has successfully been used to deposit TiN films from $TiCl_4$ and NH_3 precursors [34]. However, despite

these advantages TiN metal gate was not deposited by ALD in our work at STMicroelectronics. The main limitations of ALD were the slow deposition rate and high cost, which has limited its use in the semiconductor industry. However, ALD is becoming the preferred choice for applications that require high global thickness non-uniformity and conformal step coverage, such as for extremely scaled devices with gate last integration scheme. Nevertheless, for our work that uses the gate-first processes, such conformal step coverage is not essential. For this reason, Radio-Frequency PVD (RF-PVD), allowing deposition at room temperature and at higher deposition rates, have been chosen for metal gate applications at STMicroelectronics.

Physical Vapor Deposition

PVD processes include a variety of thin film deposition methods that are used to deposit thin films atomically by means of fluxes of individual neutral or ionized species. Cathodic arc deposition, electron beam physical deposition, evaporative deposition, pulsed laser deposition and sputter deposition are some examples of physical vapor deposition methods. In our work, TiN metal gate films has been deposited by the sputtering method.

Sputtering is the process of removing surface atoms or molecules from a solid target by the bombardment of ions and involves ejecting those atoms or molecules from a source (target) onto a substrate (silicon wafer). Sputtering, unlike evaporation, does not require melting the metal to be deposited. Therefore, refractory metals such as titanium (Ti) and tungsten (W), that are very difficult to melt, can be used. Sputtering also preserves the original composition of the target material. In addition, sputtering allows the deposition of many different materials by combining materials from solid and gaseous sources, which are introduced into the vacuum chamber either before or during deposition. This form of sputter deposition is called reactive sputter deposition.

The sputtering process is carried out in a vacuum chamber in order to avoid contaminants to interfere with the deposition process and to establish the pressure required by the sputtering plasma. The high vacuum pumping (in the 10^{-8} Torr range) is done by cryopumps. In a simple DC sputtering system, the target plays the role of the cathode and the substrate, that of the anode. When a DC voltage equal to a voltage known as breakdown voltage (Vb) is applied between both electrodes, a plasma is created. The electrons collide with argon and create positively charged argon ions (Ar+), which are strongly attracted to the negatively charged cathode (the target). The argon ions collide with the target surface and some of them causes surface atoms of the target to be ejected. The sputtered atoms travel to the substrate where they are deposited as a film.

For a given gas, the minimum voltage Vb necessary to create a stream of electron between the electrodes is a function of the product of the pressure (p) and the electrodes gap distance (d) (parameter pd in Fig. 1.9). The curve of voltage versus pd is called Paschen's curve. Examples of Paschen's curves obtained for different gases are shown in Fig. 1.9. The pressure has to be high enough to keep a high density of Ar+ ions in the plasma to sputter the target but low enough to reduce the collisions of the sputtered atoms with Ar atoms. A pressure too low or too high have to be compensated by higher breakdown voltages, as described in Figure 1.9.

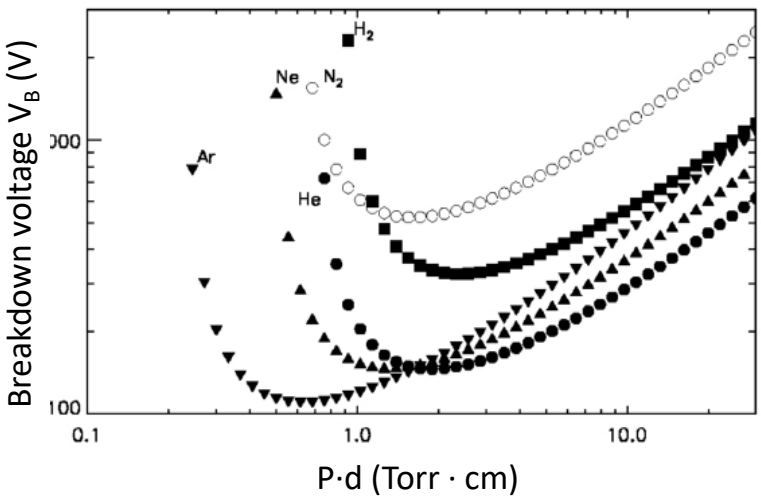


Figure 1.9: Paschen curves obtained for Helium, Neon, Argon, Hydrogen and Nitrogen [35]

Sputtering effects can be enhanced by adding magnets behind the cathode to a simple DC sputtering system. In this case, the sputtering rate is improved because the Ar ionization efficiency is enhanced. Electrons are confined because they tend to follow the magnetic field lines, ionizing more argon atoms in their path. In this way, both electron and Ar+ ions density are increased in a DC magnetron system. On one hand, with the increased electron density, Vb can be decreased from 5000 V - 10000 V in a simple DC sputtering system to only 400 V – 800 V in a DC magnetron system. On the other hand, due to the increased density of Ar+ ions, it is possible to lower the sputtering chamber pressure. At lower pressures, the sputtered atoms have fewer collisions on their path to the substrate which results in an increased deposition rate. Finally, the electron confinement in a magnetic field near the target also reduces electron bombardment of the substrate. This results in much less heating of the substrate.

In addition, a Radio Frequency (RF) generator operating at a frequency of 13.56 MHz (standard in industry) can be coupled to a DC power. The main interests of RF-PVD deposition are the reduction of the breakdown voltages and the improvement of film deposition uniformity. Indeed, the voltage needed to ignite the plasma is reduced because oscillating electrons are able to ionize more Ar gas. A direct consequence of the reduction of the breakdown voltage is that the metal atoms are ejected with less energy, limiting the damage that could be caused on the high-k layer by bombardment on the wafer during the plasma sputtering. This is a critical point because any significant damage can be detrimental for gate stack integrity and transistor electrical properties. In addition, less high energetic electrons are provided in a RF sputtering plasma compared to DC plasma for the same density of electrons, as shown in Figure 1.10.

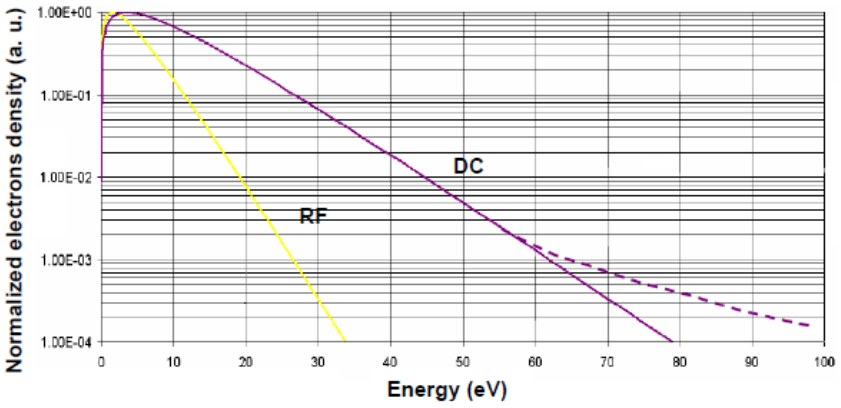


Figure 1.10: Energetic distribution of electrons in RF (yellow) and DC (purple) plasma [36]

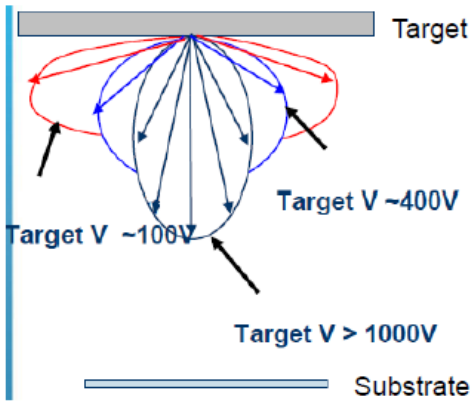


Figure 1.11: Representation of sputtering angle vs. breakdown voltage [36]

The breakdown voltage reduction also allows to increase significantly the sputtering angle of the ejected metal atoms, as shown in Figure 1.11. Consequently, the thickness uniformity can be improved to about 1% for thicknesses < 100 Å.

The Endura platform supplied by Applied Materials has been used in this work for TiN metal gate deposition in RF-PVD process chambers (Figure 1.12). Apart from TiN deposition chamber, two other chambers are present to deposit aluminum (Al) and lanthanum (La). Each RF-PVD process chamber contains a target material (Ti, Al, La) for thin films deposition.

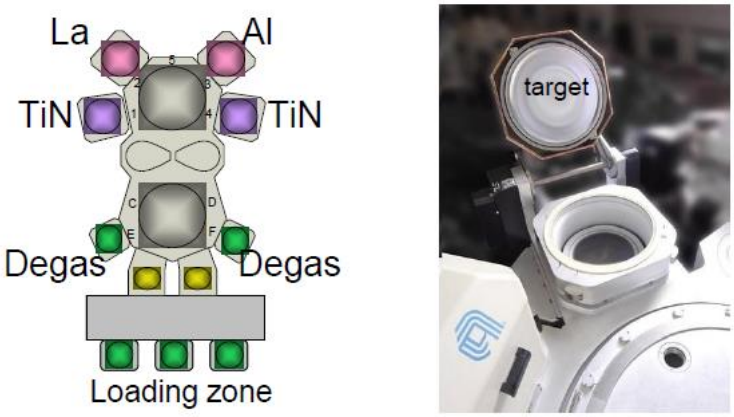


Figure 1.12: On the left, Endura platform with 4 RF-PVD chambers for TiN, lanthanum and aluminum deposition. On the right, photo of a deposition chamber

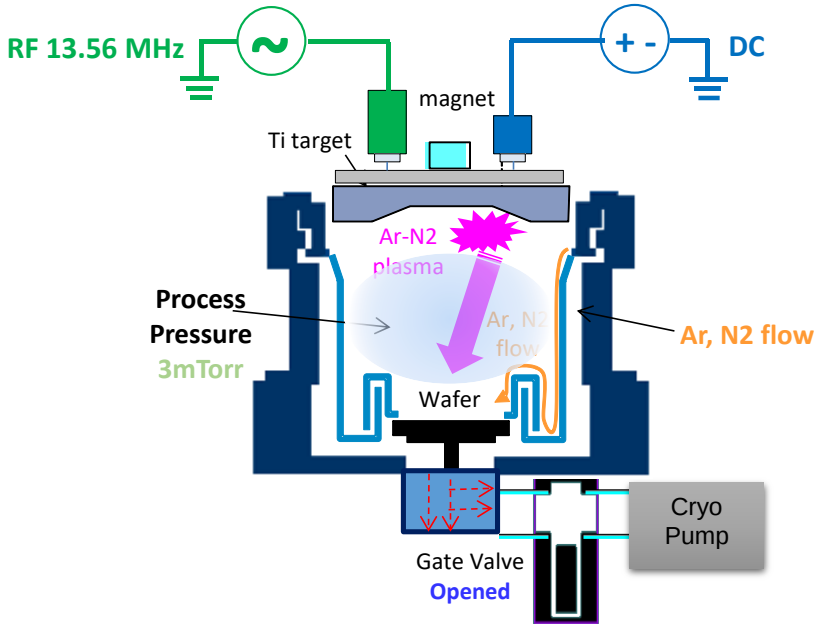


Figure 1.13: Schematic of a RF-PVD chamber used to deposit TiN metal gate layers

In general, all RF-PVD chambers are composed of a RF generator, a DC power supply, a cryo pump system for ultra-high-vacuum and an Ar mass-flow controller to regulate the argon flow into the chamber. Reactive RF sputter chambers used to deposit TiN contains, additionally, a nitrogen mass-flow controller, as shown in Figure 1.13. The stoichiometry of the deposited thin films can be controlled by adjusting the flow of these gases. The RF and DC power can be controlled by tuning the parameters of their respective generators. The partial pressure of Ar and N_2 gases (and

so the total chamber pressure) can be controlled by the flow of these gases from the respective mass-flow controllers and by adjusting the gate valve position (mid-gate to gate-open positions). In chapter 4, the impact of RF power and partial pressure on TiN properties will be investigated.

TiN films deposition

TiN is formed by reactive sputtering of a pure Ti target in a nitrogen-containing ambient, typically Ar/N_2, in a RF-PVD chamber with combined DC and RF generators. Unlike, TiN films deposited by ALD or CVD, the nitrogen and titanium do not combine in the gas phase, but on the surface of the target, chamber walls, and on the wafer. The TiN metal gate layer used in this work is deposited with a gradient profile of nitrogen and its deposition proceeds as follows: first, the Ti target is completely de-nitrided by pure Ar sputtering. Next, during film deposition, the chamber is backfilled with an Ar/N_2 mixture and plasma is initiated with a high volumetric flow rate of nitrogen ($N_2/(Ar + N_2) = 0.7$) in order to progressively cover the surface of the target with nitrogen. As a result, only titanium is sputtered from the surface of the target for the first monolayers deposited over the high-k. As nitrogen covers the target surface, TiN begins to sputter and nitrogen composition gradually increases in the TiN film being deposited. At this point, a noticeable drop of the sputter rate occurs due to the decrease in sputtering yield caused by the nitridation of Ti target surface. Accordingly, the deposition rate decreases from 1.6 Å/s to 0.4 Å/s during the TiN deposition with a total pressure of about 3 mTorr and a DC and RF power of 700W and 600W, respectively. A gradient profile of nitrogen into TiN results in better gate leakage and better reliability [37].

Deposition of Aluminum and Lanthanum monolayers

Al and La monolayers are deposited between the high-k and TiN gate for gate stack work function engineering. These additives forms dipoles at the IL and HK interface and hence modulate the WF_{eff}, this topic will be discussed in detail in later sections. While TiN is deposited with combined DC and RF powers, Al and La are deposited with only RF sputtering at smaller deposition rates because desired thicknesses are much thinner (from 2Å to 6 Å) compared to TiN. RF power and Ar flow conditions are chosen to optimize the deposition rate and the uniformity of the monolayers. Increasing the RF power and reducing the Ar flow results in a higher deposition rate, but an increment of the Ar flow improves the uniformity, as a result of the reduction of deposition rate. In the selected deposition conditions, the deposition rate is extremely low, in the range between 0.1 Å/s and 0.3 Å/s.

1.2.3 Metal gate integration in 14 nm Fully-Depleted SOI devices

As CMOS technology scales down, two approaches have been pursued by the industry to overcome the fundamental limits of traditional planar bulk transistors. One is the introduction of a Tri-gate or FinFET transistor at the 22 nm/16 nm node [38]. The other is the Fully Depleted Silicon On Insulator (FDSOI) architecture, shown in Figure 1.14, which provides a simplified planar manufacturing process compared to 3-D FinFET technology. Although FinFET architectures offer impressive drive currents per footprint at low supply voltages because of the 3-D conduction channel and excellent electrostatic control, they have high gate and parasitic capacitances, proportional to the 3-D effective width increase, which negatively impacts both the speed and active power consumption [39]. In addition, complexity of FinFET technology has created new challenges for many design teams because their current tools and techniques may not have enabled them to design their IP blocks optimally for FinFET processes, delaying the time to market.

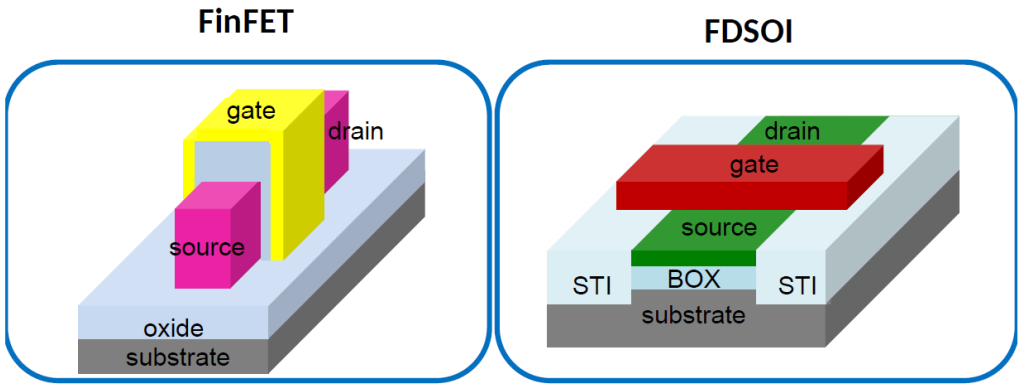


Figure 1.14: Scheme of the FinFET and the Fully-Depleted Silicon-on-Insulator [40].

In contrast, FDSOI technology is derived from bulk design rules and its process technology [41] and the design transfer is thus less complicated. Only few analog parts need to be adopted to the FDSOI technology. It is a planar technology that reuses ninety percent of the process steps used in the bulk counterpart and the overall manufacturing process in FDSOI is 15% less complex, leading to lower cycle time and reduced manufacturing costs. The manufacturing tools for FDSOI are identical to the last generation of bulk processes. FDSOI transistors are manufactured on an ultra-thin buried oxide (BOX) layer of 20 nm. Silicon channel thickness is about 6 nm. Shallow Trench Isolation (STI) feature authorizes the lateral insulation between adjacent semiconductor device components.

1.2.4 Introduction of additives for workfunction engineering

The effective work function (WF_{eff}) values of metal electrodes must satisfy the V_T requirements of specific devices. In undoped-channel 14nm FD-SOI devices, WF_{eff} at only 100 mV from the midgap are required [32] [29], which makes a very fine adjustment of the effective work function of metal electrodes even more necessary.

Lanthanum (La) for NFETS

In order to reach the WF_{eff} specifications for NFETs, La has been introduced into the gate stack. La_2O_3 and La-silicate have been studied as alternate gate dielectrics [42][43]. Insertion of La_2O_3 capping layers deposited above the high-k layer has also been extensively studied [10][44]. Narayanan et al. proposed to insert nanoscale capping layers containing lanthanum into HfO_2/TiN stacks [10] in a conventional gate first flow (Figure 1.15 a). In this way, lanthanum induced large V_{FB} shifts towards N+, from 4.43 eV (without La) to 4.05 eV (with La). However, it is indicated that the La thickness is a critical parameter to optimize, it must be thick enough to shift the WF_{eff} towards N+ but not so much that lanthanum near or in the Si/ SiO_2 interface degrades the mobility. Narayanan proposed that the WF_{eff} shifts can be either due to positively charged mobile oxygen vacancies formed upon the substitution of Hf^{4+} in HfO_2 with lower valence La^{3+} and/or a dipole at the metal/high-k interface due to differences in electronegativities between La and Hf. In contrast, Yamamoto et al. [45] experimentally evidenced that the most plausible origin for the V_{FB} shift is the formation of a dipole layer at the HfLaOx/ SiO_2 interface (Figure 1.15 b).

The potential of metallic lanthanum deposited by RF-PVD inserted into TiN metal gate has also been evaluated in the multilayer TiN/La or TiN/La/TiN structure [46][37] [6] [47](Figure 1.16) and in the alloyed TiLaN metal [48]. Apart from the WF_{eff} shift, these structures were proposed for another benefit: the interlayer scavenging effect. La ions remove oxygen from the interlayer to decrease the overall EOT of the devices and so suppressed the interlayer growth to extremely small values of around 1nm.

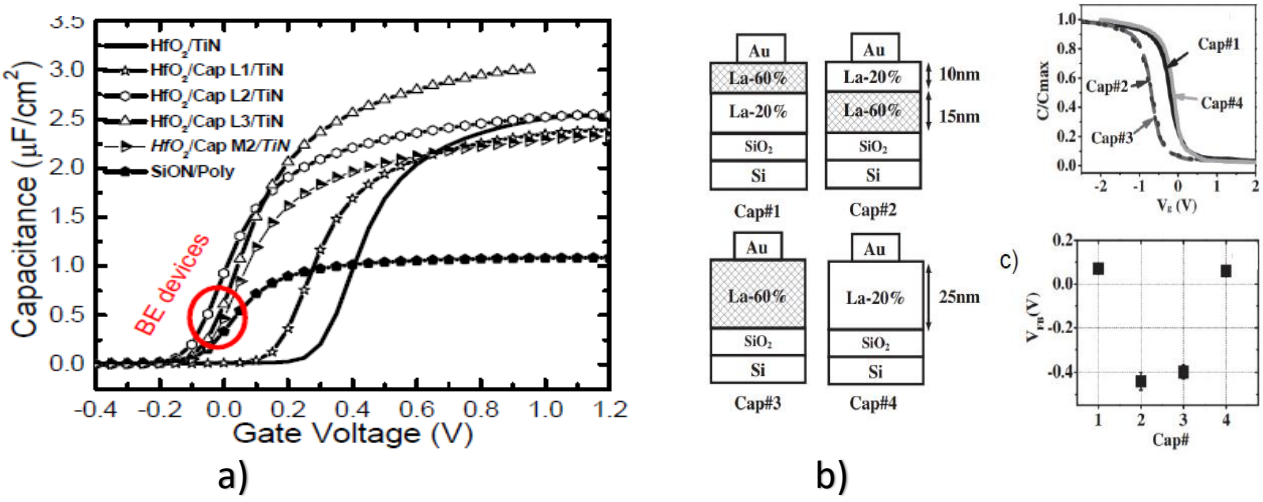


Figure 1.15: a) nFET CVs as a function of increasing La cap thickness showing the Wf_{eff} tunability by La [10] and b) Schematic description of MOS capacitors with different La concentration profiles in HfLaOx films, their CV curves and V_{FB} values [44]

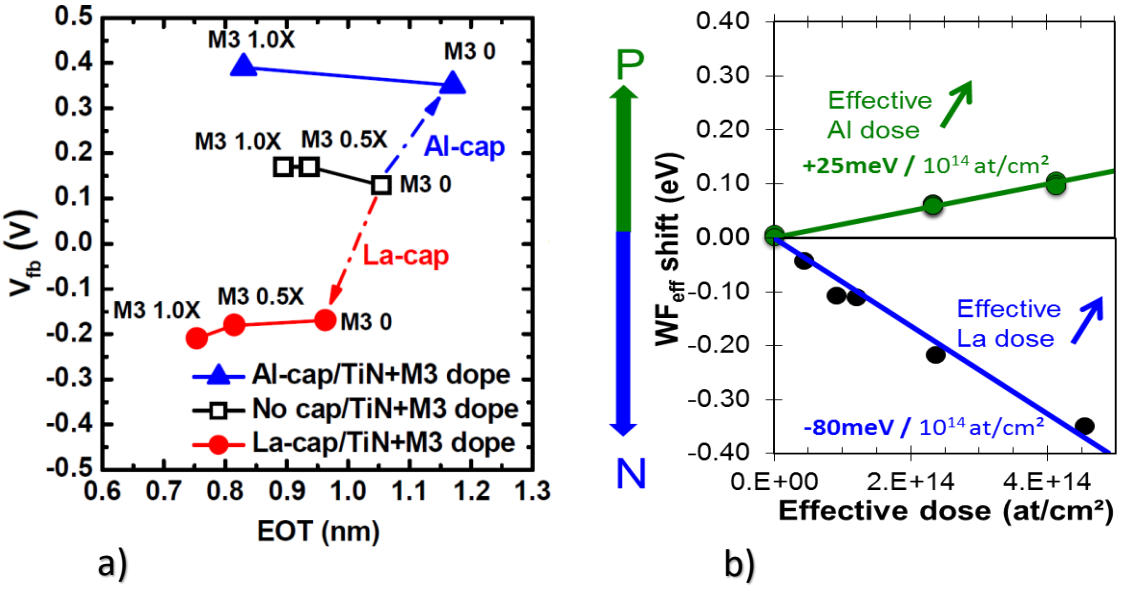


Figure 1.16: a) V_{FB} vs. EOT plot for TiN gate with La and Al cap layers [46] and b) Effective workfunction (Wf_{eff}) vs La or Al dose diffused inside the gate stack [47]

Aluminum (Al) for PFETS

In order to reach the WF_{eff} specification for PMOS, Al has been extensively used to shift the WF_{eff} towards P+. The introduction of Al_2O_3 as an alternate high-k dielectric has been studied [49][50][51]. Al_2O_3 layers has also been incorporated either at the SiO_2/Hf-based oxide interface or above Hf-based high-k dielectric. When a thin Al_2O_3 layer was deposited between the SiO_2 and HfO_2 layers, WF_{eff} increased up to +900 meV compared to SiO_2 reference [5]. In this work it was proved that this Wf_{eff} shift was indeed due to a dipole effect at the SiO_2/high-k interface. Deposition of a thin Al_2O_3 layer on the top of HfO_2 or HfSiON also induced an increase of the WF_{eff} towards P+, but the shift was smaller (100-200 meV) [52] compared to capacitors with Al_2O_3 below HfO_2. The shift towards P+ in a gate-first integration has been explained by the diffusion of Al towards the bottom HfO_2 interface [52][53], creating a dipole layer at the HfO_2/ SiO_2 interface as for Al_2O_3 directly deposited at this interface (Figure 1.17).

The introduction of metallic aluminum deposited by RF-PVD and inserted between two TiN layers (TiN/Al/TiN structure) has also been previously studied [46][47][54][55] (Figure 1.16).

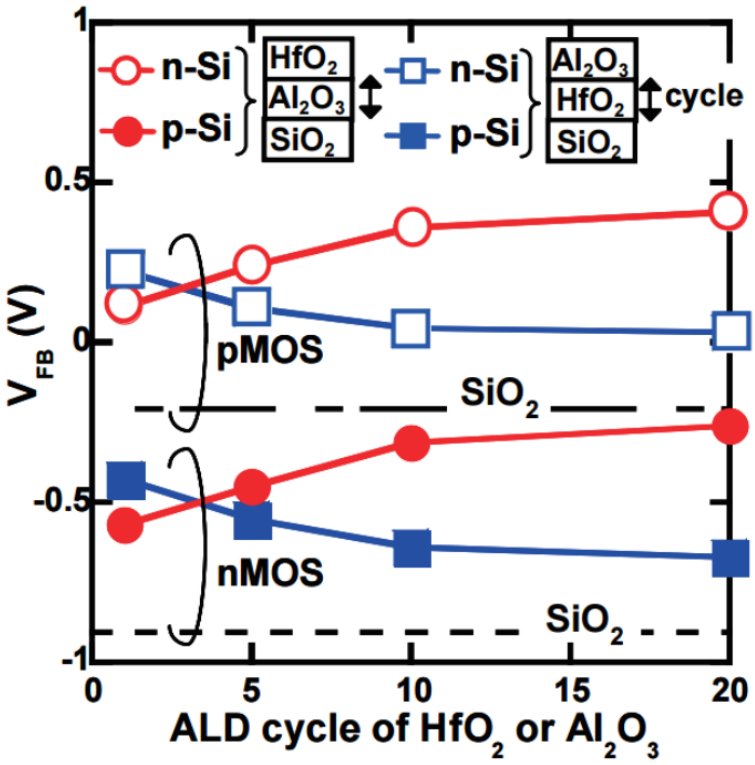


Figure 1.17: V_{FB} variations of n-MOS and p-MOS capacitors with HfO_2/ Al_2O_3 dielectrics [53]

Sacrificial metal gate-first process

In this work, La and Al additives are introduced into the HKMG stack by the sacrificial gate process. Once interlayer and high-k dielectrics have been deposited on the Si substrate, metal gate containing the additives is deposited following the steps given below, and shown in Figure 1.18.

1. First, the sacrificial multi-layer metal gate stack, composed of a TiN layer and monolayers of La or Al additives in between, is deposited in the Endura RF-PVD system described in the section 1.2.2.

2. Next, poly-Si is deposited as a capping layer on the top, followed by a thermal treatment under N_2 atmosphere at certain temperature and for a duration, to activate the diffusion of additives into the High-k/SiON stack. This step is called the drive-in-anneal (DIA).

3. The sacrificial gate stack is then removed by a chemical wet solution.

4. Finally, a poly-Si/TiN electrode without additives is deposited, followed by gate patterning. Later, spike annealing at 1047°C is done, which is responsible of the Source-Drain (S-D) dopant activation.

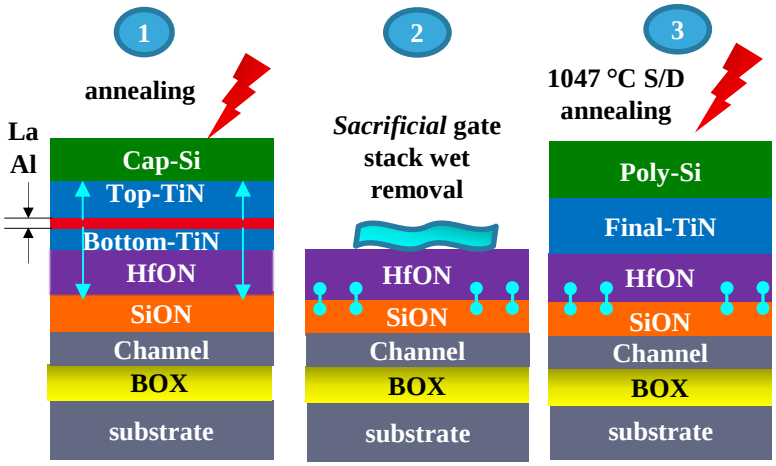


Figure 1.18: Sacrificial gate-first approach: 1) sacrificial gate deposition and annealing, 2) sacrificial gate stack wet removal and 3) final gate deposition and source-drain annealing

1.3 Process flow

In this work, depending on the gate stack electrical properties to be investigated, two types of devices have been used. First are the MOS transistors (MOSFET) that are required in order to study device reliability issues such as NBTI, PBTI and TDDB. These devices are as shown in Figure 1.1, containing the MOS capacitor with the source and drain regions. Second are the MOS capacitors that are used to study the WF_{eff} of the gate stack. The advantage of these devices are that they are fabricated with a simplified process flow containing only the key process steps required to extract and study the WF_{eff}, thus significantly reducing the process steps and time. The process flow for these two types of devices is presented below:

1.3.1 MOS transistor Process flow

In order to study device reliability issues, we have utilized 14 nm FD-SOI MOS transistors. A transistor is required to study different reliability parameters, as these require to measure the channel current (Id-V_G curves). Figure 1.19 shows the various devices included in the FD-SOI technology offering multi-V_T solutions. This results due to a combination of two metal gates with two different WF_{eff} (N-type with La vs P-type with Al), together with two different Well types (N-Wells vs P-Well) and channels (Si vs SiGe).

Figure 1.20 shows the process flow consisting of the standard 14 nm FDSOI flow [39], starting from Silicon Trench Isolation (STI) and P-Well or N-Well implants carried out on SOI wafers, before HKMG stack deposition. The gate dielectrics consist of a 1.2 nm SiO_2 as IL, followed by a 2 nm thick HfON layer deposited by Atomic Layer Deposition (ALD) and decoupled plasma nitridation (DPN), as detailed in section 1.2.1. Then, metal gate stack is deposited following a sacrificial gate first approach of Figure 1.18. Devices were completed with spacers, N or P-type S/D formation, S/D dopant activation annealing at 1047 °C, NiPt silicide, contact trench formation,

back-end first level interconnects (M1) to allow electrical measurements and a 400 °C forming gas anneal.

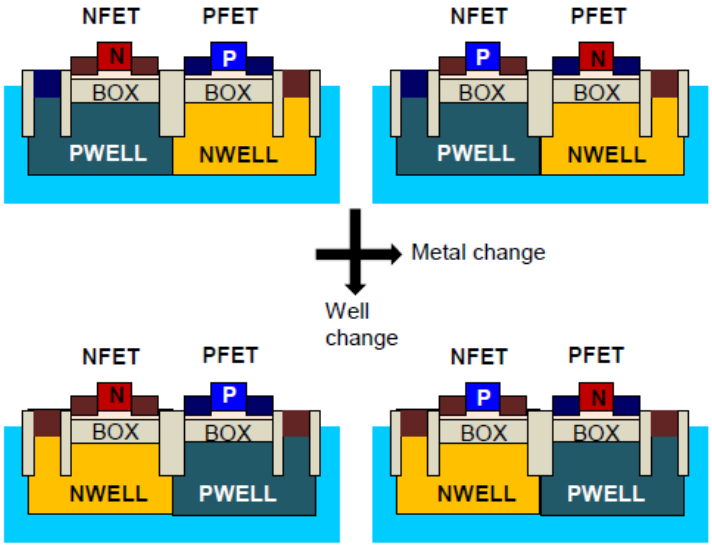


Figure 1.19: Schematics CMOS configuration to achieve 4 different V_T with two metal WF_{eff} (N & P types) and two WELL types (NWELL vs PWELL). Adapted from [32]

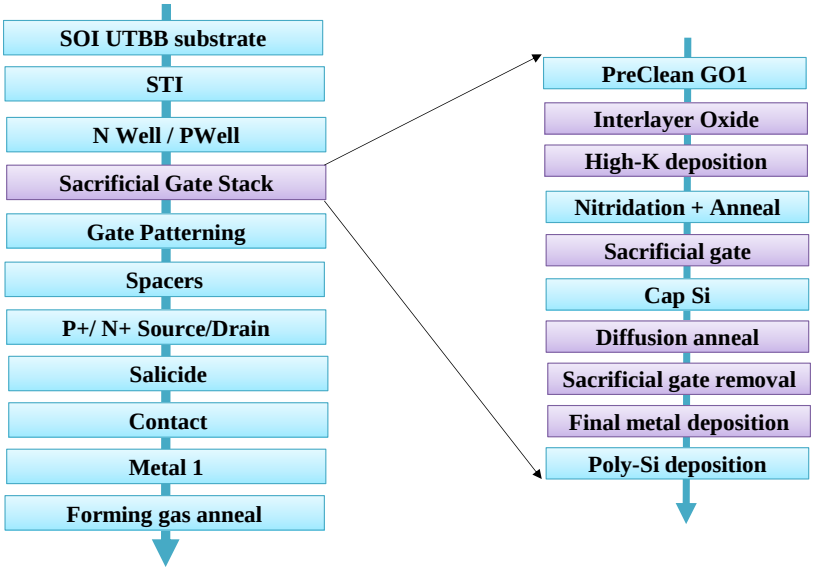


Figure 1.20: Integration scheme of the 14 nm FD-SOI MOSFET process flow [39]

1.3.2 Process flow simplification for MOS capacitor

The elementary MOSFET of the 14 or 28 nm FDSOI technology is the result of around hundreds of manufacturing steps. Process flow simplification is therefore mandatory in order to easily study the impact of the metal gate deposition processes and materials on key gate stack electrical parameters. Therefore in our studies, only the key steps which can influence the gate

stack physical or electrical properties, are kept in the process flow to reduce manufacturing steps and R&D cost.

Accordingly, MOS capacitors, fabricated on bulk Si substrate, are used to study the WF_{eff} of the HKMG stack from the capacitance voltage measurements (C-V). The simplified process flow is shown in Figure 1.21 and consists of P-Well or N-well implants carried out on silicon substrate. Then, Silicon Trench Isolation (STI) SiO_2 (500 nm) oxide is deposited and cavities are patterned by photolithography and etching. These cavities will act as the device active area. HKMG stack is deposited, consisting of 1.2 SiO_2 IL, HfSiON or HfON as high-k and metal gate stack deposited by the sacrificial gate first approach of Figure 1.18. Additional metal stack of Ti/TiN/W layers is deposited that act as the contact layer to allow electrical testing, followed by a 400 °C forming gas anneal. Finally chemical mechanical polishing (CMP) is performed in order to isolate the individual capacitors. In this way full capacitors containing the necessary intrinsic properties of the HKMG stack are fabricated. These devices will be used for the effective workfunction (WF_{eff}) studies from C-V measurements.

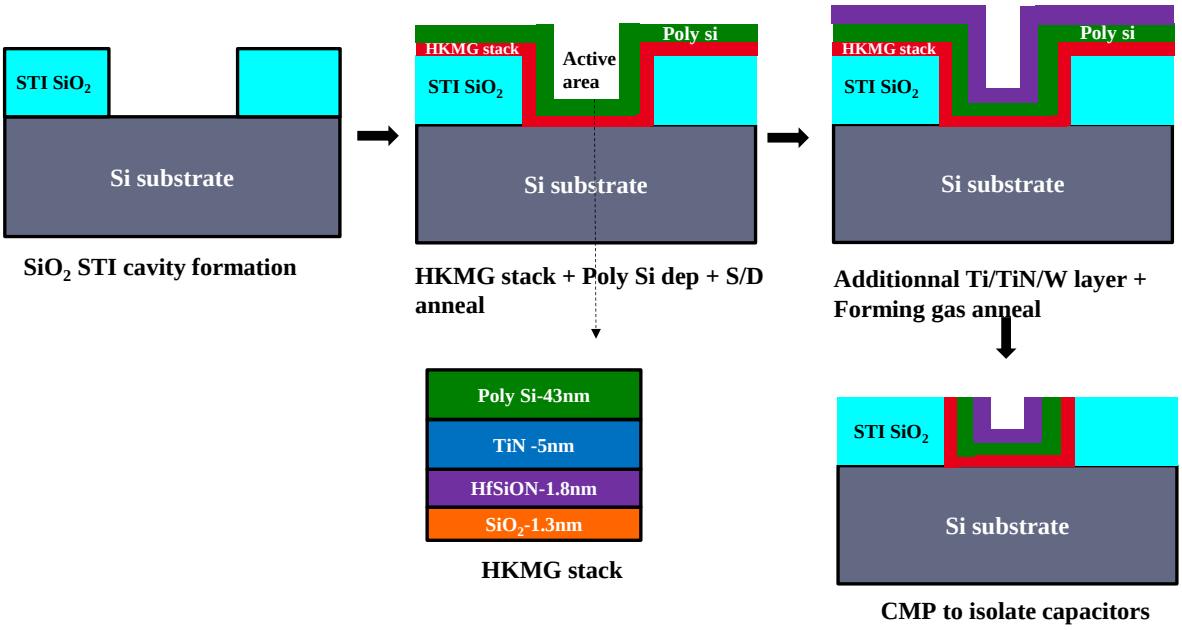


Figure 1.21: Simplified process flow of MOS capacitors used for WF_{eff} studies

1.3.3 Beveled Interlayer Oxide

Beveled oxide for gate stack IL is a powerful technique used to differentiate between the interfacial fixed charges and gate stack dipoles from the effective work function obtained by C-V measurement [17]. The bevel IL process results in thickness variation of the IL with wafer radius, as shown in Fig.1.22. Thus, we obtain nominal IL thickness (1.2nm) on the edges which is then close to the technology, and thicker at the center. As the IL thickness is changed, EOT is varied accordingly along the wafer radius, which will modulate the WF_{eff} (equation. 1.11).

The bevel process consists of a thermal oxidation to form 10 nm SiO_2, followed by a wet circular cleaning performed to form the bevel profile. Finally a 1.2 nm nominal IL is deposited on

top to obtain nominal IL thickness at the wafer edges. Figure 1.22 a shows the gate stack profile along the wafer radius. Figure 1.22 b compares the IL thicknesses on a nominal and bevel IL, measured by ellipsometry, and shows that the bevel IL thickness at the edges is the same as on a nominal IL process.

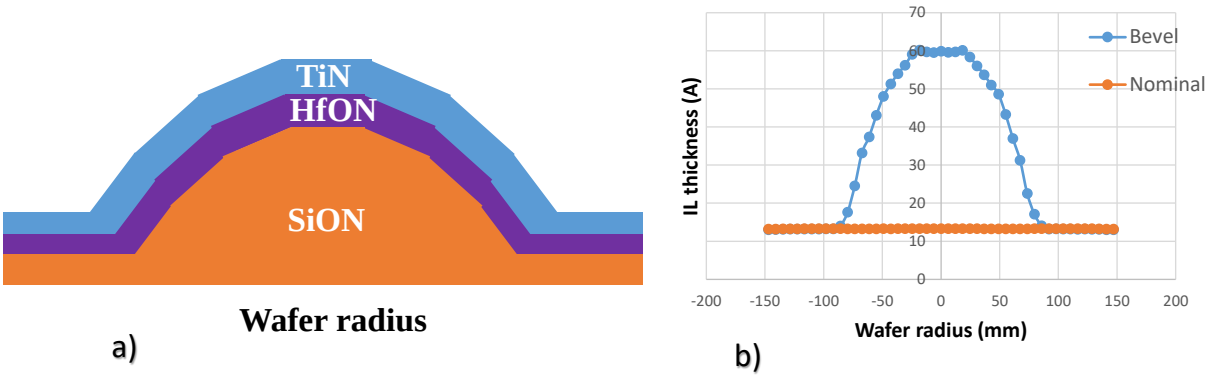


Figure 1.22: Schematic showing the variation of IL thickness along the wafer radius

1.4 MGG induced V_T variability

Metal gate has been introduced into advanced CMOS devices due to its advantages discussed in section 1.2.2, but it also brings in the metal gate granularity (MGG) parameter that could cause V_T variability. Under temperatures that are normally used in semiconductor device fabrication (see the process flow in section 1.3), metal gate grains could grow up to a few nanometers in size, having multiple grain orientations. Metal grains with different crystallographic orientations will have different work functions (WF_M) at the interface between the metal gate and high-K [56][57] (Figure 1.23).

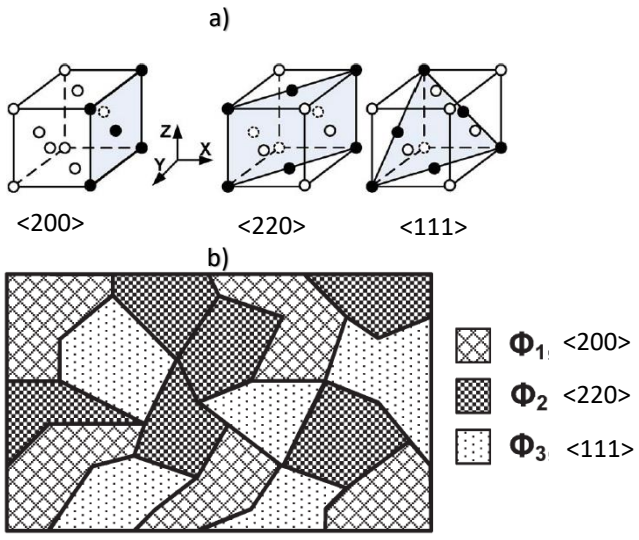


Figure 1.23: a) Cross-sectional view of surface planes along different orientations and b) Schematic of a hypothetical metal gate consisting of grains with three different orientations and hence, different WF_M values of Φ_1, Φ_2, and Φ_3 and occurrence probabilities of P1, P2, and P3, respectively [58]

The dependence of WF_M on the crystal orientation can be explained as follows: the electron density of the metal grain does not terminate at its surface but rather will spill outside. This creates a negative charge outside the metal surface, and so a positive charge inside the metal close to the surface Figure 1.24 b. This seperation of charges creates a dipole at the metal surface and increases the energy required to remove an electron from the metal, and so increases the WF_M [59]. The strength of this dipole depends on the surface atomic density of the crystal, which is different for different crystal orientations as shown in Figure 1.24 a for a FCC crystal.

a)

Orientation	Number of atoms	Area	Surface density
<111>	2	$\frac{\sqrt{3}}{2}a^2$	$\frac{4}{\sqrt{3}a^2} \approx \frac{2.31}{a^2}$
<100>	2	a^2	$\frac{2}{a^2}$
<110>	2	$\sqrt{2}a^2$	$\frac{\sqrt{2}}{a^2} \approx \frac{1.41}{a^2}$

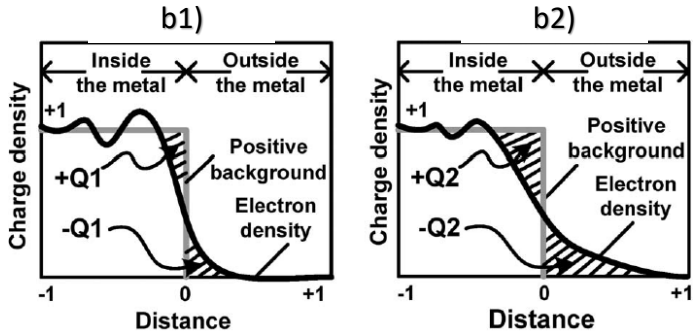


Figure 1.24: a) Surface density for different FCC crystal orientations [58]; b) Charge density at the surface of the metal crystal, for a less densely packed crystal (weak dipole-b1) and densely packed crystal (strong dipole-b2) [59]

Concerning the TiN crystal having a NaCl type structure, it can be thought of as made up of two interpenetrating FCC lattices of Ti and N atoms. The Ti atoms occupy the octahedral sites in the N sublattice, and the N occupy the octahedral sites in the Ti sublattice. Therefore, the surface of a <111> TiN lattice will be contain either just Ti atoms or N atoms, and so the Ti surface density becomes half of the value shown in Figure 1.24a. Ti Surface desnity of <100> or <200> will not change as all these planes will always contain the same Ti atoms. Thus the WF_M of TiN <200> crystals is higher than TiN <111>.

A transistor will contain only a small number of grains due to its small gate dimensions, that are in the range of few tens of nanometers. Moreover, during the metal gate growth, grain orientation of each grain will be determined randomly. As the WF_M of a grain depends on its orientation, the combined effect of randomness in grain orientation and low number of grains will cause the overall WF_M (and so V_T) of the fabricated metal gate to be a probabilistic distribution rather than a definitive value [58]. It has been shown that the distribution of the WF_M is described by a Gaussian distribution [58]. As a result, the standard deviation of the WF_M distribution is inversely

proportional to the number of grains. The number of grains increase with the area of the metal gate and hence the WF_M variability increases as the device area decreases. This will cause variation of WF_M or V_T between devices undergoing same manufacturing process and so called as V_T variability induced by MGG. This variability has been well explained and correlated to the metal gate microstructure (grain size and orientation) and gate area both experimentally [60][61] and by modelling [62][63].

V_T variability is one of the most critical challenges for future CMOS technology nodes [64][65]. It originates mainly from the contributions of the random dopant fluctuations (RDF), line edge roughness (LER), oxide thickness fluctuations (OTF) and MGG. With the introduction of the FDSOI architecture, that can tolerate an undoped or lightly doped channel, V_T variability due to the RDF contribution is dramatically reduced [66][67][68]. Even though the RDF component is dramatically reduced, other sources of variability such as LER and MGG still remains important. Moreover, these contributions could become even more important as the device dimentions decrease, shown in Figure 1.25.

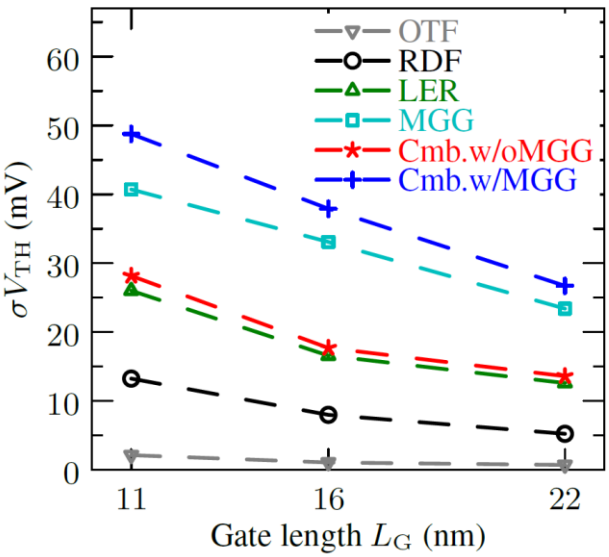


Figure 1.25: Variations of the different contributions to V_T variability obtained by simulations, showing random dopant fluctuations (RDF), oxide thickness fluctuations (OTF), line edge roughness (LER), and metal gate granularity (MGG) [68]

In this work, TiN is used as the gate material with two main orientations: <111> and <200> having WF_M of 4.4 and 4.6 respectively [69]. Chapter 2 covers the X-Ray Diffraction technique used to characterize the microstructure of the TiN gate. In chapter 3, variations in the RF-PVD TiN process parameters (section 1.2.2) have been proposed with the aim of modulating its microstructure that can have an impact on MOSFET V_T variability.

1.5 MOSFET reliability

Reliability is defined as the probability that a device performs a given function, under specified conditions for a given period of time. Due to aggressive scaling in device dimensions for performance improvement, advanced CMOS transistors in the nanometer range have resulted in major reliability issues due to increased electric fields, device working temperatures and introduction of the HKMG stack. These include hot carrier injection (HCI), stress induced leakage current (SILC), time dependent oxide breakdown (TDDB), and bias temperature instability (BTI). These reliability mechanisms cause a shift in MOS transistor parameters, such as: threshold voltage V_T, transconductance and channel mobility. For reliability testing, the degradation phenomena is accelerated by putting the device in extreme conditions of thermal and electrical constraints that are well beyond normal conditions of use. It is then possible to extrapolate the degradation caused at these extreme conditions, to determine device lifetime under normal conditions of use. In this thesis we have focused on MOSFET's BTI and TDDB phenomena and how they are modulated by process parameters of the gate stack. A detailed discussion on the BTI and TDDB mechanisms is presented below:

1.5.1 Bias temperature instability (BTI)

BTI refers to the drift in transistor electrical parameters during its operation. It is caused by the charging of defect states in the gate oxides and at its interface [70]. The defects could be both pre-existing and generated during device operation. The trapped charges in these defects result in a shift of device parameters, such as its drain current, threshold voltage V_T, channel mobility, transconductance, and subthreshold slope. It is driven by gate bias and occurs even at low source-drain voltage V_{DS}. Moreover, its mechanism is strongly accelerated by temperature. BTI in n-MOSFET, which are positively biased in circuits, is referred to as Positive Bias Temperature Instability (PBTI), while Negative Bias Temperature Instability (NBTI) takes place in p-MOSFETs that are negatively biased.

Negative Bias Temperature Instability (NBTI)

NBTI as the shift of PMOS transistor threshold voltage V_T and other performance parameters has been known since 1966 [71][72]. Only during the last few years, NBTI became one of the largest CMOS reliability concern due to the introduction of new materials, increased electric fields and operating temperatures in the transistors [73][74][75]. The main characteristics of NBTI are as follows:

- It results in an increase of the absolute V_T and decrease of the drain current (I_d) and transconductance (G_m) with stress time. It occurs when negative bias is applied to the gate

(PMOS inversion state). Figure 1.26 and Figure 1.27 shows device biasing condition during NBTI stress and shift of these parameters with stress time.

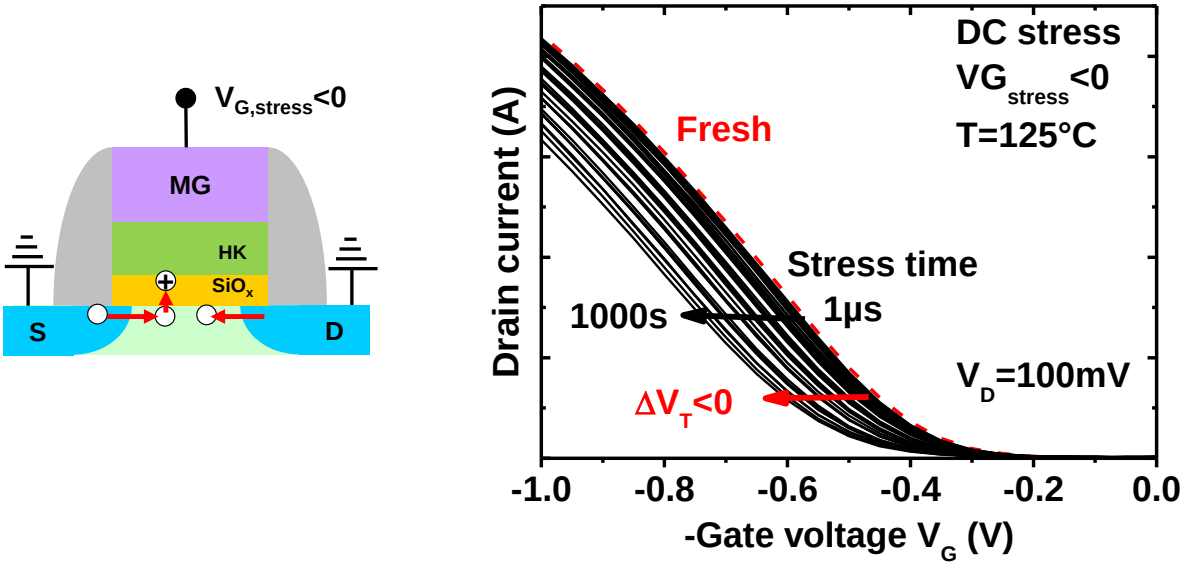


Figure 1.26: a) Schematic showing PMOS device under NBTI stress; b) Variation of drain current versus gate voltage curves with NBTI stress time

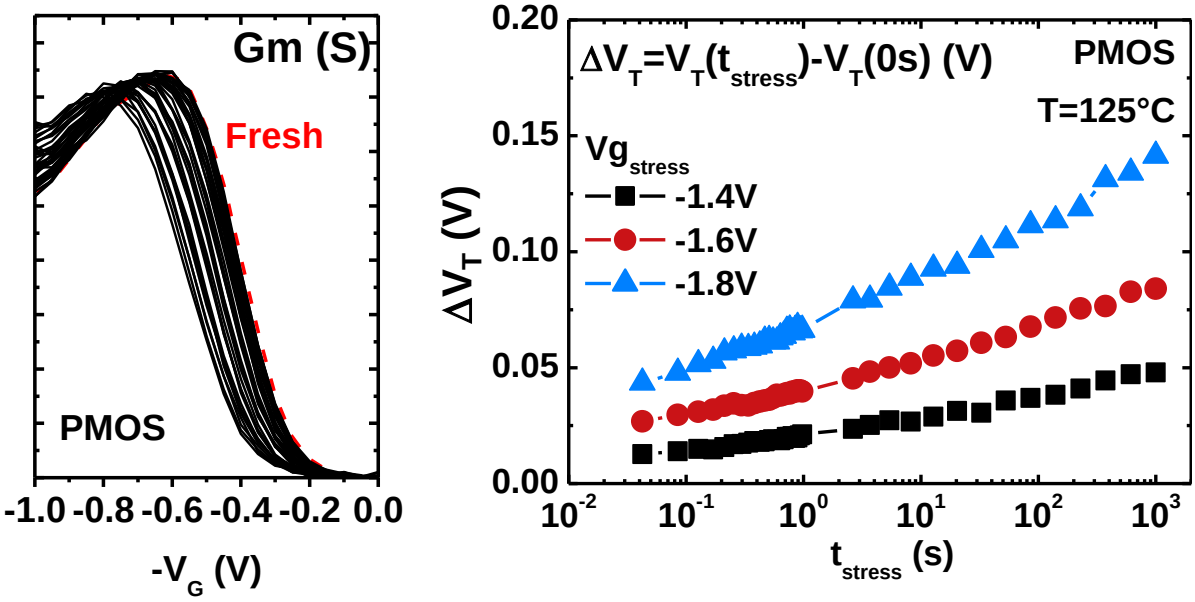


Figure 1.27: a) Variation of transconductance (G_m) versus gate voltage curves with NBTI stress time; b) Drift of V_T with stress time at different stress voltages, showing its power law dependence

- The shift in V_T (ΔV_T) with stress time follows a power law and is strongly activated by gate bias and temperature (T), as shown in Figure 1.27 b. The power law is as follows:

$$\Delta V_T(t, V_G, T) = A.\eta.V_G^{\gamma} e^{-\frac{E_a}{kT}} t^n \quad (1.14)$$

where t is the stress time and n is the power law time exponent.

- It is attributed to the creation of positive charge trapping at the SiO_2/Si interface (qN_{it}) and in the interlayer gate oxide (qN_{ot}). Thus, the threshold voltage shifts due to the formation of these charges in the gate stack:

$$\Delta V_T = -(qN_{it}+qN_{ot})/C_{OX} \quad (1.15)$$

- Interface traps N_{it} can be Si dangling bonds that are formed from the breaking of Si-H bonds at the SiO_2/Si interface due to a combined effect of electric field, temperature and holes. These traps are distributed in the Si band gap and are usually of the acceptor-type in the upper half and donor-type in the lower half of the Si band gap. Figure 1.28 a, shows the energy band diagram of a PMOS device at the flatband condition, with electrons occupying states below the Fermi energy E_F. A this condition, the states in the lower half of the band gap and those above E_F are neutral, and the states between midgap and the Fermi energy are negatively charged. Figure 1.28 b, shows the energy band diagram of a PMOS device in the inversion condition. At this condition, the interface traps between mid-gap and the Fermi level will be positively charged (called as unoccupied donors). Thus, interface traps in a PMOS device in the inversion regime will be positively charged, leading to negative threshold voltage shifts [76]. Application of a negative bias stress generates donor states in the lower half of the band gap [74][77]. When these stress generated interface states are charged and discharged, drifts in PMOSFET electrical characteristics occur: $\Delta V_T = -qN_{it}(t)$. Figure 1.29 shows the increase of PMOS V_T and N_{it} with stress time, for negative bias applied to the device [70]. It evidences the relationship between ΔV_T and N_{it}.

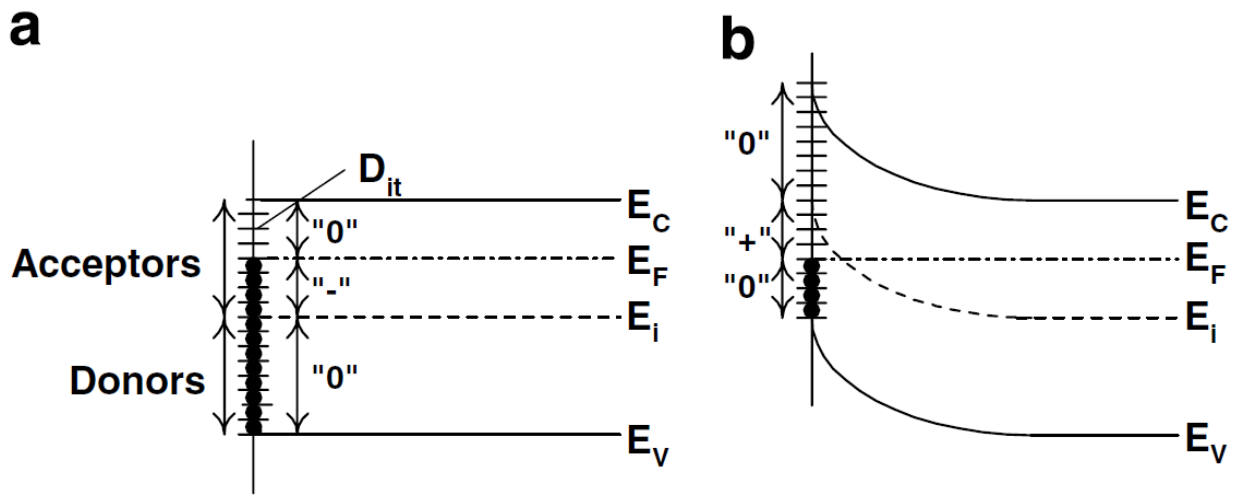


Figure 1.28: Band diagrams of the Si substrate of a PMOS device showing the occupancy of interface traps a) negative interface trap charge at flatband and b) positive interface trap charge at inversion [73]

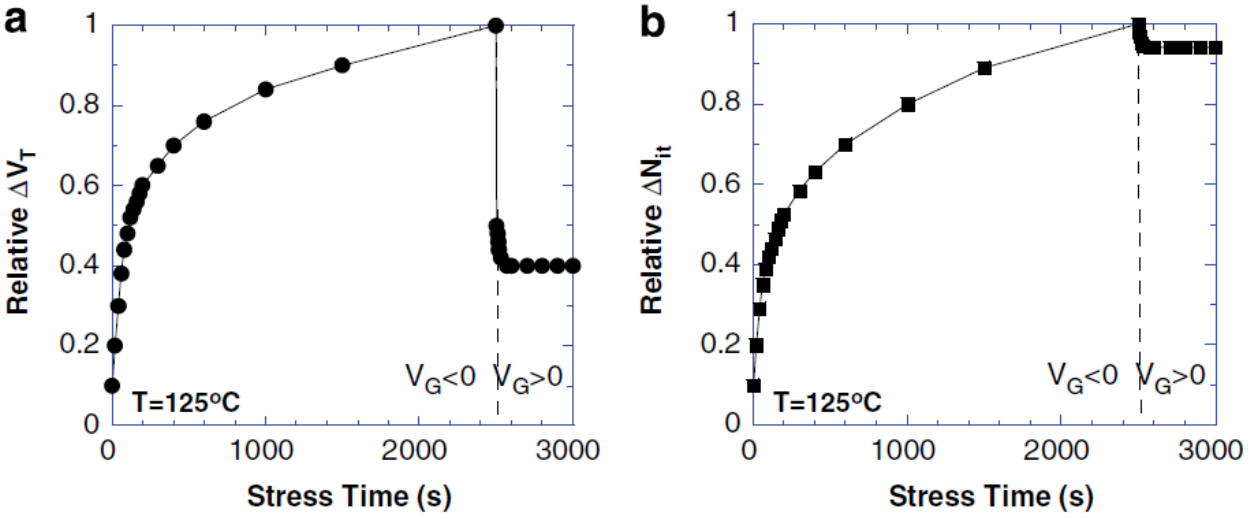


Figure 1.29: Relative shifts for a) V_T and b) N_{it} versus stress time for negative and positive gate voltages [70]

- Oxide charges qN_{ot} is the second contributor to the V_T shift (apart from N_{it}) and their role on NBTI is not entirely clear. In some studies it was believed to be H+ ions coming from breaking of Si-H bonds and getting trapped in the oxide near the SiO_2/Si interface. However, recent studies have shown that these positive charges are formed due to hole trapping [70][78][79]. Precursors for this hole trapping might be present before the stress is applied and gets occupied by holes during stress.

- Some of the NBTI defects (or V_T shift) can be recovered or annealed by removing the stress conditions or by applying a positive bias (called the recoverable part), the rest is the permanent damage (called the non-recoverable part). Huard et al. [70] have shown that the non-recoverable part is largely due to the trapping in interface states that are created during the stress, as they remains almost constant after recovery (Figure 1.29 b). They attributed

the recoverable part to the formation of oxide charges due to hole trapping (the saturation of V_T in Figure 1.29 a).

- NBTI is very sensitive to the interlayer oxide and to the quality of its interface with the Si substrate. Gate oxide and interface process parameters and materials affect NBTI degradation, such as presence of Nitrogen, Fluorine, Boron and water.

The mechanism of NBTI is still not completely clear and deviations from the perfect power law have been observed. Several models and their improvements have been proposed, here these will discussed briefly.

Reaction-Diffusion Model

Reaction–Diffusion (R–D) model was the earliest of models proposed for NBTI [80][81][82]. The mechanism is schematized in Figure 1.30. It is caused by the breaking of Si-H bonds at the SiO_2/Si interface and release of hydrogen, due to temperature and electric field (the reaction phase). In the subsequent diffusion phase, hydrogen diffuses from the interface into the oxide. This results in the formation of dangling bonds at this interface and are termed as the interface traps N_{it}. The model assumes classical (Gaussian) diffusion of neutral hydrogen in the gate oxide as the rate-limiting step. Gaussian distribution means that the allowed hydrogen states in the oxide bulk, during diffusion, has the same energy and thus the same hopping time for hydrogen diffusion. Solving the rational diffusion equation gives:

$$N_{it}(t) = A\, t^{0.25} \qquad 1.16$$

where, t is the stress time and A is a pre-factor which depends on the electric field and temperature.

This model leads to a power law dependence that is observed experimentally, but only considers the part of ΔV_T associated to the interface states and does not take into account the contribution of positive oxide charges formed during stress. Moreover it gives a constant time exponent of 0.25, but deviations from this constant exponent have been observed experimentally. The models presented next will deal with these issues

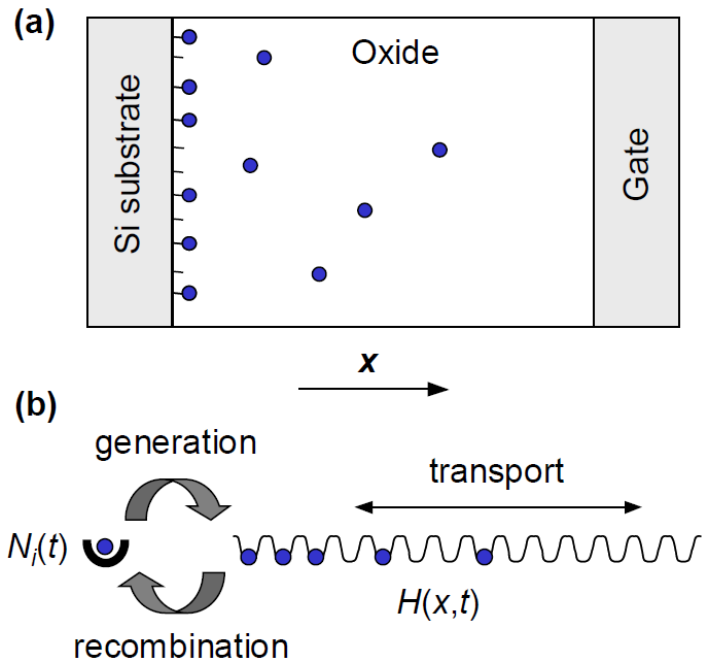


Figure 1.30: Reaction-diffusion model of NBTI: a) Schematic view of the MOS stack showing N_{it} generation; b) The physical mechanism involved in hydrogen generation and its transport [83]

Disorder controlled Kinetics Model

The classical reaction diffusion model presented above assumes a Gaussian diffusion in the MOS oxide layers. Studies have shown that the particle kinetics in a disordered system such as SiO_2 should be considered as dispersive instead of Gaussian. Energy of the localized hydrogen states in the bulk of the gate oxide are dispersed with a density-of-states (DOS) g(E) and thus will result in a wide distribution of hopping times of hydrogen. Thus, these localized states will affect transport of hydrogen through the oxide bulk [80][83][84] (Figure 1.31).

Solving the diffusion equation with assumption of dispersive transport gives:

$$N_{it}(t) = A\, t^{\alpha/4} \quad , \alpha = K_b T/E_0 \qquad 1.17$$

where α is the dispersion parameter, E_0 is characteristic DOS width, T is the temperature and A is the pre-factor depending on the electric field, temperature and dispersive parameters.

Kaczer et al [83] showed that the assumption of dispersive transport allows to explain a number of experimental observations that cannot be otherwise accounted for by the classical Reaction-Diffusion model. These include the temperature and material dependence of NBTI time exponent, log (t)-like dependence for NBTI recovery and the ability to link the NBTI power-law exponent

with the NBTI relaxation rate. The model was also extended to the transport of charged particles (Hydrogen+).

This model was able to explain the variation of the time exponent but is still limited to the interface states contribution and ignores the one from oxide charges.

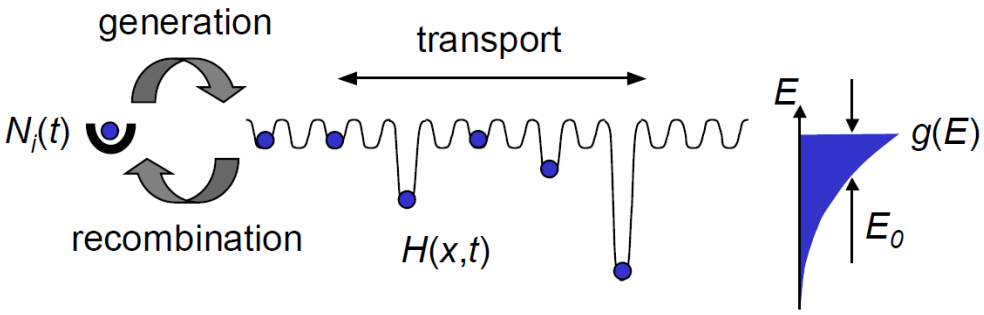


Figure 1.31: Schematic representation of the NBTI model with Disorder-controlled kinetic of hydrogen [83]

Hole trapping model

Huard et al. [70] [85], Grasser et al. [86] and Kaczer et al. [87] have shown that apart from the interface states, positive oxide charges contribute to the threshold voltage shift and have related it to the hole trapping mechanism. However, the temperature and voltage dependence of the recovery behavior, observed correlation between interface states and oxide charges [88] and, lack of characteristic time scale in NBTI recovery could not be explained [78].

NBTI recovery lacks a characteristic time scale and rather follows a log behavior in time. Grasser et al. [78] and Kaczer et al. [79] linked this lack of characteristic time scale in NBTI recovery with the low frequency noise which has a response on many timescales and is caused by carrier trapping with many time scales. This similarity between NBTI recovery and low frequency noise might be an indication that the same kind of defects are responsible for NBTI degradation. Distance of traps from the SiO_2/Si interface determines the tunnel time and hence the low frequency noise response. In these models, the authors have assumed that holes can be captured via a multiphonon emission (MPE) process into deep near-interfacial layer traps, which are probably oxygen vacancies (E0 centers) [89]. This mechanism is characterized by its exponential electric field dependence and a strong temperature activation. In this model, defect creation proceeds via a two stage process when stress is applied to the device, as shown in Figure 1.32. In the first stage, holes can be trapped into near-interfacial oxygen vacancies via the MPE mechanism. In the second stage, once positively charged (state 2), the E' center can attract the H from the interface and thus reaching to state 4. This step locks in the E0 center and creates a dangling bond at the interface, which are filled by charges that depends on the Fermi-level of the substrate [79].

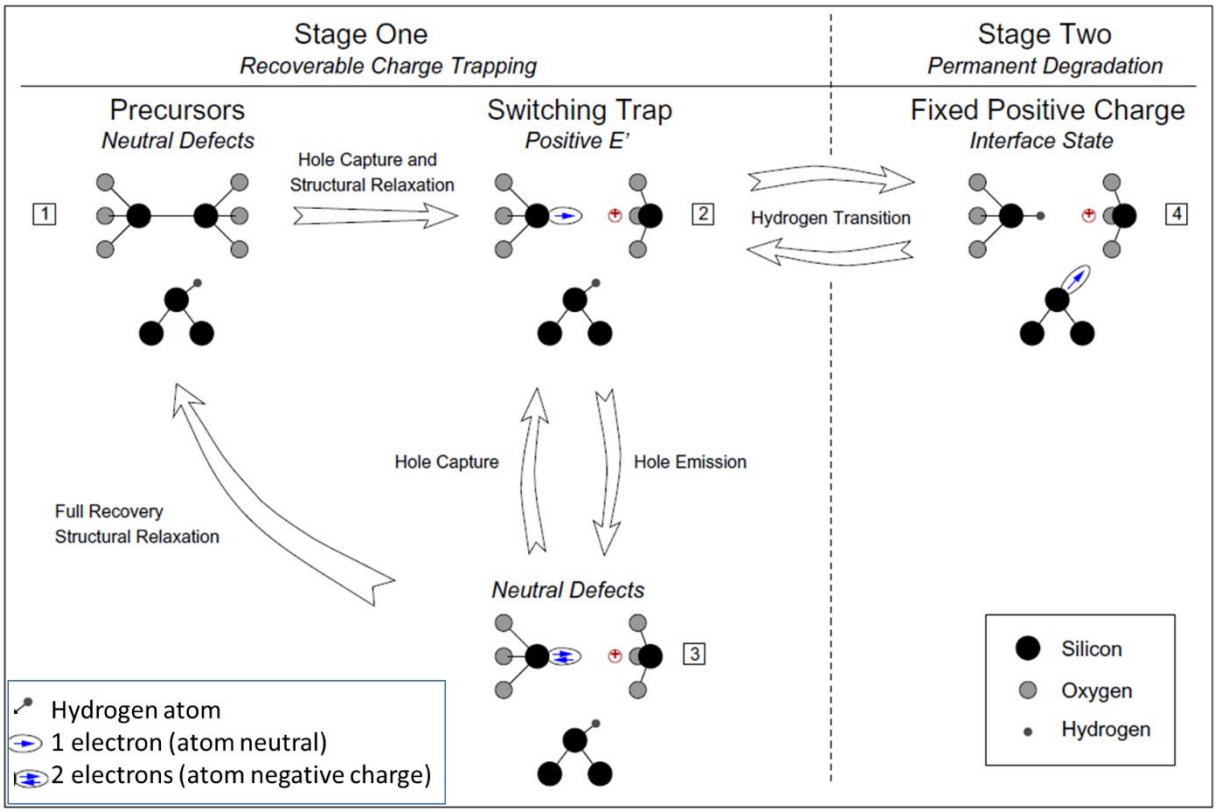


Figure 1.32: Schematic explaining the model for switching oxide trap coupled to the creation of a dangling bond at the interface [78]

So the V_T shift due to NBTI stress is composed of two traps: first, hole traps (metastable states) which recover, and interface states which are stable (no recovery). The NBTI recovery corresponds to holes de-trapping and the saturation level corresponds to interface traps, as found experimentally by Huard et al [70]. In conclusion, this model is able to provide a mechanism for the experimental observations of NBTI.

Recently Grasser et al. [90] have updated their model that is presented above. In this work the authors suggest that the H diffusing from the gate may create border traps and hence influence the hole capture and interface traps creation mechanism.

Positive Bias Temperature Instability PBTI

PBTI is the V_T shift phenomena taking place on N-MOSFET devices under positive stress (Figure 1.33). The shift in V_T (ΔV_T) with stress time follows a power law in time and is strongly activated by gate bias, and like NBTI follows the power law in Eq. 1.14. It is attributed to trapping mechanism in localized defects in the high-k oxide. This trapping phenomena is enhanced by electric field, temperature and by the high-k thickness. It has been attributed to oxygen vacancies in the high-k, by identifying the energy position of these defects [91][92]. Some of these traps are reversible, that can be de-trapped, and other are permanent. Moreover, during stress at high voltages, an opposite shift in V_T (ΔV_T <0) has been observed and is termed as the turn-around

effect [93]. It might be caused by the generation of positive charges in the high-k by the electrons that are injected from the substrate at high Vg [94]. The shift of V_T under PBTI stress varies very fast in a high-k oxide, compared to SiO_2, due to higher defects in the high-k. Thus, PBTI has become one of the most critical problems in the integration of high-k in the MOS stack.

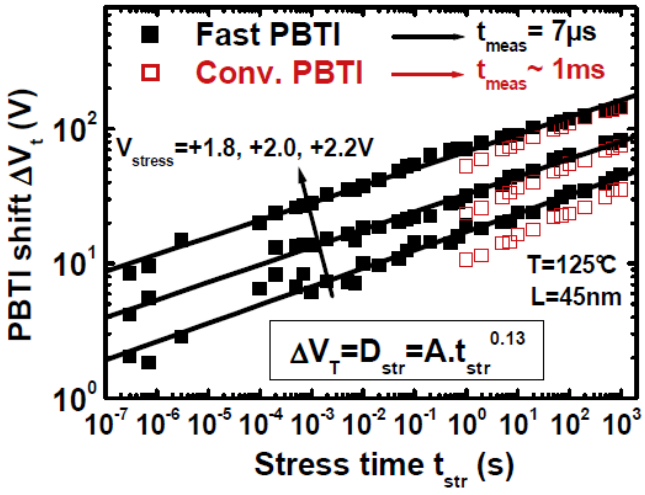


Figure 1.33: V_T shift induced by PBTI stress evaluated by the fast and the conventional techniques [95]

Bersuker et al [92] proposed a trapping model in which the pre-existing traps can be filled via several processes, as illustrated in Figure 1.34 a, in two stages. A first quick tunneling process (Pc) where the trapping characteristic is in the order of microseconds due to the smaller tunneling barrier and higher density of states in the Si conduction band. Second process (P_T) is the transport of electrons between traps, activated by temperature and is of Poole-Frenkel type. Figure 1.34 b compares the experimental and modeled PBTI V_T shift, and a close correlation is observed.

PBTI shift in high-k dielectrics can be considered as the sum of a reversible component D_{rev} that can be fully recovered by applying a negative bias after stress, and an irreversible component D_{irrev} that is permanent. D_{rev} is caused by 1) pre-existing fast and slow traps, where fast traps are temperature independent (1 in Figure 1.35.); 2) stress induced slow traps that are strongly activated by temperature (2 in Figure 1.35). D_{irrev} are caused due to stress generated electron traps that cannot be discharged and can be considered as fixed negative charges (3 in Figure 1.35) [95][94]. In the case of thick HfO_2 with a poly Si gate, PBTI is mainly caused by pre-existing defects and the generation of defects due to stress is relatively low [96]. Garros et al. found that for the case of HfSiON with La additive [95], trapping in pre-existing fast traps is negligible by comparing fast and conventional measurement techniques as shown in Figure 1.33. Mitard et al.[94] and Ribes et al. [97] gave the assumption that the defects responsible for electron trapping in High-k are the negative "U-traps" [98].

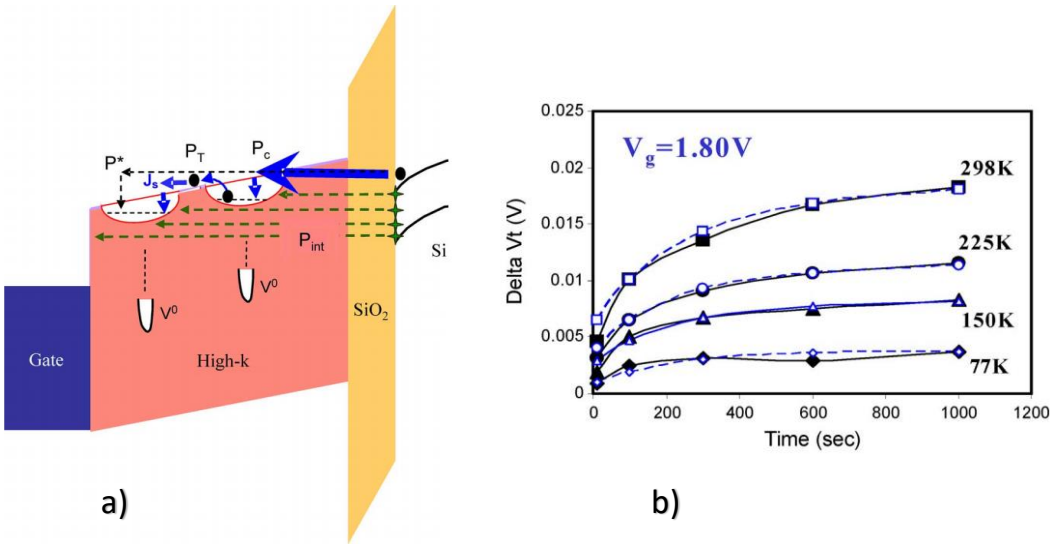


Figure 1.34: a) Schematic of energy band diagram with the electron trapping processes responsible for PBTI and b) experiment and modeled (open symbols) V_T shift during the PBTI stresses at different temperatures [92]

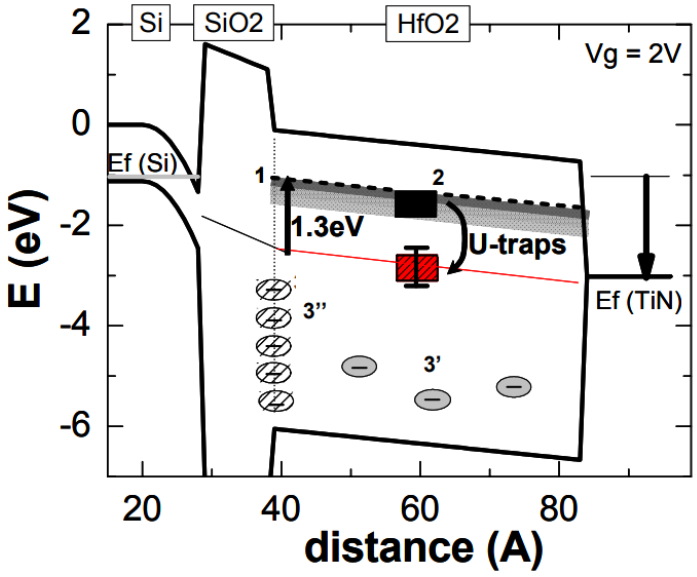


Figure 1.35: Energy band diagram of the HKMG stack showing different traps responsible for PBTI [94]

1.5.2 Time dependent oxide breakdown (TDDB)

TDDB is the gate oxide breakdown phenomena occurring in both N-MOSFET and P-MOSFET devices. The gate oxide breaks down as it loses its insulating properties due to the application of gate voltage for long durations of time. From an electrical point of view this breakdown means a strong increase in gate oxide tunneling current due to the formation of a conducting path through the gate oxide and into the substrate. The conduction path is formed by defects generated during electrical stress.

Figure 1.36a and b shows device biasing for TDDB tests and the variation of leakage current with stress time respectively, for a NMOS device. In the early stages of stress, the gate current

increases slightly with time due to generation of new defects in the gate oxide and this behavior is termed as stress induced leakage current (SILC). As the stress is continued, a strong increase (or sometimes decrease) in current is observed and is termed as a breakdown event. This event can be a soft breakdown (SBD) which is caused by formation of a weak conducting path containing only a few traps, or can be a hard breakdown (HBD) which is caused by a complete defect formation path, when the number of defects reaches a critical value. Depending on the tolerance in device performance, device failure or time to breakdown can be defined from either a SBD or HBD event. Both SBD and HBD are localized and randomly distributed all over the device area, and are spatially uncorrelated to each other [99][100]. In our studies, the time to breakdown (Tbd) is measured by the occurrence of a HBD breakdown event.

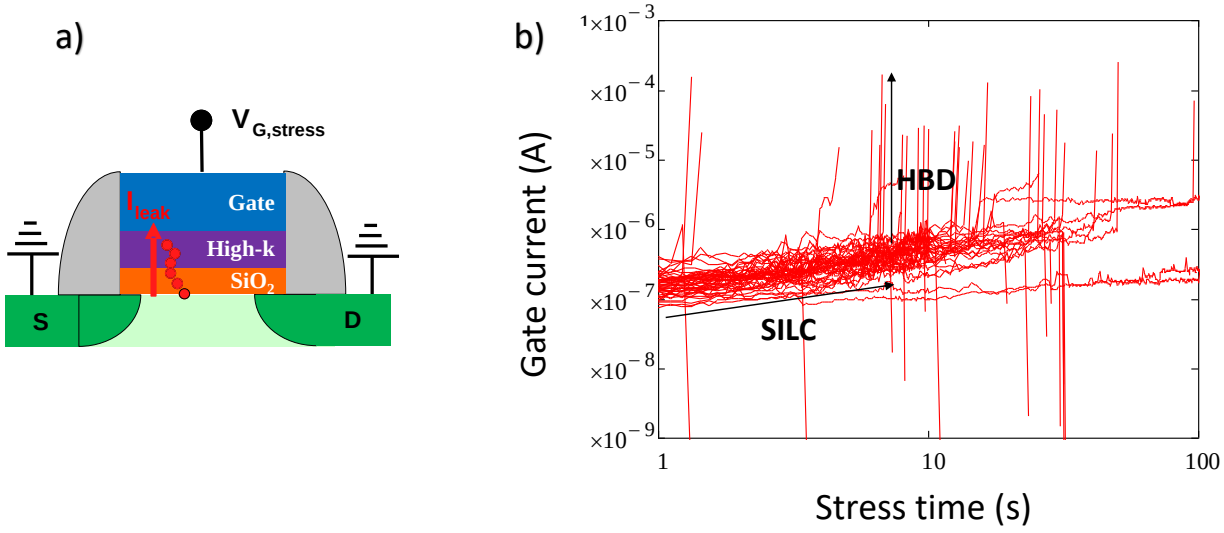


Figure 1.36: a) Device biasing during a TDDB measurement; b) Increase of leakage current during TDDB stress, showing SILC regime and then the hard breakdown

The charge to breakdown (Q_{bd}) is defined as the total charge injected into the oxide until breakdown occurs. It's given by the following relation:

$$Qbd = \int_0^{Tbdi} J(t) \cdot dt = Tbdi \cdot average(J) \qquad 1.18$$

where J(t) is the leakage current density and Tbdi is the time to breakdown of a device i among a sample.

In Figure 1.36 , it can be observed that the time to breakdown are relatively dispersed over time and thus it becomes a statistical phenomenon. In general, the Weibull distribution is considered as the most widely accepted distribution function that can describe the MOSFET time to breakdown. Studies conducted by Wu et al. [101] on about 1000 samples, have justified the use of Weibull distribution (Figure 1.37). Weibull scale W is given by:

$$W(F(t)) = \ln(-\ln(1-F)), \qquad W(F) = -\beta\ln(\eta) + \beta\ln(t) \tag{1.19}$$

where F(t) is the distribution function representing the cumulation of samples broken until stress time t; η is the characteristic lifetime for 63% breakdown of the sample in the distribution (t63% or T_{bd}). So, the plot of W(F) as a function of ln(t) is a line of slope "β" and ordinate "-β.ln (η)". Figure 1.36 represents the Weibull scale values versus time (the respective failure percentage are also mentioned).

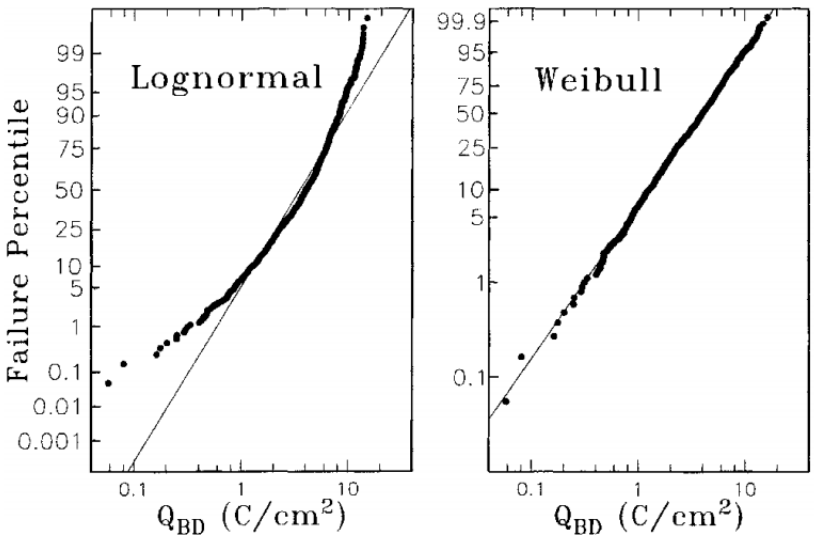


Figure 1.37: Log-normal and Weibull distribution fit of the experimental data [101]

The breakdown phenomena is mainly described by the percolation model presented below:

Percolation model

The first proposal of the idea that gate oxide breakdown is the result of a conduction path of defects, formed in the gate oxide, was given by Suñé et al [102]. Then the percolation model was updated and generalized to three dimensions by Degraeve [103]. By using this percolation model approach, it is possible to show that the breakdown phenomena follows the Weibull probability law defined in Eq. 1.19. Finally, this concept of percolation became generalized and underwent various improvements [104][105].

Here we will describe a simplified description of the percolation model with the aim to understand its principle, which can lead to the formation of breakdown times that follows the Weibull distribution [105]. This description, illustrated in Figure 1.38, models the dielectric as discrete cells of size a_0 (characteristic dimension of defects), each cell having a probability to

become defective during stress. Thus, by random events, the cells will generate defects during stress until a percolation path is generated and give rise to breakdown.

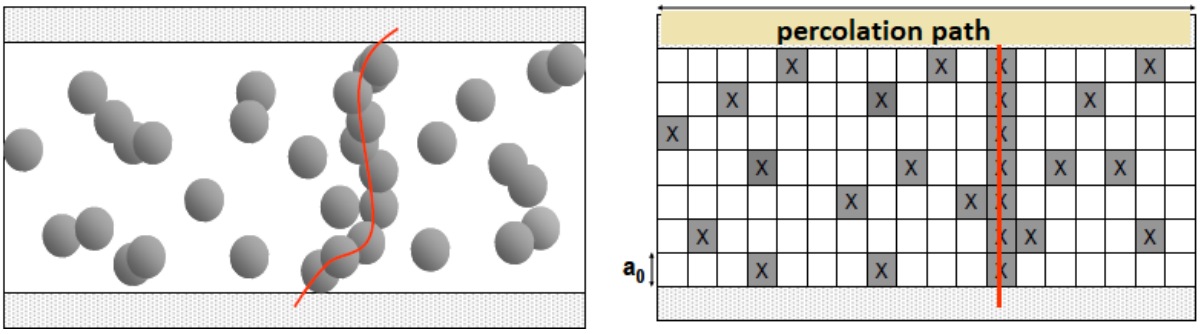


Figure 1.38: Schematic diagram illustrating the percolation model. It consists of dividing the oxide in several cells which will generate defects and thus forming a percolation path [105]

From this model a Weibull distribution is obtained as in Eq. 1.19. The values of Weibull parameters β and η are related to the device properties as:

$$\beta = (\alpha/a_0)\, t_{ox} \quad \text{and} \quad \eta \propto \exp(-\ln(A_{ox}/a_0^2)/\beta) \tag{1.20}$$

where a_0 is the defect size, t_{ox} is oxide thickness, A_{ox} is the oxide surface area and α is a typical parameter of the oxide.

The charge to breakdown Q_{bd} (corresponding to T_{bd}), the defect generation probability (λ) and critical defect density for breakdown (N_{bd}) are related by the relation:

$$Q_{bd} = N_{bd}/\lambda \tag{1.21}$$

Many studies has proposed different values for these breakdown parameters. The model parameters differ from one study to another [105] [106] [107] [108] [109] [110] [111].

Breakdown models

The thermochemical model (E-Model): According to this model, oxide breakdown due to generation of defects is governed by the oxide field. The bonds in the oxide can break under the electric field influence, thus the defects density can reach its critical value for breakdown (N_{BD} in Eq. 1.21) [112]. This model predicts an exponential behavior in electric field for the time to breakdown.

The Anode hole injection model (1/E model): According to this model, electrons ejected from the metal gate reach the anode and generate hot holes through impact ionization. These holes then tunnel through the gate oxide (meaning electrons tunnel from gate oxide to the substrate) and generate defects which are responsible for oxide breakdown [113]. This model predicts an exponential behavior in (1/electric field) for the time to breakdown.

The hydrogen release model: The hypothesis that hydrogen atoms released during stress are at the origin of breakdown, was proposed in the case of thin oxides [114]. Figure 1.39 illustrates the principle of this mechanism. It begins with electron injection from the gate, then they transfer their energy to Si-H bonds at the SiO_2/Si interface and release of hydrogen species (H^0 or H^+) takes place. These species diffuses through SiO_2 and creates defects in the oxide.

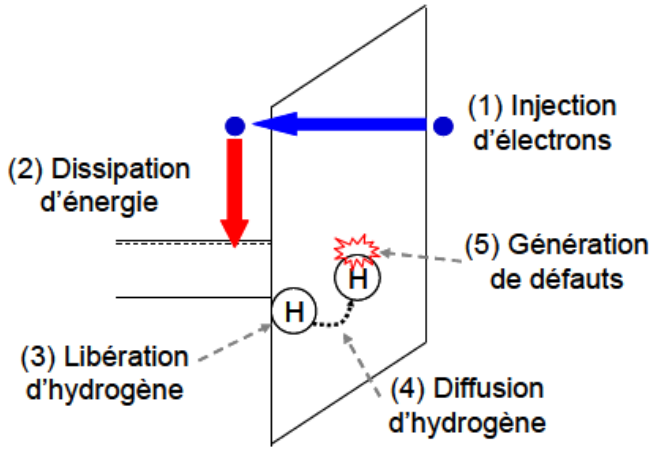


Figure 1.39: Schematic diagram of the hydrogen release breakdown model [115]

Studies based on SILC measurements have shown a correlation between hydrogen release and oxide breakdown [110][116]. This is particularly the case of thin oxides where these models are based on multi-vibrational excitation. Ribes et al. proposed that only one electron is not sufficient to break the Si-H bond by direct excitation, but a cooperation between them is required [97]. They linked this assumption to the multi-vibrational excitation of Si-H bonds by electrons [116] [117]. Thus the defect generation probability (λ) (in Eq. 1.21) is linked to the oxide current density and it becomes:

$$\lambda \propto \left(\tau_e^{\frac{E_d}{n\hbar\omega}-1} \cdot \frac{E_d}{n\hbar\omega} \right) \cdot E^{K.\frac{E_d}{n\hbar\omega}} \cdot \left(\frac{I}{e} \right)^{\frac{E_d}{n\hbar\omega}-1} \qquad 1.22$$

where K = 6 for holes and K = 4 for electrons [97], with τ_e the lifetime of excitation, E_d the dissociation energy of the bond and nħω the energy supplied by an electron. The excitation current I and the energy of the carriers E depend on the operating mode of the transistor. Accordingly Q_{bd} depends on the square of the current density ($Q_{bd} \propto J^{-2}$) [88]. This model is termed as multi-vibrational hydrogen release (MVHR).

For HKMG stacks (SiO_2/high-k), the mechanisms can be quite different depending upon the injection regimes as shown in Figure 1.40:

a) Injection of carriers from the substrate (carriers see the high-k/metal gate interface at the oxide output/exit)

b) Injection of carriers from the metal gate (carriers see the SiO_2/Si interface at the oxide output/exit).

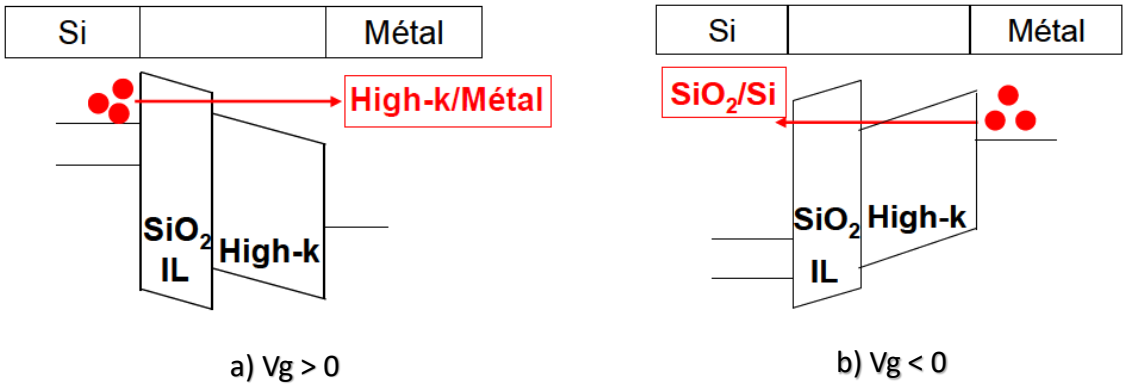


Figure 1.40: HKMG stack band diagrams in the case of the (a) injection regime from the substrate and (b) injection from the metal gate [115]

Model in case of injection from the gate

The MVHR model is normally used to explain the degradation of SiO_2 under electrical stress. It can also be applied to the breakdown of oxides based on SiO_2/high-k stack during injection mode (PMOS inversion) from the gate, as the electrons that cross the oxide see the Si/SiO_2 interface. Indeed, the validity of this model was confirmed for the SiO_2/high-k stack by Ribes et al [97]. The authors also show that the high-k does not affect the N_{bd} but only the λ, and that the breakdown of the HKMG stack was governed by the charge to breakdown (or current) and the breakdown of the interfacial layer.

Model in case of injection from the substrate

Carrier injection from the substrate applies to a NMOS transistor under positive bias (inversion regime). Degradation in this case is caused mainly by the gate current, consisting of electrons, injected from the conduction band of the substrate towards the gate, thus degrading the gate oxides. In the beginning, it was reported that device breakdown in this case is governed by the degradation of the interfacial layer, like in the case of injection from the gate. According to Rafik [115] the activation energy and voltage acceleration factors extracted for HfSiON were similar to as observed for SiO_2 and as provided by the MVHR model (Figure 1.41), but were quite different for HfO_2 high-k (Figure 1.41). Hence it might be possible that the MVHR model is valid for substrate injection regime in the case of HfSiON but not for HfO_2.

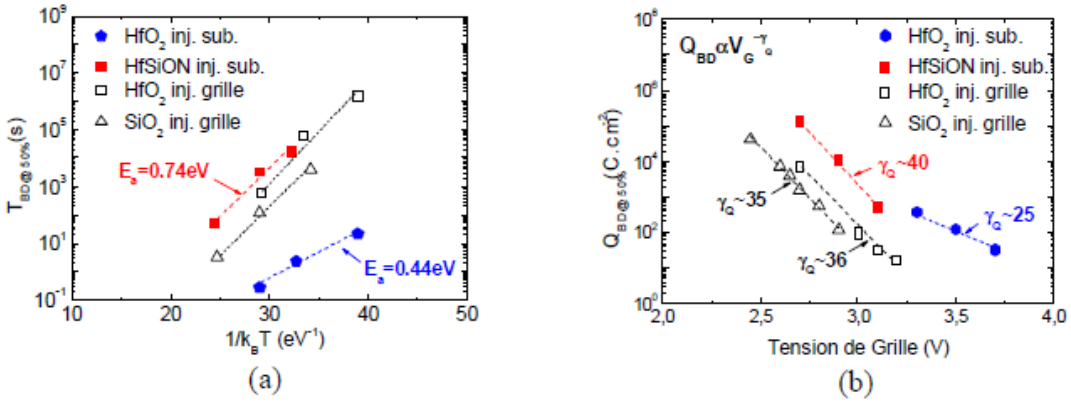


Figure 1.41: a) Activation energies extracted on different samples of SiO_2 / high-k stacks [115]

For the case of HfO_2 as high-k, many studies have been conducted in order to understand the breakdown mechanism [118][119]. From SILC measurements, it has been shown that the traps induced by stress in the high-k layer or at the interfaces enhances the trap assisting tunneling (TAT) and gives rise to a conduction path. In SILC measurements the initial Id-Vg curves is measured and the device is put under stress, during which many Id-Vg curves are measured at regular intervals of time. SILC is the normalized change in gate current Ig (t) from the initial current Ig (0) during stress time (t_s), and is defined as follows:

$$\text{SILC } (\Delta Ig/Ig) = (Ig\ (t_s) - Ig(0))/Ig(0) \qquad 1.23$$

SILC current varies depending on the measured voltage, and a peak appears in the SILC spectrum when the Fermi level of the substrate is aligned with the energy level of the defects in the oxide (Figure 1.42).

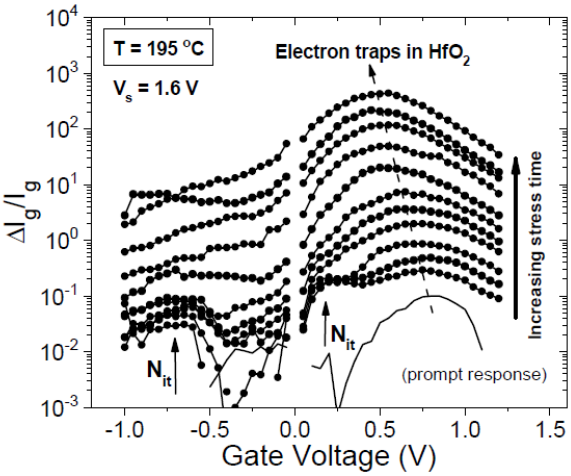


Figure 1.42: SILC spectrum showing the increase in the leakage current during stress [118]

Energy band diagram of the device can be formed at a certain gate voltage, and by knowing the peak position in the SILC spectrum, the energy level of the traps can be obtained. Based on the evolution of SILC spectrum during stress and its peak positions, the author showed that the peak position in the SILC spectrum corresponds to the energy level of oxygen vacancies in the bulk of HfO_2 [119] (Figure 1.43). This result is confirmed by the activation energy of SILC, which is attributed to the formation of oxygen vacancies in the HfO_2 oxide [118]. The authors report that these bulk oxygen vacancies in the high-k increase much more rapidly than interface traps and causes the breakdown of the high-k, followed by the interfacial layer degradation [108].

It was also proposed that the trapping in defects responsible for PBTI might be related to the defects responsible for SILC and breakdown of the gate oxide [33][107][108]. Cartier [107] found that SILC varies as the third power of PBTI shift and so the same defects might be responsible for both degradations.

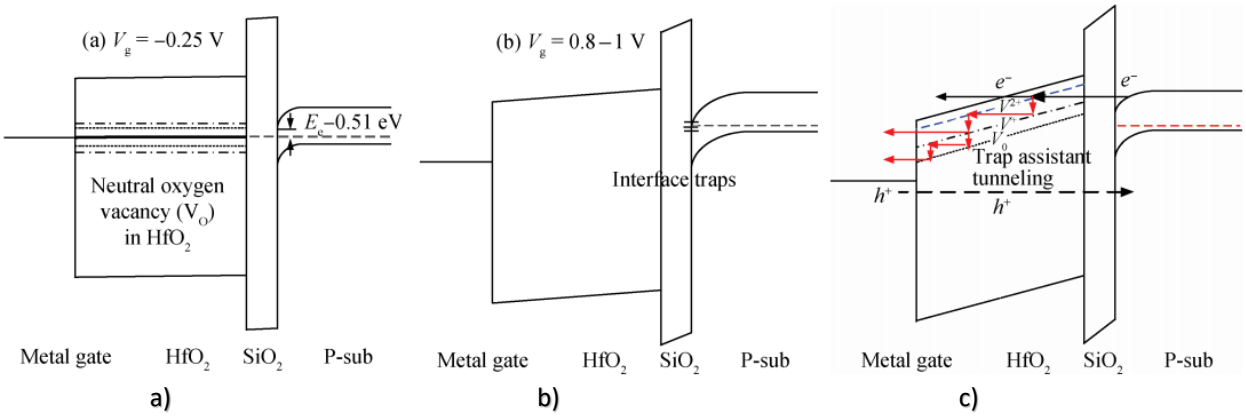


Figure 1.43: a) Band diagram of the SiO_2 / HfO_2 oxide NMOS transistor illustrating the energy level a) of the oxygen vacancies (b) of the interface states (c) of the oxygen defects assisting the conduction [119]

1.6 Conclusion

This chapter had focused on building an understanding of the working of the device used in this thesis. We mean: its fabrication flow and various electrical and physical phenomena that will be will investigated in the later chapters. In section 1.1 the working of the MOSFET device and its most important performance parameters such as the effective work function (WF_{eff}) or the flatband voltage (V_{FB}), threshold voltage and the impact of charges and dipoles in the gate stack on these parameters were introduced.

In section 1.2, processes for various layers used to fabricate the MOSFET device and which can influence its performance parameters were presented. Particularly in section 1.2.2, TiN metal electrode deposition by RF-PVD was introduced. In section 1.2.4, the introduction of additives by annealing for WF_{eff} engineering was presented, the impact of the annealing parameters on WF_{eff} will be investigated in chapter 5. In section 1.3, the process flow and its simplification, used for the different devices that we will analyze in later chapters, has been discussed.

In section 1.4, it is shown that a part of MOSFET V_T variability is related to the microstructure of the metal gate layer. Impact of TiN metal process parameters, introduced in section 1.2.4, on its microstructure and device electrical properties will be investigated in chapter 4.

In section 1.5, BTI and TDDB reliability phenomena occurring in the MOSFET devices was presented. Reliability of MOSFETS is very important as it can affect many of its performance parameters, thus the impact of additives that are used for WF_{eff} engineering, introduced in section 1.2.4, on reliability will be studied in chapter 3.

2. Characterization techniques for the gate stack

This chapter introduces the various electrical and physicochemical techniques used in this work to characterize the gate stack. Section 2.1.1 presents the capacitance voltage (CV) measurement technique, and an analytical and automatic method to extract flatband voltage (V_{FB}) and equivalent oxide thickness (EOT). A method to individually extract the charges and dipoles present in the gate stack will also be presented. Section 2.1.2 presents the four probe method used to extract sheet resistance of films.

Section 2.2 presents the various non-electrical techniques that have been used in this thesis. In section 2.2.1, the X-ray Photoelectron Spectroscopy (XPS) technique is introduced that will be used to obtain the band energies of the gate stack layers. Next in section 2.2.2, X-Ray Fluorescence (XRF) technique is presented that will be used to measure the additive dose and metal gate elemental composition. Further in section 2.2.3, the X-Ray Diffraction (XRD) technique is

introduced that will be used to obtain information on the microstructure of crystalline films, by combining the In-plane and out of plane configurations. Lastly, section 2.2.4 presents the technique to measure the stress added by the deposition of a film on the Si substrate.

2.1 Electrical characterization

2.1.1 CV measurements and electrical parameters extraction

MOS stack capacitance is the combination of two capacitances in series: the gate oxide capacitance (C_{OX}) and the semiconductor capacitance (C_{Si}), which is dependent on the gate voltage. Thus the total measured capacitance (C_{exp}) becomes:

$$\frac{1}{C_{exp}(V_G)} = \frac{1}{C_{OX}} + \frac{1}{C_{Si}} = \frac{EOT}{\varepsilon_{ox}} + \frac{1}{C_{Si}(Q_{Si})} \qquad 2.1$$

where C_{Si} depends upon the charge in the Si substrate (Q_{Si}) and Si surface potential (Ψ_{Si}) and so on the device operating regime (section 1.1.1). It is the variation of charge Q_{Si} with respect to Ψ_{Si}.

$$C_{Si} = -dQ_{Si}/d\Psi_{Si} \qquad 2.2$$

Ψ_{Si} is related to gate voltage V_G and flatband voltage V_{FB} as follows:

$$V_G = V_{FB} - (Q_{Si}/C_{OX}) + \Psi_{Si}$$

Depending on the operating regime, variation of C_{Si}, and so of the total measured capacitance C_{exp}, with gate bias V_G will be different as shown in Figure 2.1

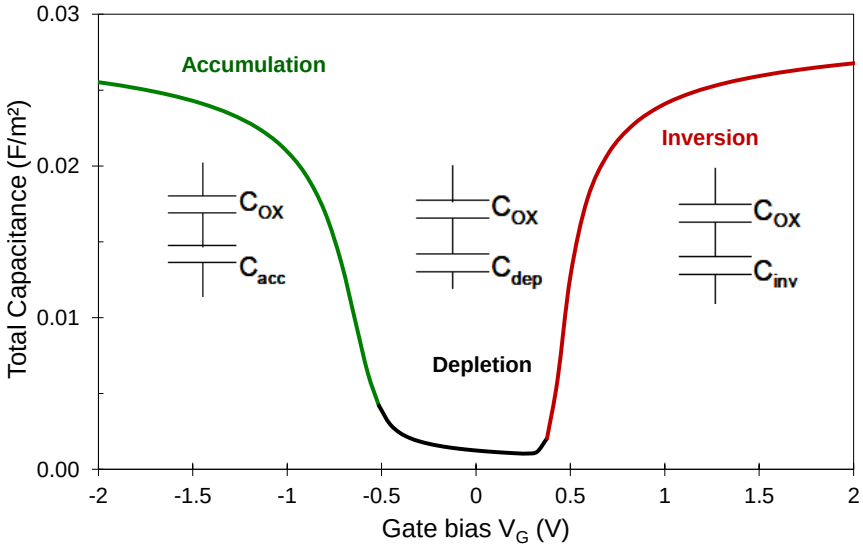


Figure 2.1: Variation of the total measured MOSFET capacitance

Experimental set up

To measure the capacitance, a MOS device is typically connected to a C(V) analyzer, such as the Agilent 4284A or 4980A, as shown in Figure 2.2. The C(V) analyzer applies a high frequency (between 100 Hz to 1 MHz) AC signal $V_g(t)$, which is superimposed on a main continuous voltage V_G, to the gate contact (High terminal). Induced current is measured through the substrate via another probe needle or the prober chuck, that is grounded (Low terminal), and the capacitance is calculated. In case of a MOS capacitor with connection only on gate and backside substrate, this setup authorizes the measurement of only the accumulation and depletion capacitances (inversion capacitance cannot be measured). This is due to the absence of the source and drain regions that normally provides sufficient amount of minority carriers, able to respond to the gate voltage variations.

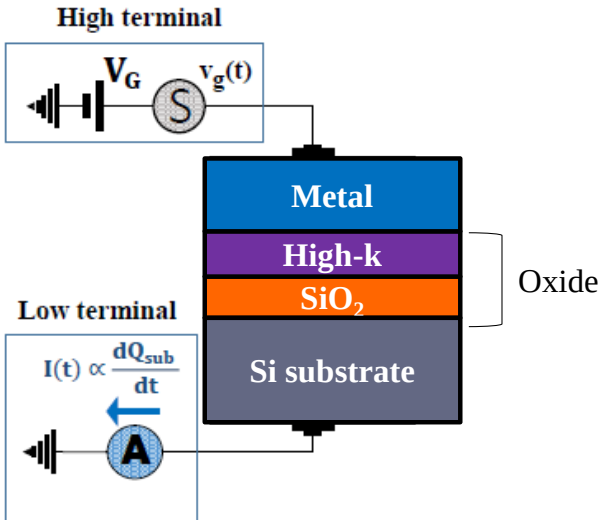


Figure 2.2: Experimental configuration for C-V measurements of MOS devices

Parameter extraction by Poisson-Schrodinger quantum simulations

The EOT and V_{FB} of a device can be obtained by the measurements of capacitance versus voltage (C-V) curves, but high leakage currents and interface states can lead to a wrong evaluation of these parameters [120]. Maserjian function [121] is able to deal with C-V curves that are deformed by high leakage currents but leads to inaccurate extraction in the case of stretched C-V curves [120], as it is based on capacitance derivative. Here we will describe, and further use for our studies, a methodology developed for a fast and reliable extraction of the EOT and V_{FB} from the experimental C-V characteristics, even if only a small part of the C-V characteristics is relevant due to leakage currents and interface states [122].

Dependence of C_{Si} with the charges in the semiconductor Q_{Si} and its relationship with the potential at the semiconductor surface Q_{Si} (Ψ_{Si}) can be described by the Poisson equation for electrostatics, according to classical electrodynamics. For ultra-thin oxide thickness and low field strengths, such interactions are better described by quantum Poisson-Schrodinger simulations. Indeed, for an accurate extraction of these electrical parameters on MOS devices with nanoscale oxides, one should take into account the effects of quantum confinement at the dielectric/substrate interface. These effects become significant when the thickness becomes comparable to the De Broglie wavelength of the carriers. The De Broglie wavelength for a free electron is equal to 1.2 nm at room temperature [123]. LETI has developed a one dimensional Poisson-Schrodinger solver which leads to the charge Q_{Si} and capacitance C_{Si} dependences versus the potential at the semiconductor surface Ψ_{Si} for various substrate doping levels (N_{sc}) [123]. The results are stored in a database, which are used for C-V analysis [124].

From the C-V measurements of a MOS device, two voltages V_G1 and V_G2 are chosen in between which the C-V characteristics are not affected by leakage currents or by interface states. For instance, V_G1 can be the maximum accumulation bias not-altered by high leakage currents and

V_G2 the minimum accumulation bias not influenced by interface states, as shown in Figure 2.3. At these bias conditions, Eq. 2.1 can be rewritten as follows:

$$\frac{\varepsilon_{ox}}{C_{exp}(V_G1)} - \frac{\varepsilon_{ox}}{C_{Si}(Q_{Si}1)} = EOT = \frac{\varepsilon_{ox}}{C_{exp}(V_G2)} - \frac{\varepsilon_{ox}}{C_{Si}(Q_{Si}2)} \quad 2.4$$

$$\frac{\varepsilon_{ox}}{C_{exp}(V_G1)} - \frac{\varepsilon_{ox}}{C_{exp}(V_G2)} = \frac{\varepsilon_{ox}}{C_{Si}(Q_{Si}1)} - \frac{\varepsilon_{ox}}{C_{Si}(Q_{Si}1 - \Delta Q_{Si})} \quad 2.5$$

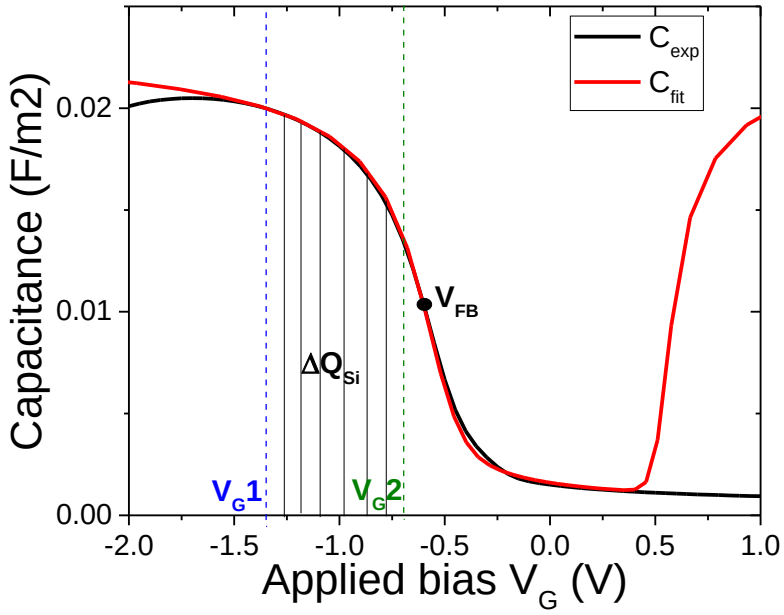


Figure 2.3: Variation of measured MOS capacitance (black) and its fitting by Poisson-Schrodinger simulations (red), to calculate the flatband voltage and EOT

The challenge is to find the substrate charge densities $Q_{Si}1$ and $Q_{Si}2$ that correspond to gate bias V_G1 and V_G2 respectively. The difference between $Q_{Si}1$ and $Q_{Si}2$ (ΔQ_{Si}) can be obtained by integrating C_{exp} from V_G1 to V_G2, as done in Eq. 2.6 below:

$$\Delta Q_{Si} = \int_{VG1}^{VG2} C_{exp}(V_G)\, dV_G \quad 2.6$$

By using the calculated relationships $Q_{Si}(\Psi_{Si})$ and $C_{Si}(\Psi_{Si})$ from the Poisson-Schrodinger database, $Q_{Si}1$ can be determined from Eq. 2.5. Indeed, $Q_{Si}1$ takes the value required to equalize

both sides of the equation. EOT is then obtained from Eq. 2.4. The flat band condition (V_{FB}) is finally calculated from Eq. 2.7 given below:

$$V_G1 = V_{FB} - Q_{Si}1 \cdot EOT/\varepsilon_{ox} + \Psi_{Si}1(Q_{Si}1) \quad 2.7$$

and its value corresponds to:

$$V_{FB} = V_G1 + Q_{Si}1 \cdot EOT/\varepsilon_{ox} - \Psi_{Si}1\,(Q_{Si}1) \quad 2.8$$

where the Ψ_{Si} value associated to V_G1 is taken from the Poisson-Schrodinger database.

Finally, the effective work function corresponds to:

$$WF_{eff} = qV_{FB} + q\Phi_{Si} \quad 2.9$$

Devices with beveled oxide

In section 1.3.3, the concept and process for beveled interlayer was introduced. The IL thickness varies along the wafers radius and so does the capacitance, as shown in Figure 2.4. From each of these C-V characteristics, we can extract a flatband voltage or effective work function WF_{eff} (Eq. 2.8 and Eq. 2.9) and an EOT (Eq. 2.4), by fitting the C-V curves with Poisson-Schrodinger quantum simulations.

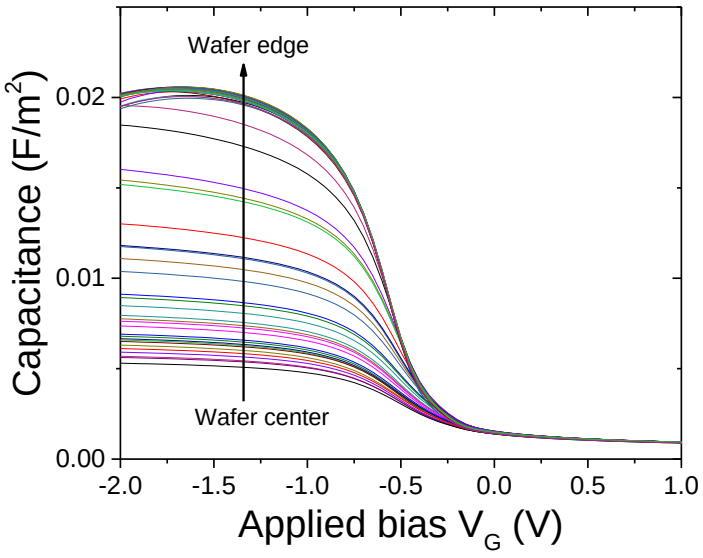


Figure 2.4: C-V characteristics of devices with varying EOT along the wafer radius

The WF_{eff} vs EOT curves can be plotted for each wafer, as shown in Figure 2.5. Linear behavior proves the absence of bulk charges in the beveled oxide. From Eq. 1.13 for WF_{eff} and Figure 2.5, fixed charges $Q_{Si/SiO2}$ at the Si/SiO_2 interface can be calculated from the slope of the WF_{eff} vs EOT curve. Moreover, The extrapolation of V_{FB} to EOT=0 allows the assessment of only metal

workfunction ($q\Phi_M$) and voltage drop induced by interfacial dipoles (δ), without the influence of SiO_2/Si interface fixed charges [5]. Metal workfunction can change due to change in its microstructure or thickness, this will be investigated in chapter 4. Interface dipoles δ can be modulated by additives or other process changes, particularly in the case of evaluation of additives incorporation in a sacrificial gate-first approach. In this case, assuming that the final TiN microstructure is unaltered (and so $q\Phi_M$), the WF_{eff} shift will be explained only by fixed charges or dipoles.

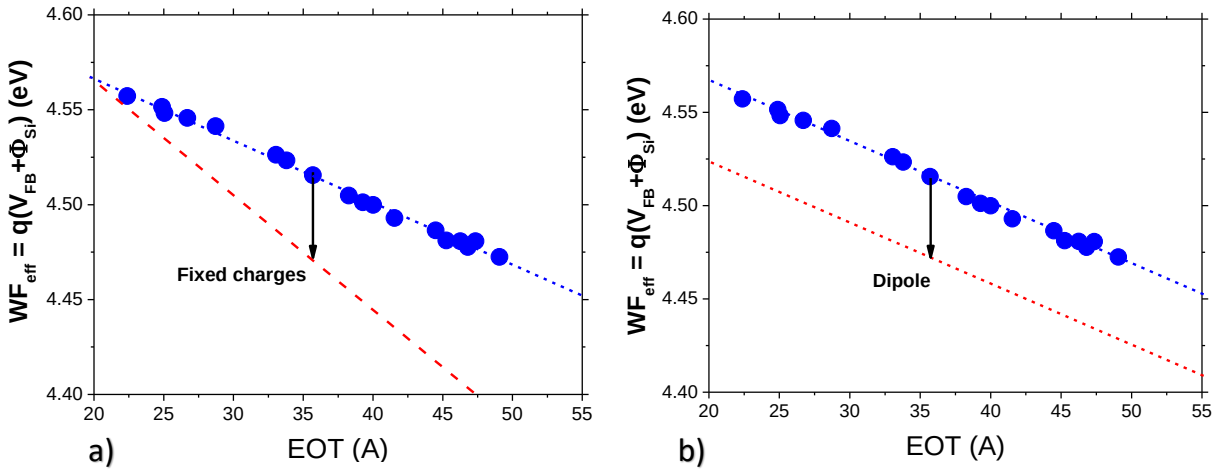


Figure 2.5: WF_{eff} vs EOT shift explained by a) fixed charges; b) a dipole effect or metal work function change

2.1.2 Sheet resistance measurements by four probe method

Four point probe is a method used to measure sheet resistance, bulk resistivity or thickness of films. The main principle is the use of separate pairs of probes for current and voltage sensing, by passing a current (I) through two outer probes and measuring the voltage (V) through the inner probes. The separation of the current and voltage probes eliminates the wire and contact resistances from the measurements. A current is passed through the two outer probes and that induces a voltage drop on the two inner probes, which is measured and used to calculate sheet resistance.

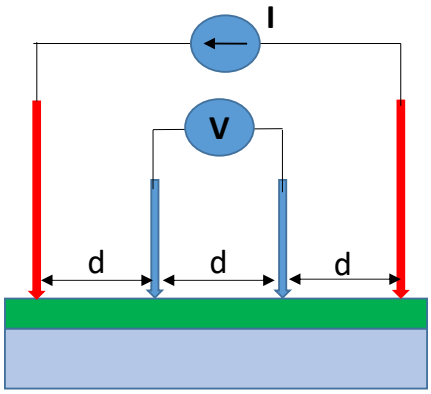


Figure 2.6: Four point probe configuration for sheet resistance measurements

The sheet resistance Rs of a film of thickness t, width W, length L, bulk resistivity ρ and total resistance R is:

$$Rs = R \cdot (W/L) = \rho/t \qquad 2.10$$

If we consider that the film dimensions are much larger than the size of the probe tips, then the sheet resistance can be calculated directly be equation below[125]:

$$Rs = \frac{\pi}{\ln(2)} \cdot \frac{V}{I} \qquad 2.11$$

Bulk resistivity is obtained by multiplying the film thickness by the sheet resistance (Eq. 2.10). Thus, if the thickness is known then resistivity can be calculated from these measurements. In the other case, if the resistivity is known then thickness can be easily calculated. In this work, the four probe technique has been used to calculate the sheet resistance of metal gate TiN films, processed by varying the RF-PVD parameters.

2.2 Physicochemical and Stress characterization

In this work, we have utilized several physiochemical techniques and stress measurements in order to characterize the HKMG stacks. From these characterizations, we are able to extract the physical, chemical and mechanical properties of the gate stack such as band energy levels, diffused additive dose, film composition, grain microstructure and mechanical stress.

2.2.1 X-ray Photoelectron Spectroscopy (XPS)

XPS is a non- destructive surface analysis technique that involves irradiating a material with a beam of X-rays while simultaneously measuring the kinetic energy and number of electrons that escape, by the photoelectric effect, from the material being analyzed. The interest of this technique is mainly in the analysis of the binding energy of core level (CL) electrons, which makes it possible to obtain important information on the electronic structure and the chemical environment of the atom in question. The analysis depth of the order of 10 nm, makes XPS particularly well suited for the study of physicochemical properties of the gate stack used in our studies, which is only a few nanometers thick.

Principle

XPS is based on the principle of the photoelectric effect. The sample surface is irradiated by a monochromatic X-ray of energy hν and the photons are absorbed by the atoms of the material to be analyzed. During irradiation, Incident photons can transfer some of their energy to the core level electrons of the material and they can be ejected as shown in Figure 2.7. The emitted electrons are collected and their energy spectrum is analyzed.

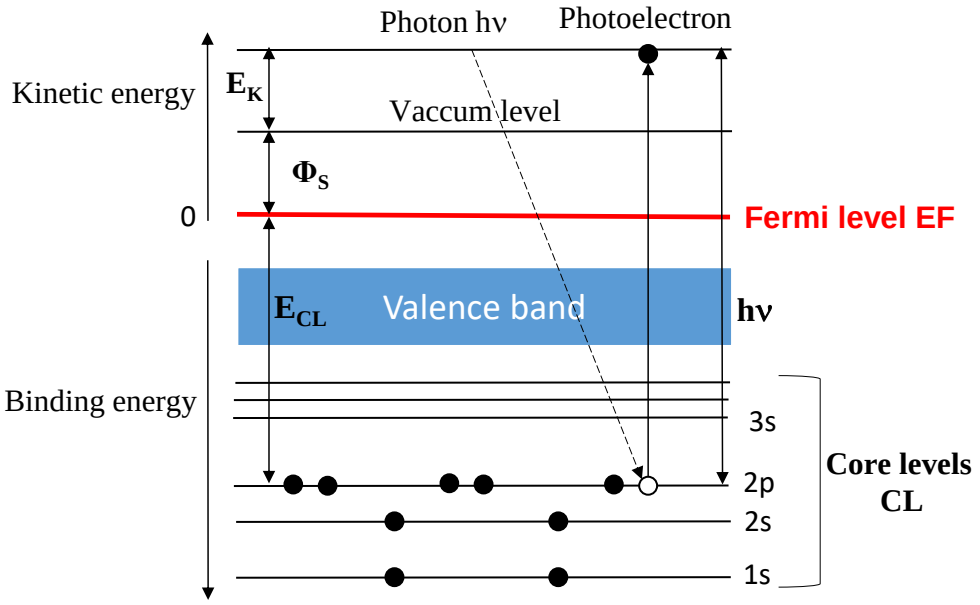


Figure 2.7: Energy diagram of the material undergoing photoelectric effect by irradiated beam

Vacuum level corresponds to an energy level of an ejected electron at rest, just outside the sample surface. The workfunction Φ_S corresponds to the difference between the material Fermi level EF and its vacuum level, E_{CL} is the binding energy of the core level electrons with respect to the EF and E_K is the kinetic energy of the emitted electrons. It is important to notice that in XPS analysis, energy of core levels are always reported from their fermi level. This is the contrary on the band diagram of MOS devices (Figure 1.2), where the energy of electrons and bands are reported from the vacuum level. The measurement of E_K must allow the determination of E_{CL}, from the law of conservation of energy expressed by Einstein's relation:

$$h\nu = (E_{CL} + \Phi_S) + E_K \qquad 2.12$$

Binding energy depends on the atoms and also on its chemical environment. The photon energy hν must be adapted to the type of electronic states studied. For this purposes, X-ray sources that most often consist of magnesium anodes (Mg Kα = 1253.6 eV) and aluminum (Al Kα = 1486.6 eV) are used. The process of photoemission in a solid can be described in three steps:

- Optical excitation of an electron from a core level to a free level (outside the atom) within the crystal after absorption of a photon of sufficient energy (E_{CL}). The probability for this transition is proportional to the ionization cross section, which depends on the energy of the incident photon and the atom.

- Propagation of the ejected electron to the surface. The photoelectron can undergo various interactions before reaching the surface. The electrons that does not undergo

any inelastic scattering gives rise to the main peak in the XPS spectrum. The electrons that undergo inelastic scattering processes during their path, loses a part of their energy and thus lose any information about their original core energy level. These scattered electrons contribute to the continuous background in the energy spectrum.

- Emission of the electron from the surface. The electrons have to cross a potential barrier (Φ_S) while crossing the sample surface, thereby reducing its kinetic energy.

XPS Sampling Depth

The transport of electrons to the sample surface, after ejection from an atom, is affected by the inelastic mean free path of the electrons (λ). λ is the average distance a photoelectron travels in the solid between two inelastic collisions and is related to the probability that an electron will reach the vacuum without scattering. The value of λ depends on the nature and volumetric mass density (mass per unit volume) of the material in which the electrons travels, λ is expected to be lower for dense materials. λ also depends on the energy of the ejected electron, as shown in Figure 2.8. If a beam is incident on a sample at an angle θ and by assuming that I_0 is the total XPS signal obtained for an infinitely thick layer, then the intensity contribution (I_s) of a surface layer of thickness d, to the total intensity of an XPS signal, is given by the relation:

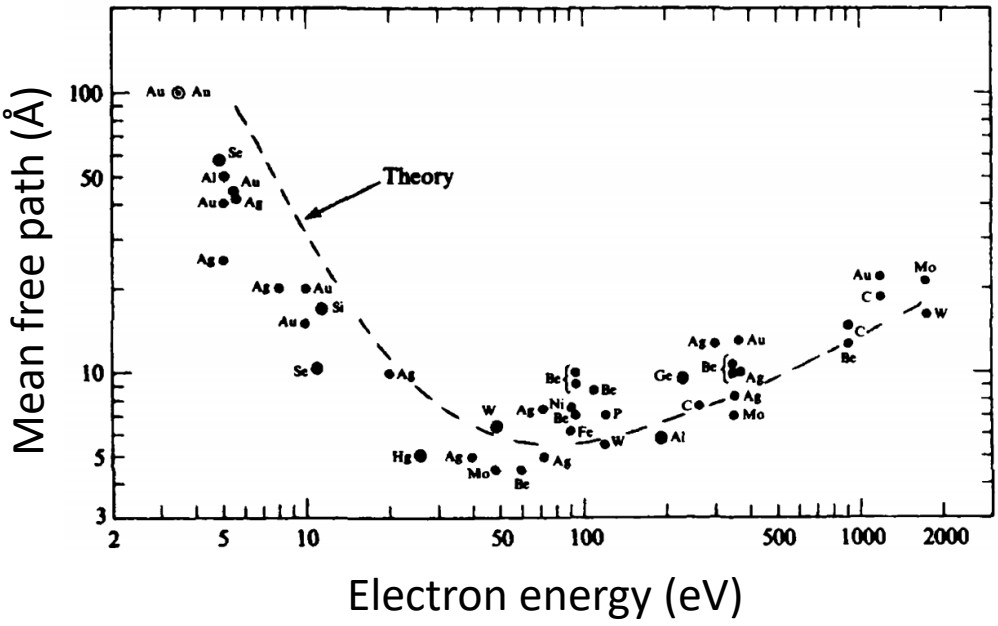


Figure 2.8: Electron mean free path versus its energy [126]

$$I_s = I_0 \cdot e^{-d/\lambda \sin(\theta)} \qquad 2.13$$

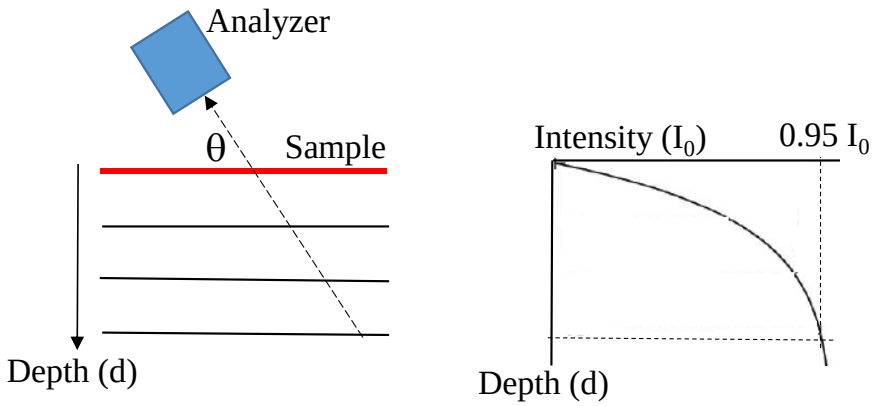


Figure 2.9: Contribution of a part of the sample thickness to the total XPS intensity

Sampling depth is defined as the depth from which 95% of all ejected photoelectrons are scattered by the time they reach the surface, which is equal to $3\lambda \sin(\theta)$ (from Eq. 2.13). Most λ's are in the range of 1-3.5 nm for Al Kα radiation, so the maximum sampling depth (3λ) for XPS under these conditions is 3-10 nm.

Measurements of the core levels (CL) spectra

The measurement of the kinetic energy of the electrons with respect to the vacuum level is more complicated because it requires to know Φ_S for each sample (Eq. 2.12). When the sample is electrically connected to a spectrometer, as shown in Figure 2.10, the Fermi levels of the spectrometer and sample are equal. Thus the measured kinetic energy $E_{K,meas}$ will be:

$$E_{K,meas} = E_K - (\Phi_{SP} - \Phi_S) \qquad 2.14$$

where Φ_{SP} is the workfunction of the spectrometer.

Inserting E_K from Eq. 2.14 into Eq. 2.12, we get:

$$h\nu = E_{CL} + \Phi_{SP} + E_{K,meas} \qquad 2.15$$

Thus, by knowing the workfunction of the spectrometer, the energy of the photons and the measured kinetic energy of the photoelectrons, the binding energy of the electron can be determined without the knowledge of the workfunction of the samples.

In order to calibrate the spectrometer, a measurement is done on a Copper (Cu) film and the offset of the spectrometer is adjusted to get the reference value for Cu-2p 3/2 at 932.62 eV [127]. Cu has a FCC crystal type with orientations <110>, <112>, <100> and <111> with respective workfunction of 4.48, 4.53, 4.59 and 4.94 eV [128]. The workfunction of a polycrystalline Cu is usually taken at 4.7eV [129]. Therefore it means that when a E_{CL} measurement is reported from a

calibrated XPS tool, the CL electron energy is referred to the Fermi level and assuming that its Φ_S is equal to the one of Cu.

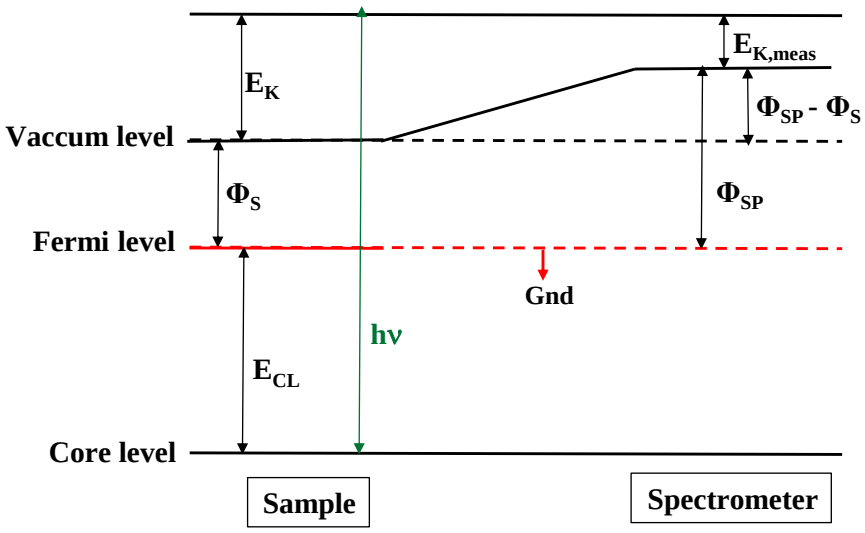


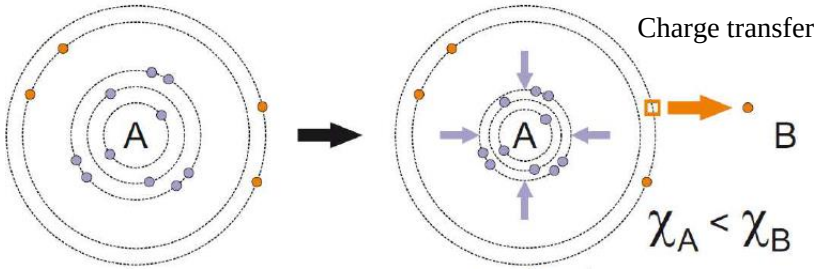
Figure 2.10: Schematic representation of energy levels during an XPS acquisition. The sample and the spectrometer are electrically connected so that their Fermi levels are equalized.

Analysis of atomic chemical environment

The binding energy of an electron in the core level depends on the nature of the element and the electronic level considered. In reality, the binding energy of an electron in a core level may differ from its theoretical energy, and this difference is mainly observed due to effects of the chemical environment or bonding of the considered atom. To illustrate this point, consider an atom A surrounded by identical atoms of A. The electronic arrangement of this atom is shown schematically in

Figure 2.11. When atom A is bonded to an atom B, which is more electronegative than A ($\chi_A < \chi_B$), higher electronegativity of B tends to attract the electrons involved in the covalent bond towards itself. Consequently, contraction of the core electronic levels takes place in atom A, which increases its CL binding energy (

Figure 2.11).



Contraction of core level energy

Figure 2.11: Schematic representation of the effect of atomic bonding on the shift of the binding energy of core levels

Spin orbit coupling

The orbital motion of the electron in a CL creates an internal magnetic field proportional to its kinetic moment L, which can interact with its intrinsic magnetization (the magnetic moment S), associated with the spin of the electron. This interaction is called spin orbit coupling. In a multi-electronic system, it increases the degeneracy of energy levels by taking into account the total kinetic moment J, such that J = L + S. The quantum number J has values between | L - S | and L + S. Thus, in the case of silicon, the electrons of an orbital p can have a J value of 1/2 or 3/2. The relative intensity of the components of the doublet, produced by the spin-orbit coupling, depends on the degeneracy of the state J, and is given by the expression (2J + 1). For the core level of 2p, the ratio of the intensities of the 3/2 to 1/2 peaks is 2. The difference between the components of a doublet depends on the strength of this coupling, and is equal to 0.6 eV for Si 2p.

XPS measurements and instrumentation

The XPS equipment is schematized in Figure 2.12. Its working principle is as follows [130]:

- It uses an X-ray tube as the source of radiation containing a LaB_6 filament. The heating of the filament causes the emission of electrons by thermoionization. The electrons from the filament are accelerated by a high voltage and strikes the Al anode that creates a point source of X-rays.
- The ellipsoidal shaped quartz crystal monochromator refocuses the X-ray beam onto the sample.
- The electron beam is scanned over the Al anode that causes the X-ray beam to scan on the sample surface.
- The electronically scanned X-ray beam enables the fast collection of X-ray beam induced secondary electron images (Field of view up to 1400 μm) that can be used for sample navigation and analysis area definition.
- The binding energy of the electrons is analyzed by a hemispherical detector. It consists of two concentric hemispheres of different radii, covered with a conductive coating. The application of a potential difference between the hemispheres gives rise to a radial electric field which influences the trajectory of the photoelectrons. The geometry of the analyzer, as well as the potential difference, define the analysis energy Ea (also called energy of passage) of the analyzer. This is the energy that an electron entering the analyzer must have in order to come out along a path of median radius. Electrons whose kinetic energy is less than Ea strike the inner sphere. Likewise, those whose energy is greater than Ea are stopped by the outer sphere. The resolution depends on the passing energy Ea, the mean radius, the acceptance angle of the photoelectrons and the width of the input and output slits. All photoelectrons emitted from the surface of the sample do not have a kinetic energy equal to Ea of the analyzer. Therefore, a potential electrical delay is applied at the input of the analyzer, so that the kinetic energy of the photoelectrons becomes equal to Ea. The most frequently used analysis method is the one in which the pass energy Ea is kept constant and a varying delay potential is applied on the electrons, so that their energy becomes equal to Ea. This allows to work with a constant energy resolution over the entire range of kinetic energy studied.

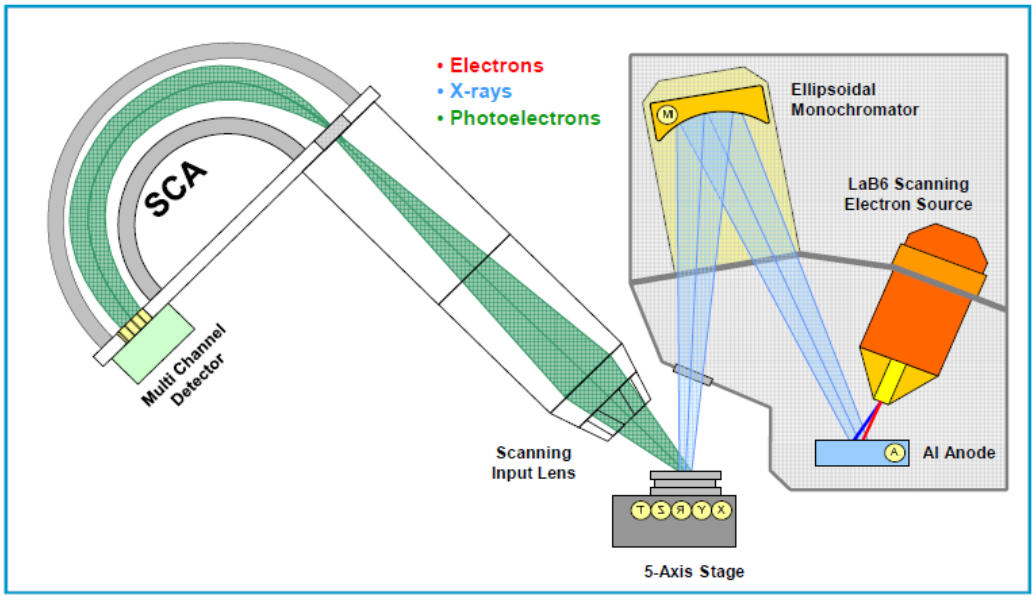


Figure 2.12: Schematic showing the working of the XPS equipment [130]

Effect of bias on the XPS spectra

The band energy levels of a material are shifted as a bias is applied to the sample. This applied bias acts as an additional potential barrier for electron ejection and so the kinetic energy of the photoelectrons is affected it. The effect of a bias applied to a p–n junction diode is shown in Figure 2.13 , where the diode was grounded from the n-side and the potential was applied from the p-side. The shift of the Si-2p XPS peak is clearly observed with the application of a bias [131].

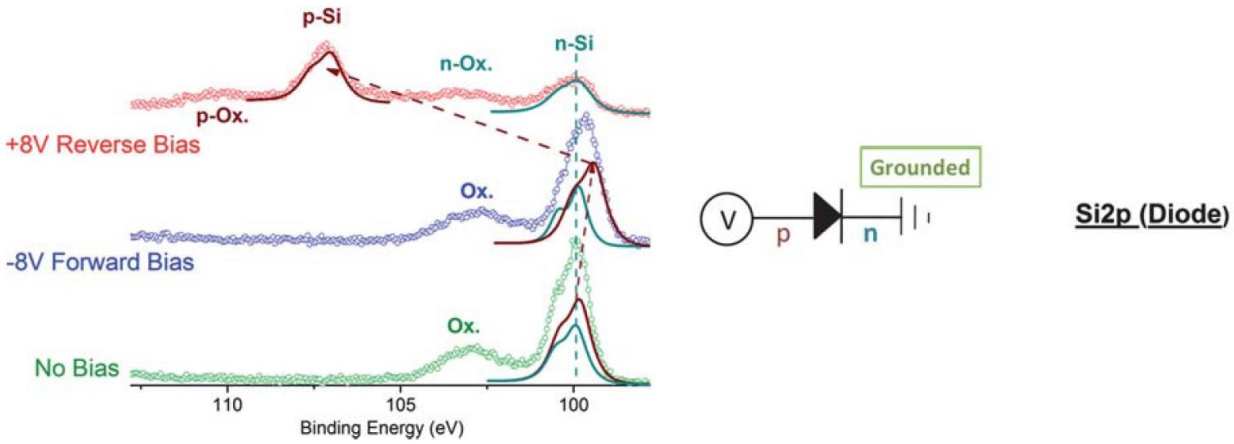


Figure 2.13: The Si-2p region of the p–n junction of a Si-diode recorded under: no bias, - 8 V forward and +8 V reverse bias. The inset schematically displays the electrical connection [131].

In our work, XPS has been used mainly to study the band energy levels of the HKMG gate stack in chapter-6. Also, it has been used to measure the effective dose of Al additives diffused into the gate stack in chapter-5.

2.2.2 X-Ray Fluorescence (XRF)

XRF is a non-destructive analytical technique used to determine the chemical and elemental composition of materials and its film thickness. It is based on the principle of analyzing the fluorescent (or secondary) X-ray radiation emitted from a sample when it is excited by a primary X-ray source (Figure 2.14). A X-ray beam with high enough energy to excite the inner core electrons of the atoms is created by a X-ray tube, which is then illuminated on the sample to be analyzed. This X-ray beam interacts with the core level electrons and supplies its energy to them, which results in electron excitation to higher states or ejection of these core electrons. The eletronic structure thus created is electronically unstable and in order to regain its stability, electrons from higher enegy levels fills this vacancy left in the inner level. This electron transition from a higher to a lower energy level gives rise to emission of a photon, called fluorescent or secondary X-ray beam.

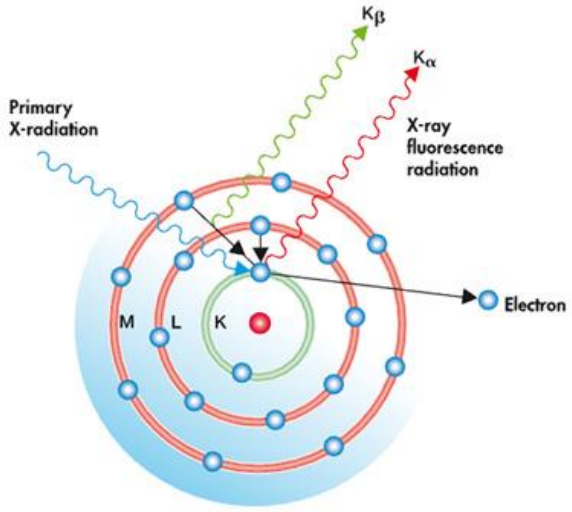


Figure 2.14: Principle of the X-Ray Fluorescence [132]

The energy of this secondary beam is equal to the specific difference in energy between two energy levels involved in the transition. The spacing between the energy levels or orbital of an atom is unique to the atoms of each element and hence the energy of the secondary beam is also unique, and this forms the basis of the XRF analysis. The main transitions taking place in XRF measurements are: K_α for L- K transition, K_β for M-K transition and L_α for M-L transition (Figure 2.14). XRF signal is directly related to the quantity of each element or atoms present in the sample. Therefore, it can be used as a direct measurement of film thickness, diffused additive dose or composition analysis of compound films.

The main limitation of this technique is that the fluorescence process is inefficient, and the secondary radiation is much weaker than the primary beam. This becomes a major challenge in measuring thin layers because of its long acquisition time needed for the XRF detector to acquire a statistically significant number of fluorescence counts. Moreover, the intensity of the secondary beam depends on the elemental mass of the atoms present in the sample. Secondary beam intensity from lighter elements such as Na, Mg, Si, etc. is relatively low and also have low penetration depth, and so will be difficult to detect even at relatively small depths. Whereas, heavier elements such as Cu, Ag, Au etc give rise to relatively higher intensity beam and also have larger penetration depth, and so are able to be detected from deep inside the sample. One limitation of heavier elements is that they will strongly absorb the emitted X-rays. So the best scenario for XRF measurements is of the heavy elements present in a matrix mainly composed of light elements. Another limitation of XRF technique is that it cannot be used for depth profiling (the order of stacking of multi layers), even while having a depth of analysis in the μm range. Furthermore, as the XRF signal efficiency depends on the element being analyzed and the collection efficiency of the detector, the XRF signal has to be calibrated by known thickness and composition standards.

In our studies, XRF technique has been used to measure the effective diffusion of additives into the gate stack after diffusion annealing. Moreover, it has been used to measure the composition of Ti versus N atoms in the TiN gate. XRF measurements were carried out on a RIGAKU WaferX300 equipment, with X-ray tube containing a rhodium target, operated at 40 kV, 90 mA and a spot size of 40 mm. XRF measurements were performed on many points on blanket wafers with nominal dielectrics thicknesses. For effective additive dose, the measurements were carried out at 2 steps during the process: a) after sacrificial gate stack deposition and b) after diffusion annealing and sacrificial gate stack removal (Figure 1.18). This helps to evaluate the effective dose incorporated into the high-k/SiO_2 stacks as a function of the as-deposited dose. La doses were determined from its characteristic spectral corresponding to M-L transition (L_α). Al doses were determined from its characteristic spectral line corresponding to L-K transition (K_α)[40].

2.2.3 X-Ray Diffraction (XRD)

Principle

XRD enables the identification of the crystalline structure of solid films. The sample is irradiated by monochromatic X-ray radiation, impinging at a certain incidence angle with respect to the sample surface. The radiation is diffracted by the crystals in the sample into many specific directions. By measuring the angles and intensities of these diffracted waves, a three-dimensional picture of the density of electrons within the crystal can be obtained and so the positions of the atoms in the crystal, called the crystal structure or orientation. This allows to evaluate the orientation of crystals, crystallinity fraction and average size of crystal grains in a micro/poly-crystalline sample.

In a perfect crystal we have different crystal planes with different orientations with respect to the sample surface. Each set of crystal planes are defined by miller indices h, k, l that are separated by a spacing d_{hkl}. Each set of planes gives rise to constructive interference for certain values of the

incidence angle θ_{hkl} at which the Bragg condition for that plane is satisfied in Eq. 2.16, i.e. when the difference of path lengths between the scattered waves is an integer multiple of the wavelength of the incident radiation, as shown in Figure 2.15.

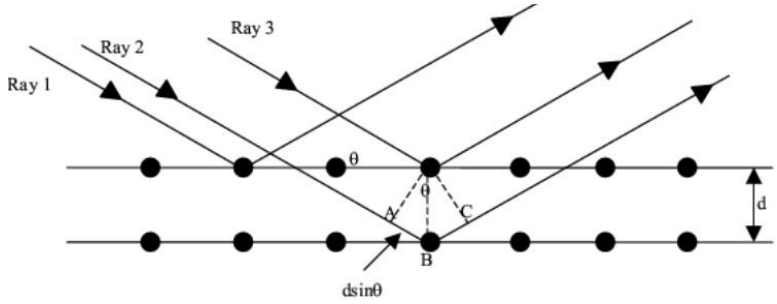


Figure 2.15: Diffraction of incident X-rays from parallel planes separated by a distance d [133]

$$2d_{hkl} \sin\theta_{hkl} = n\lambda \qquad 2.16$$

where λ is the wavelength of the incident radiation.

A typical XRD diffraction pattern shows the intensity of X-ray radiation detected as a function of the detection angle. The presence of a peak at a certain value of 2θ can be related to the Bragg diffraction by a particular crystal plane. Moreover, in the case of a poly-crystalline material the width of the diffraction peaks can be related to the average size of the crystals, i.e the bigger the crystals, the sharper are the peaks.

The intensity of the diffraction peaks (I_{hkl}) from a set of planes is determined by the crystal system and its crystal orientation (h,k,l miller indices). The peak intensity of diffracted beam from a set of crystal planes is affected by the scattered waves from all the atoms in the unit cell and thus by diffraction from other planes. The structural factor (F_{hkl}) determines the influence of atomic arrangement on the peak intensity from a set of planes (h,k,l). Thus, different peaks in the diffraction pattern might not have the same peak intensity and are proportional to F_{hkl}. Sometimes, destructive interferences can occur between the reflections of different planes and the intensity of that set of planes will not be observed in the XRD pattern. In case of a face-centered cubic (FCC) structure (which TiN or NaCl has), a peak is observed when the h,k,l indices are all odd or all even. Thus, the intensity of a peak in the diffraction pattern will be determined by the square of the structure factor F_{hkl} which will be different for various crystal systems and for their family of planes.

Peak Intensity in TiN crystal structure

Due to F_{hkl}, diffraction peak intensities corresponding to different crystal orientations in the sample will not have the same values, even if these crystals are present in same quantity in the sample. TiN, used as the metal gate, have a cubic NaCl-type FCC crystal structure [134]. As mentioned earlier, this crystal structure only have diffraction peaks when all h,k,l indices are either odd or even, otherwise the peak will not be visible in the XRD pattern for other Miller indices.

Thus, TiN (111), (200), (220), (311), (222), (400), (331), (420), (422), (511) planes are observed but others such as (100), (110), (210) are not observed in the diffraction pattern [40]. Information on F_{hkl} for various crystals, in their powder form, can be found in the International Centre for Diffraction Data (ICDD) database. This includes relative intensities of different crystal planes in a powdered sample, the d_{hkl} spacing and their angle of diffraction. The most intense diffraction peaks of TiN are summarized in Table 2.1, these correspond to measurements done by a monochromatic X-ray source with $\lambda = 1.540$Å.

h	k	l	I_{hkl}	$2\theta_{hkl}$	d_{hkl}
1	1	1	**72**	36.6	2.449
2	0	0	**100**	42.6	2.12
2	2	0	**45**	74.06	1.279

Table 2.1: Relative intensities (I_{hkl}) and peak positions (2θ) for the observed diffraction peaks in TiN [134]

XRD equipment and geometric configurations

A typical XRD equipment mainly consists of 1) A source of X-ray consisting of a X-ray emission tube with a copper anode; 2) A monochromator; 3) A X-ray detector. In most cases the sample and the detector can be rotated by a goniometer, in order to change the incidence and detection angles. In our studies, depending on the information to be obtained on the microstructure of the samples, two XRD configurations have been used: the θ-2θ and the in-plane configuration as shown in Figure 2.16. In both these configurations, the source and the detector are always on the circumference of a circle with the sample in the center. Ω is the angle between the incident X-ray beam and the sample and 2θ is the angle between the incident beam and the detector angle, as shown in Figure 2.16 a. Crystal information will be obtained along the direction of the diffraction vector, which is parallel to the bisector of incident and diffracted beam and perpendicular to the sample surface. This means that only the crystal planes that are oriented perpendicular to the diffraction vector will be observed in the XRD pattern [40].

Figure 2.16 a shows the θ-2θ XRD configuration. The angle of incidence Ω and the detector angle 2θ (at which X-rays are collected) are always related by $\Omega = \theta$. To achieve this, the X-ray source is kept fixed and the sample and detector are rotated by θ°/s and 2θ°/s respectively. In this configuration only diffraction from crystal planes that are parallel to the sample surface is detected, as the Bragg condition can be satisfied only for these planes. Thus, we can study the size of vertical grains and their relative amounts present in the sample.

Figure 2.16 b shows the in-plane XRD configuration. The angle of incidence Ω and the detector angle 2θ are kept fixed at a very low value (few degrees), nearly parallel to sample surface. $2\theta_\chi$ and φ angles are changed together by keeping the X-ray source fixed and rotating the sample and detector by φ°/s and 2φ°/s respectively, so that the diffraction vector is always in the same direction.

As a consequence, analysis will only be performed on the same plane of the sample which is perpendicular to the sample surface. Thus, we can study the size of the horizontal grains in the sample.

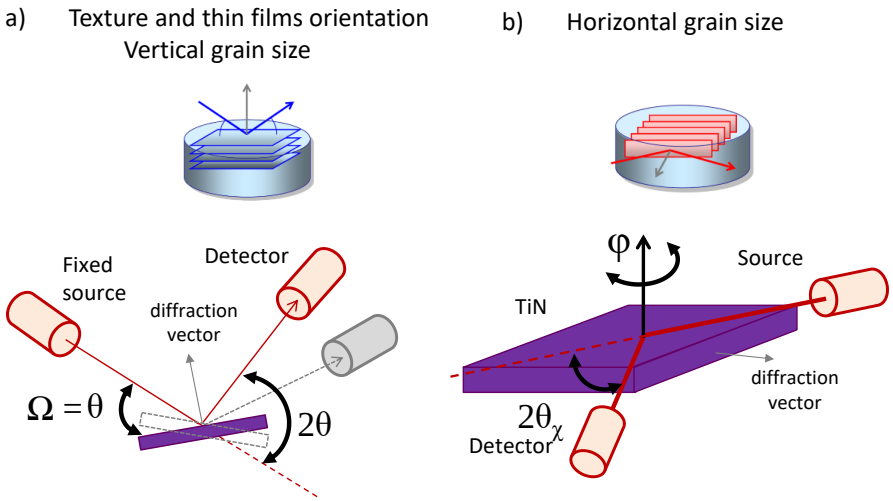


Figure 2.16: a) θ-2θ XRD configuration and b) in-plane XRD configuration [40].

Determination of crystallite size and their proportion

Figure 2.17 shows a typical XRD diffraction peak, with its parameters such as the peak position (2θ), peak height (I_{max}), peak area (A) and the full width at half maximum (FWHM). The FWHM is calculated by the difference between the two values of 2θ that corresponds to the intensities equal to half of I_{max}. In our studies, the diffraction spectra are fitted by assuming a Gaussian or Lorentzian shape.

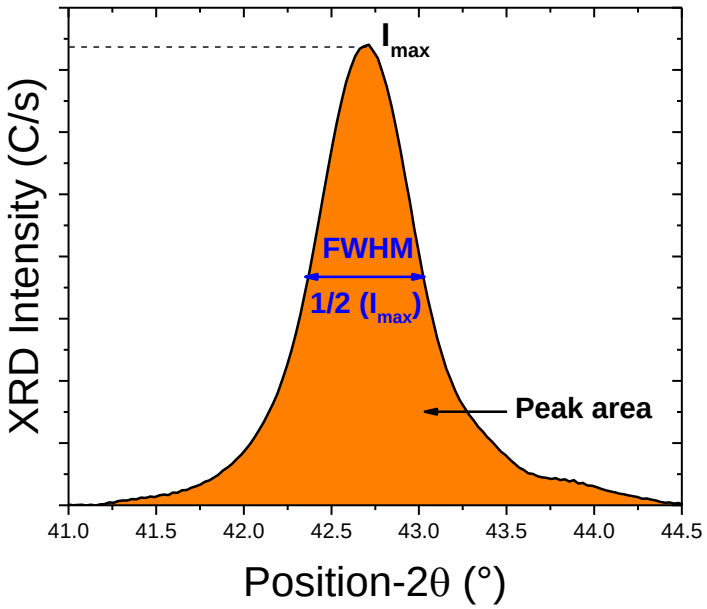


Figure 2.17: Peak in a X-Ray diffraction pattern showing its various parameters

The FWHM of the diffracted peak will be affected by the peak broadening due to instrumental imperfections and by structural defects in the crystal. In order to get rid of the instrumental contribution, a near-perfect sample or powder was first measured to determine the diffractometer's broadening, and then this value was removed from the peak width obtained for a sample.

The mean size of crystallites <D> was determined by using the Scherrer equation given below:

$$< D >= (0.9\lambda)/(\beta \cos\theta) \qquad 2.17$$

where λ is the incident wavelength, β is the FWHM, and θ is the Bragg (diffraction) angle.

The diffraction intensity should not be related to the peak height but to the peak area, since the area represents the true sum of all diffracted X-ray photons that have been detected regardless of the peak shape. Thus, the relative distribution of crystallographic orientations of the crystallites were calculated from the ratio of area under the diffraction peak, for different orientations. As discussed earlier, X-Ray diffraction intensity is not directly correlated to the amount of a specific oriented crystallites in a sample, but are influenced by the structure factor. Thus, the ratio of area under the peaks for different orientations, corrected by the relative intensities due to this structural factor has been taking into account to calculate the relative amount of different orientations.

As stated in section 1.4, each crystallite orientation is associated to a different value of work function and therefore contributes to the local V_T variability. In addition, it was also stated that a decrease of the horizontal crystallite size might be of benefit to reduce the local V_T variability. In consequence, the impact of TiN process deposition conditions on the preferred orientation of the TiN crystallites and their size will be investigated in chapter-4, with the help of the θ-2θ and the in-plane XRD configurations.

2.2.4 Mechanical stress measurements

Mechanical stress in thin films used in microelectronic devices varies significantly with its process conditions. There is a strong correlation between microstructure and stress. Mechanical stress can originate due to: 1) Extrinsic stress, caused by applied stress, thermal expansion and plastic deformation; 2) Intrinsic stress, caused by material growth, morphology, lattice mismatch and phase transformation. Mechanical stress can be of two types: Tensile and compressive, which can cause bending of the films from their normal shape.

In this work, mechanical stress σ_f of TiN films deposited over a Si substrate has been calculated by the Stoney's equation:

$$\sigma_f = \frac{E_s \cdot e_s^2}{6 \cdot (1-\nu_s) \cdot e_f \cdot R_f} \quad 2.18$$

$$1/R_f = 1/R_{fs} - 1/R_s \quad 2.19$$

where E_s is the Young's modulus and ν_s is the Poisson ratio of Si, e_s and e_f are the thicknesses of the substrate and TiN film, R_{fs} and R_s are the radius of curvature of the wafers before and after TiN deposition respectively.

In our work, this method will be used to calculate the stress added by the deposition of metal gate TiN films on Si substrates, processed by varying the RF-PVD parameters.

2.3 Conclusion

In conclusion, the various techniques to characterize electrical, physical and mechanical properties of the gate stack have been presented and reviewed. The methodology for a fast and quick extraction of the effective work function (WF_{eff}) and the equivalent oxide thickness (EOT) is introduced, by fitting the ideal (or relevant) part of the experimental capacitance voltage (C-V) curves by Poisson-Schrodinger quantum simulations. Moreover, the gate stack charges and interface dipoles can be individually obtained by including the interlayer bevel process into the gate stack, this method will be used to extract WF_{eff} and dipoles in chapter 4, 5 and 6. Sheet resistance technique based on the four probe measurements has been presented, it can be used to quickly calculate the sheet resistance and the bulk resistivity in chapter 4.

Spectroscopy techniques to probe the gate stack, that utilizes X-rays as their source, have been introduced. This includes the X-Ray Fluorescence (XRF) technique that will be used to measure the additive dose in chapter 5 and metal gate elemental composition in chapter 4, and X-Ray Diffraction (XRD) technique that will be used to obtain information on the microstructure of TiN crystalline film in chapter 4. X-ray Photoelectron Spectroscopy (XPS) technique will be used to analyze the band energies of the gate stack layers in chapter 6. Lastly a technique to measure the stress added by the deposition of a film on a Si substrate is presented, which will be used to measure stress induced by the TiN metal gate layer.

3. Impact of La and Al additives on MOSFET reliability

In this chapter, the impact of La and Al additives used for threshold voltage adjustment (section 1.2.4), by dipole formation at the SiO_2/HfON interface into the FDSOI gate dielectrics, on MOSFET reliability has been studied. As described in section 1.2.4, in FD-SOI devices, the effective work function (WF_{eff}) values of metal electrodes must satisfy the threshold voltage (V_T) requirements of specific devices, for which La and Al are introduced in the gate stack [121] [121] [52] [53]. Therefore, these additives are strong candidates for future FDSOI MOSFETs and therefore it becomes important to study their impact on device reliability.

First in section 3.1, the impact of these additives on NBTI and PBTI reliability (as discussed in section 1.5.1) will be studied [PK-1]. This covers first the device fabrication and secondly the test methodology, analysis for NBTI and PBTI V_T shift and the role of oxide field on the BTI reliability. Further, results on the role of La and Al additives on life-time and time to failure of the device are studied. Lastly, the physical interpretation related to the impact of these additives on the HKMG stack is discussed.

Next in section 3.2, the impact of these additives on the TDDB reliability (as discussed in section 1.5.2) will be studied. This covers the electrical measurements for TDDB, time to breakdown detection and their representation on the Weibull distribution scale. Further, the methodology to select the gate stress voltage for TDDB tests is described. Lastly, results on the role of La and Al on NMOS and PMOS device time to breakdown will be shown and their explanation by defect creation in the oxide is discussed.

3.1 Impact of La and Al additives on BTI reliability

3.1.1 Device fabrication

FDSOI N-MOSFET and P-MOSFET devices were fabricated by the process flow given in section 1.3.1 and in Figure 1.20. La and Al additives were introduced, for V_T tuning, into the HKMG stack by the sacrificial gate process shown in Figure 1.18. The gate dielectrics consists of a 1 nm thick SiON interlayer (IL) dielectric and a 2 nm thick HfON high-k layer processed by atomic layer deposition of HfO_2 followed by decoupled plasma nitridation. Then, the sacrificial gate stack consisting of La or Al layers sandwiched between a bottom pedestal TiN and a top capping TiN (45Å) layer were deposited over HfON (TiN/La/TiN-45Å or TiN/Al/TiN-45Å). TiN, La and Al layers were deposited in RF-PVD chambers, as explained in section 1.2.2. In order to modulate the diffused La or Al dose in the gate stack, thicknesses of the additive layers: La (0Å, 2Å, 4Å, 6Å) or Al (0Å, 2.2Å and 6.6Å) and bottom/pedestal TiN (0Å, 10Å, 20Å and 25Å) was varied. A thermal treatment under N_2 atmosphere at 900ºC for 10 s is performed in order to activate the diffusion of La and Al into the HfON/SiON stack. Thus, various La and Al doses used in our study were introduced into the gate stack. The sacrificial gate stack is then removed by wet etching. Finally, a poly-Si/TiN electrode is deposited on top of HfON, followed by gate patterning and S/D

annealing. Once the sacrificial gate is removed (step 2 in Figure 1.18), the effective doses of La or Al diffused into the gate stack can be evaluated by XRF diffusion measurements (described in section 2.2.2) [47], they are reported in Table 3.1. Effective dose of TiN 10Å/Al 6.6Å is lower than TiN 10Å/Al 2.2Å due to lower diffusion of Al through pedestal TiN at higher Al doses [135], which is caused by the formation of AlN clusters in the pedestal TiN at high Al doses. La-0 Å and Al-0 Å splits are termed as the reference (ref in Table 3.1) splits for all the others splits.

La Splits	La dose (at/cm²)	Al Splits	Al dose (at/cm²)
La 0Å (ref La split)	0	Al 0Å (ref Al split)	0
TiN10Å/La2Å	4.24E13	TiN25Å/Al2.2Å	1.06E14
TiN10Å/La4Å	9.13E13	TiN20Å/Al2.2Å	1.53E14
TiN10Å/La6Å	1.2E14	TiN10Å/Al6.6Å	2.32E14
La 2Å	2.36E14	TiN10Å/Al2.2Å	4.14E14
La 4Å	4.5E14		

Table 3.1: Different La and Al splits used for this study, including variations in TiN pedestal, and their corresponding effective doses

3.1.2 Electrical measurements and analysis

Threshold voltage measurements

The initial threshold voltage V_{T0} (before stress) of the MOSFET devices is calculated by the constant-current method. For this, the devices were subjected to a gate bias (V_G) sweep and a constant drain bias (V_D), and the drain current (I_D) is measured (I_D-V_G curves). V_{T0} is calculated as the gate voltage at which the drain current is equal to a constant current (I_{D0} in Eq. 3.1), appropriate for a given technology, times the ratio of gate width (W) to gate length (L) of the transistor. V_T is calculated using:

$$V_{T0} = V_G \text{ (at } I_D = I_{D0} \cdot W/L) \qquad 3.1$$

I_{D0} is selected for a given technology such that V_{T0} is in the subthreshold region of the device operation. For N-MOSFET devices, $I_{D0} = 0.2$ µA and for P-MOSFET devices $I_{D0} = -0.1$ µA. Figure 3.1 and Figure 3.2 shows the V_{T0} distribution for devices with different La and Al doses Table 3.1 respectively, and linear change in V_{T0} can be seen with La and Al dose due to dipole formation at the SiO_2/HfON interface. The split Al-4.14E14 in Figure 3.2a has a V_T that is less than expected for such Al dose, compared to other splits. This is due to a different Si substrate WELL doping

type (NWELL or PWELL in Figure 1.19) compared to other splits. This change in WELL type changes the V_T but should not affect the BTI behavior.

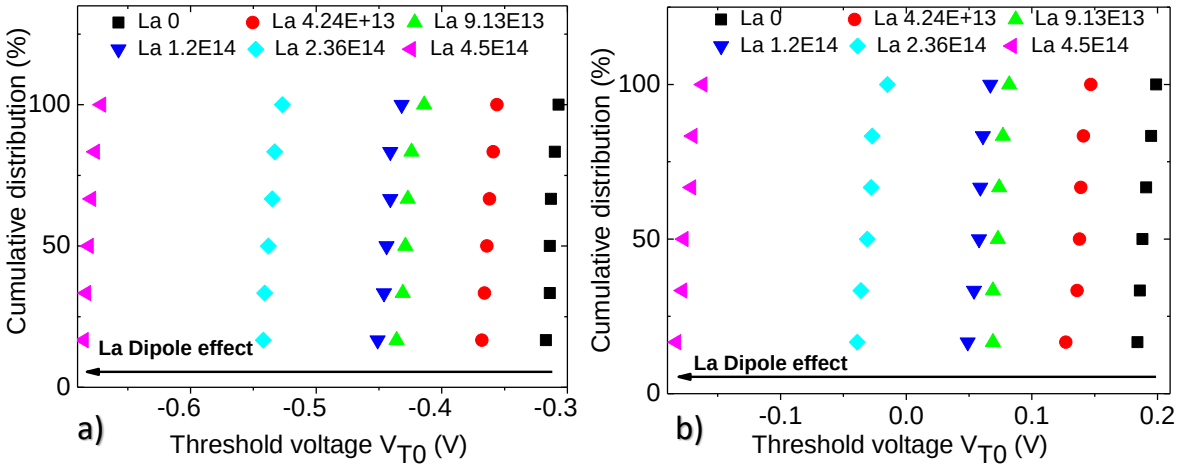


Figure 3.1: Threshold voltage distribution for devices with various splits of La doses used to investigate BTI reliability, for a) PMOS; b) NMOS

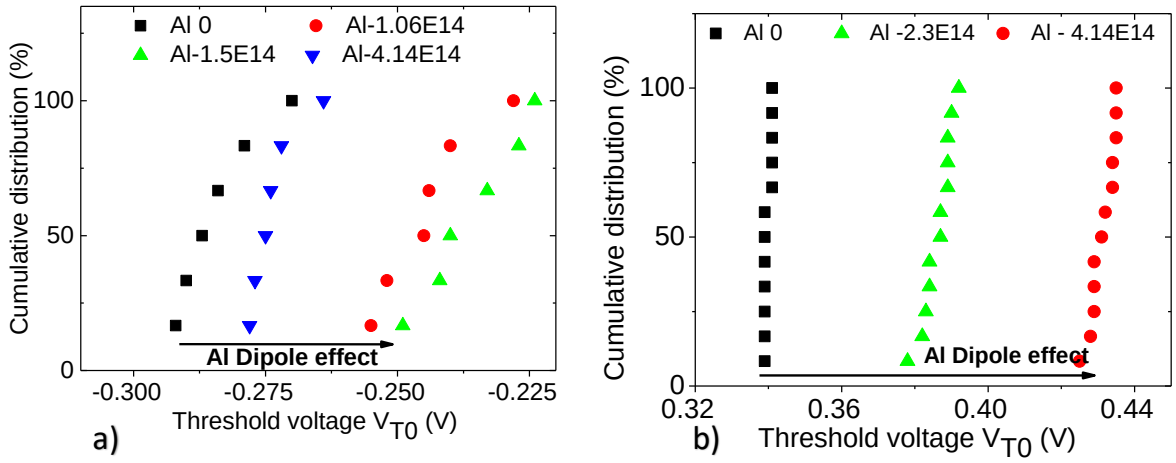


Figure 3.2: Threshold voltage distribution for devices with various splits of Al doses used to investigate BTI reliability, for a) PMOS; b) NMOS

BTI measurements

The NBTI and PBTI tests were performed by the conventional method [83]. The initial I_D-V_G characteristics are measured and then the devices are stressed by applying a high gate voltage V_G and a high temperature, and V_T shifts (ΔV_T) are measured at predefined points in time. During the ΔV_T measurement phase, the V_G is lowered to a value $V_{G,meas} \cong V_T$ and the drain current is measured. The V_T shift is then determined by the horizontal shift of the initial I_D-V_G curve so as to have this drain current at the $V_{G,meas}$, as shown in Figure 3.3.

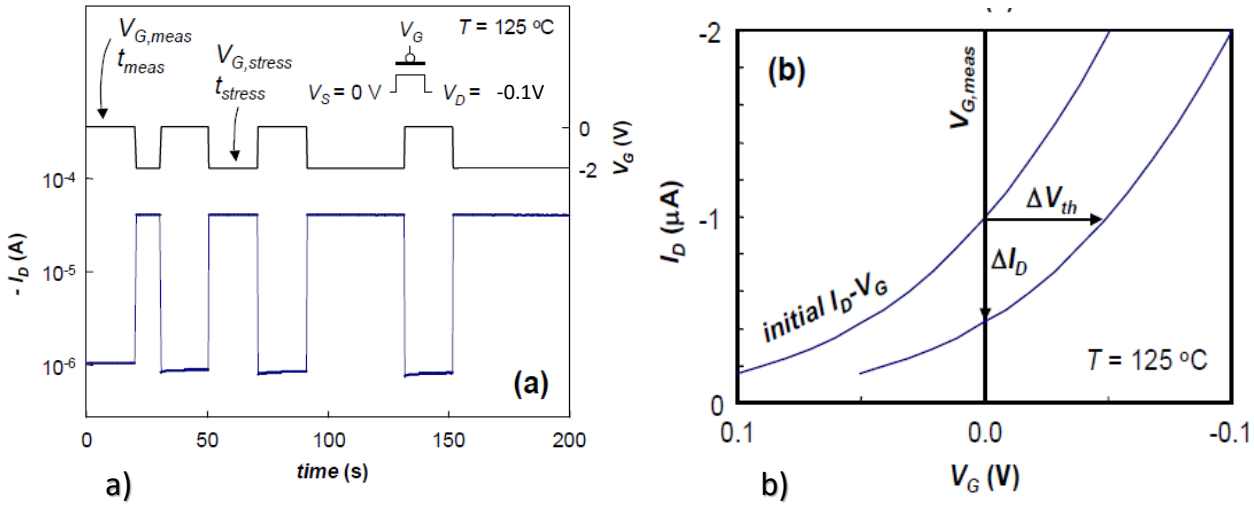


Figure 3.3: The experimental method used to study BTI V_T shift. a) A PFET device is biased and V_G is varied between stress and measurement values while I_D is monitored; b) Calculation of ΔV_T by horizontal shift of initial I_D-V_G curve [83]

In our studies, NBTI and PBTI tests were done for 1000 s at gate voltages of ± 1.2, 1.4, 1.6 and 1.8 V on PMOS and NMOS devices respectively, of length 20 nm and width 2 μm and stress temperature of 125°C. The drain voltage was kept at 100 mV.

Analysis method

In order to understand the BTI degradation results and to compare them on different devices, an analysis method (or model) and degradation parameters must be defined. For each device with different dose of La or Al and for each stress bias, BTI measurements were performed on three identical devices and the V_T shift (ΔV_T) with stress time was calculated from the method described above. ΔV_T was calculated as an average over these three devices. Figure 3.4 shows the NBTI ΔV_T for a PMOS device measured at regular intervals of time for various stress voltages, for the device with ref La dose (La-0 Å).

ΔV_T vs stress time follows a linear relation in a log-log scale (power law) as predicted by the mechanisms in section 1.5.1. By fitting the experimental data at different V_G with a power law having a same time (t) exponent b, the factor A is obtained for each V_G, $A(V_G)$:

$$\Delta V_T = A(V_G).t^b \qquad 3.2$$

The knowledge of $A(V_G)$ for different V_G then authorizes the identification of a second power law in γ: $A(V_G) = A_0 \cdot V_G^{\gamma}$. Thus the complete power law, in V_G and time, of Eq. 1.14 can be obtained. The same method has been applied for A and b parameters extraction for all the other devices. This power law can then be used to extrapolate the V_T shift for a specific stress voltage and time, specifically: V_T shift after 5 years (ΔV_{T5y}) and time to failure for 50 mV V_T shift (T_{50mv})

at the reference bias of 0.945V, which is our industry standard for the devices used in this work. ΔV_{T5y} and T_{50mv} are the two performance parameters that will be used in our studies to compare devices with different doses of La and Al.

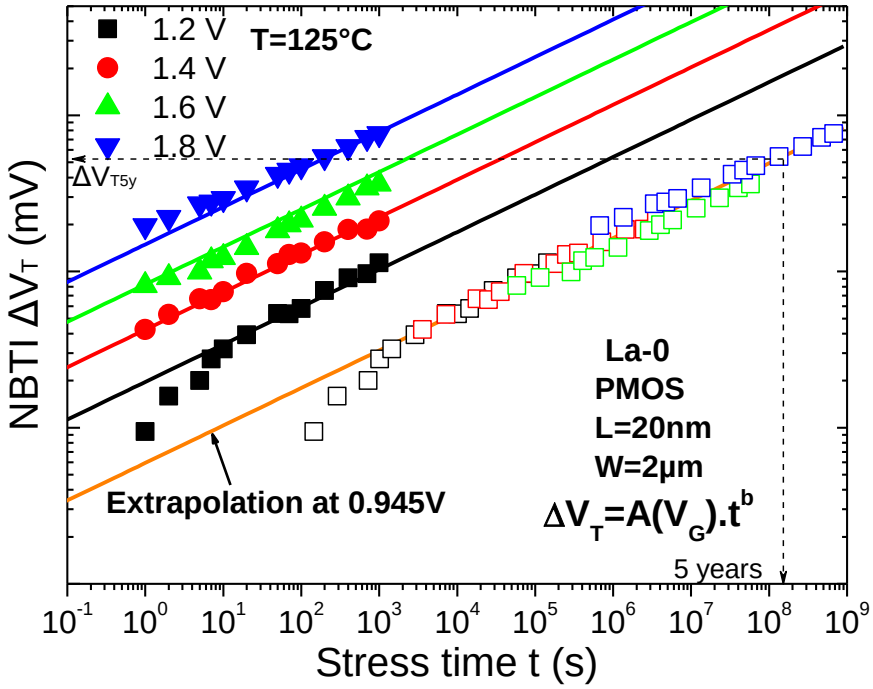


Figure 3.4: NBTI V_T shift vs. stress time for device without La at different stress voltages. Power law extrapolation is done to calculate V_T shift @5years (ΔV_{T5y}) and time to failure@50mv (T_{50mv})

Oxide field versus gate voltage in the context of BTI

Oxide electric field (E_{OX}) inside the gate stack (E_{OX}) depends not only on the applied gate voltage V_G but also on the properties of the gate stack such as the V_T and EOT. It is calculated as:

$$E_{OX} = 10 \cdot (V_G - V_T)/EOT \qquad 3.3$$

where E_{OX} is in MV/cm and EOT is in nm. Oxide field in the high-k and the IL will be different, and here E_{OX} is the effective oxide field that is related to the individual oxide fields in both layers.

La or Al dipoles cause V_T modulation, as seen in Figure 3.1 and Figure 3.2. From Eq. 3.3 this modulation in V_T will change, at the same gate voltage, the E_{OX}, when compared to the ref La or Al split (device without any additives). Therefore, different V_G have to be applied to keep the same E_{OX} for devices with different doses of La or Al. This is confirmed by band diagrams of devices with La-0 and La-4.5E14 at/cm^2 doses in Figure 3.5. These band diagrams have been calculated by simulations based on Poisson Schrödinger formalism and the methodology is presented in [136].

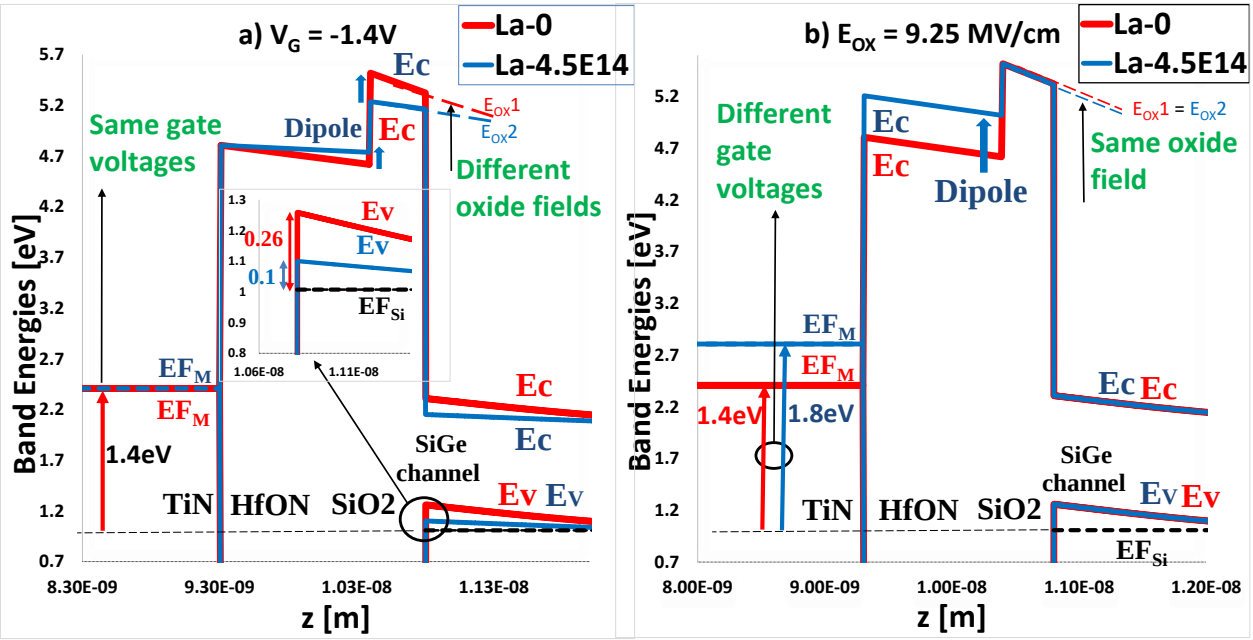


Figure 3.5: Band diagrams obtained by Poisson Schrödinger formalism for devices with La-0 and La-4.5E14 atoms/cm^2 doses at a) Same stress voltage (-1.4V), showing different E_{OX} and EF_{Si}-Ev and b) Two different bias conditions corresponding to the same Eox showing that the EF_{Si}-Ev are also the same

In Figure 3.5a, simulations at a same V_G = -1.4V are presented which evidences that E_{OX} at interfaces and EF_{Si}-Ev (difference between semiconductor Fermi level EF_{Si} and valence band Ev), are different. On the contrary in Figure 3.5b, simulation at two different bias (-1.4 V for La-0 and -1.8 for La-4.5E14 devices) corresponding to the same Eox shows that EF_{Si}-Ev are then at the same position.

In literature, NBTI and PBTI are often expected to depend upon the oxide electric field [83] [78] [79] [95] and not directly on the V_G, as has been discussed in section 1.5.1 on the BTI mechanisms and models. NBTI is attributed to hole capture into the gate oxide and takes place via a multiphonon emission (MPE) process characterized by its exponential electric field dependence. Application of an electric field shifts the total energy of holes in the valence band state and hence increases their transition probability into the oxide [79]. PBTI is attributed to electron trapping mechanism in localized defects in the High-k oxide, which is enhanced by the electric field.

If we assume that BTI is driven by the oxide field then BTI values should be compared at the same E_{OX}. If it's not the case and BTI depends directly on V_G instead of E_{OX} or in applications for which common V_G is used in a circuit having devices with unequal V_T, comparison at the same V_G will be useful. Depending on the way we intend to use the BTI results, either at same V_G or at same E_{OX} to calculate ΔV_{T5y} and T_{50mv}, two extrapolations can be made from the V_G power law of Eq. 3.2.

- First, assuming an E_{OX} dependence, we will compare the BTI results at the same reference E_{OX}. This E_{OX} corresponds to the devices with La-0 Å or Al-0 Å doses at V_G = + 0.945V

or - 0.945V for NMOS or PMOS respectively (given in Table 3.2). For devices with different La and Al doses, V_G corresponding to these reference E_{OX} is then calculated from Eq. 3.3 and then is used in the power law of Eq. 3.2 to calculate ΔV_{T5y} and T_{50mv}.

Reference Device	Reference Oxide field
1) La 0Å, PMOS (NBTI)	4.6 MV/cm
2) La 0Å, NMOS (PBTI)	6.1 MV/cm
3) Al 0Å, PMOS (NBTI)	4.6 MV/cm
4) Al 0Å, NMOS (PBTI)	4.1 MV/cm

Table 3.2: Reference oxide fields to compare NBTI and PBTI results for various La and Al doses

- Secondly, assuming a V_G dependence, extrapolations from the power law will be done at Vg = 0.945 V or -0.945 V for the NMOS or PMOS devices respectively.

The reported values of ΔV_{T5y} are with respect to the ΔV_{T5yref} (for La-0 or Al-0), i.e ΔV_{T5y}-ΔV_{T5yref}. The reported values of T_{50mv} are normalized by $T_{50mvref}$ (for La-0 or Al-0), i.e T_{50mv}/ $T_{50mvref}$.

3.1.3 Results

In this section, the effects of La and Al additives on NBTI and PBTI threshold voltage shift ΔV_T will be presented. These results will be compared at the same oxide field and at the same gate voltage. Further, a physical interpretation of the defects inside the HKMG stack will be made in order to understand and explain these results.

Effect of La on NBTI

NBTI tests were performed on P-MOSFET devices containing La doses shown in Table 3.1 and V_T shifts ($\Delta V_T < 0$) shown in Figure 3.1a, introduced by the sacrificial gate process. The devices were stressed at different stress biases for stress time of 1000 s, and the V_T shifts were measured at many points in between. Figure 3.6 shows NBTI V_T shift versus E_{OX} after 1000 s of stress for devices with different dose of La, ΔV_T increases with E_{OX} and La dose. The absolute values of V_T shift are not shown in this work due to confidentiality reasons.

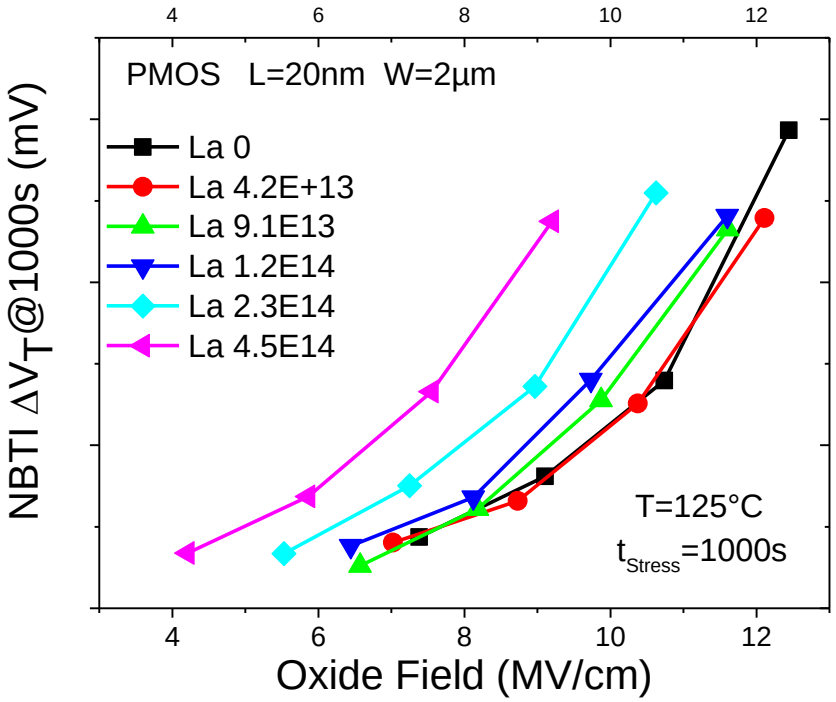


Figure 3.6: NBTI V_T shift vs. E_{OX} after 1000 s for devices with different effective La doses. La addition causes enhancement of NBTI V_T degradation

From the ΔV_T versus time measurements at different bias, similar analysis as done in Figure 3.4 was performed, for devices with different La doses. From this analysis a power law is defined for each device with parameters $A(V_G)$ and b. Then, the V_G corresponding to the ref E_{OX} = 4.6MV/cm (from Table 3.2) was calculated for each device and used in their respective power law to obtain the extrapolated parameters ΔV_{T5y} and T_{50mv} at ref E_{OX} = 4.6MV/cm. Figure 3.7 shows the ΔV_{T5y} and T_{50mv} (with respect to the reference devices in Table 3.1) for devices with various La doses and as in Figure 3.6, significant enhancement of the NBTI is observed due to La addition. As discussed on NBTI mechanisms in section 1.5.1, NBTI degradation is caused due to hole trapping in preexisting defects in the interlayer [78] [79]. Enhancement of NBTI due to the addition of La might mean that La creates defects in the interlayer due to their mixing to form La silicate.

Figure 3.8 shows the power law parameter $A(V_G)$ at the ref E_{OX} = 4.6MV/cm (normalized by the reference device) and time exponent b, calculated for each device by fitting the experimental data. Monotonic increase of A and decrease of b is observed with La dose. Increase of $A(V_G)$ shows the NBTI enhancement with La dose. As discussed in NBTI mechanisms, exponent b is dependent in a complex way on the interface states generation (by hydrogen transport) and the hole trapping, with the later contribution being dominant. If we assume that dependence of b on hole trapping does not change with La, then its decrease with La dose might mean that La addition in the interlayer decreases this hydrogen diffusion process [83].

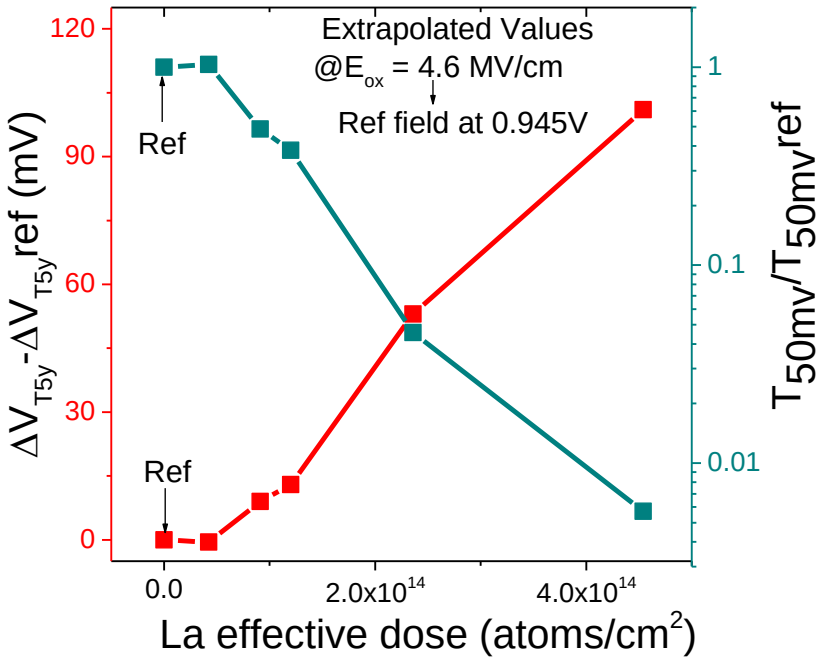


Figure 3.7: NBTI V_T shift @5years (ΔV_{T5y}) and time to failure@50mv (T_{50mv}) versus effective La dose, at reference E_{OX} = 4.6MV/cm

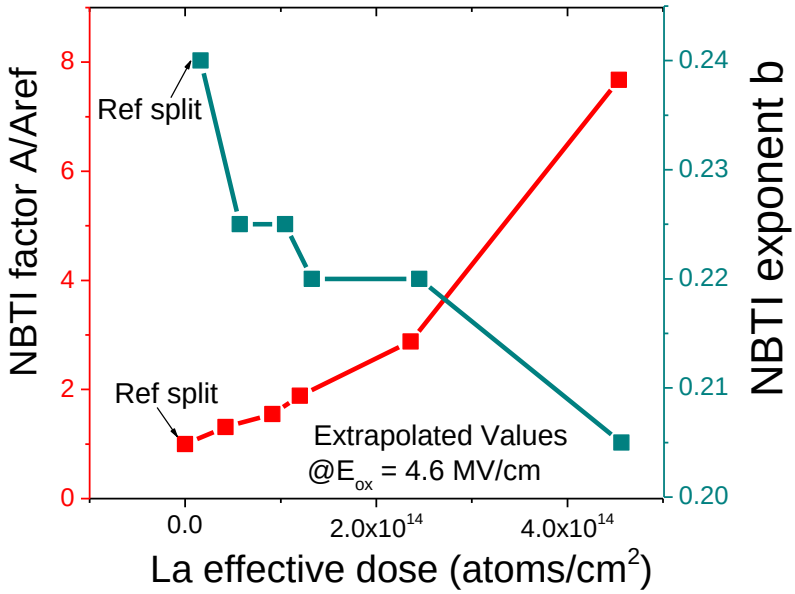


Figure 3.8: Variation of power law factor A and exponent b with La dose

The results can also be analyzed at the same gate voltage. Figure 3.9 shows NBTI V_T shift versus V_G after 1000 s of stress and

Figure 3.10 shows ΔV_{T5y} and T_{50mv} at V_G = -0.945 V, for devices with different doses of La (Same characteristics as Figure 3.6 and Figure 3.7 but with V_G instead of E_{OX}).

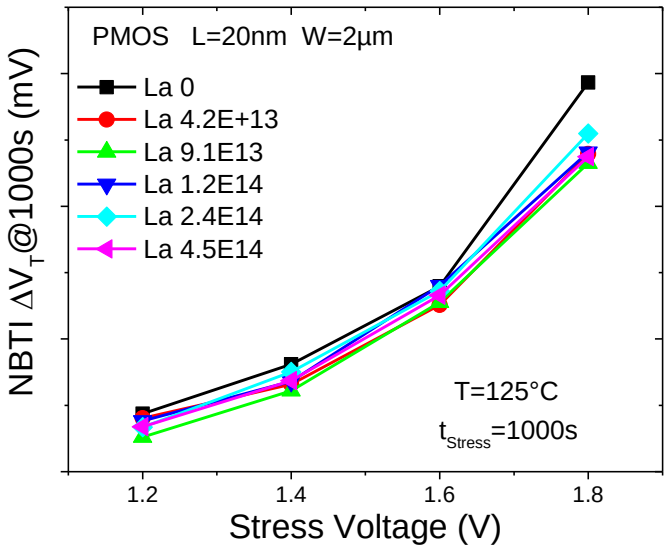


Figure 3.9: NBTI V_T shift vs. V_G after 1000 s for devices with different effective La doses. La addition causes slight decrease in NBTI V_T degradation

It is observed then that only a slight reduction of NBTI V_T shift is caused by La addition. It is important to highlight here that addition of La causes an increase of E_{OX} at the same gate voltage due to dipole formation (Figure 3.5). These results can be interpreted by two different ways:

1) If BTI mechanism is driven by E_{OX}, the results from Figure 3.9 and
2) Figure 3.10 would mean that at the same V_G NBTI degradation enhancement caused due to increasing La dose is mostly compensated by the decreasing E_{OX} due to dipole formation with La dose.

3) Another possibility is that NBTI degradation is in fact driven by the gate voltage instead of the oxide field, and so at the same gate voltage negligible effect of La would be observed.

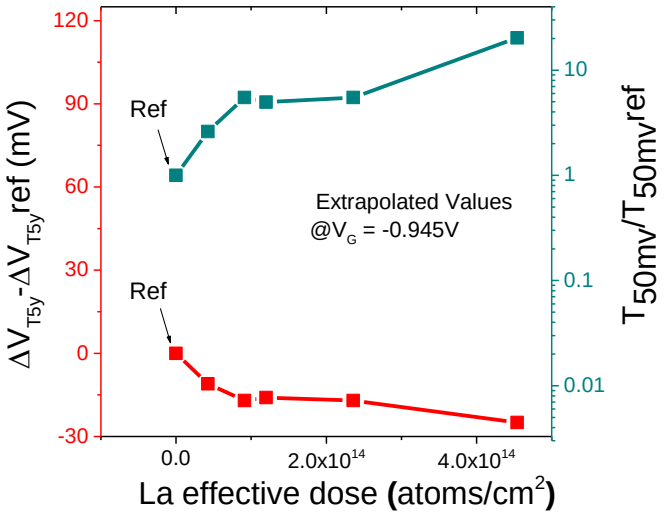


Figure 3.10: NBTI V_T shift @5years (ΔV_{T5y}) and time to failure@50mv (T_{50mv}) vs. effective La dose, at gate voltage V_G = -0.945V

The conventional BTI method used above might be impacted by the relaxation of stress generated defects between their creation and measurement, and so of BTI, due to the relatively large measurement time period (>1ms) compared to the fast BTI method with measurement time of 7 µs [95]. In order to show that the conventional method is valid in case of our devices, NBTI tests were also performed by the fast BTI method for devices with La-0Å and La-4Å. From Figure 3.11, it is clear that results from both methods are almost the same due to negligible relaxation, which means that the conventional technique is valid for the devices used in this work.

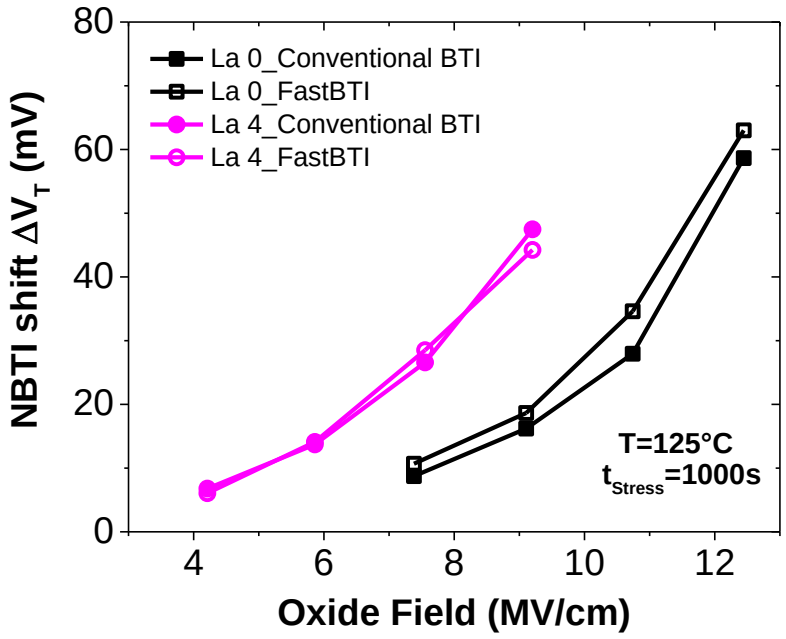


Figure 3.11: NBTI V_T shifts with oxide field for conventional and fast NBTI methods on La 0 Å and La 4 Å splits

Effect of La on PBTI

PBTI tests were performed on various N-MOSFET devices having the same La dose splits as used for P-MOSFET device in the NBTI tests. The doses are shown in Table 3.1 and V_T shifts ($\Delta V_T > 0$) shown in Figure 3.1 b. The devices were stressed at different stress biases for 1000 s and the V_T shift was measured at many points in between. Figure 3.12 shows PBTI V_T shift versus E_{OX} after 1000 s of stress for devices with different doses of La, its value increases with oxide field but decreases with La dose, contrary to the case of NBTI. Power law parameters A (V_G) and b are obtained by fitting the experimental data. ΔV_{T5y} and T_{50mv} are calculated for devices with various La doses at ref E_{OX} = 6.1 MV/cm (from Table 3.2), their variation with La dose is plotted in Figure 3.13. Figure 3.12 and Figure 3.13 evidence a preventive effect of La on PBTI ($\Delta V_{T5y} - \Delta V_{T5yref} < 0$). Nevertheless the amplitude of this effect is much smaller than the effect of La on NBTI.

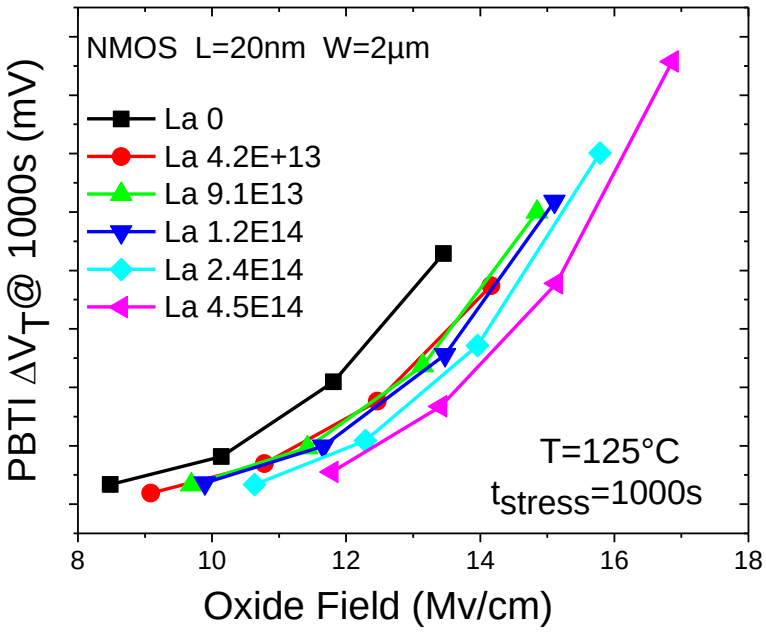


Figure 3.12: PBTI V_T shift vs. E_{OX} after 1000 s for devices with different effective La doses. La addition causes reduction of PBTI V_T degradation

As discussed in section 1.5.1, PBTI mechanism is caused by electron trapping in the defects present in the high-k. The curing effect due to La addition might be due to the passivation of oxygen vacancies in HfON by La [137]. This effect does not seem to depend linearly on the La dose but at first on its deposition process. From Table 3.1 and Figure 3.13, when La is with pedestal TiN (splits 2, 3 and 4) the values seem to saturate, it then decreases and saturates again for splits in which La is placed directly on HfON (splits 5 and 6). This might be due to the presence of another effect, other than vacancy passivation by La, that is impacting PBTI and which depends on the interface between high-k and the sacrificial gate stack (pedestal TiN or not).

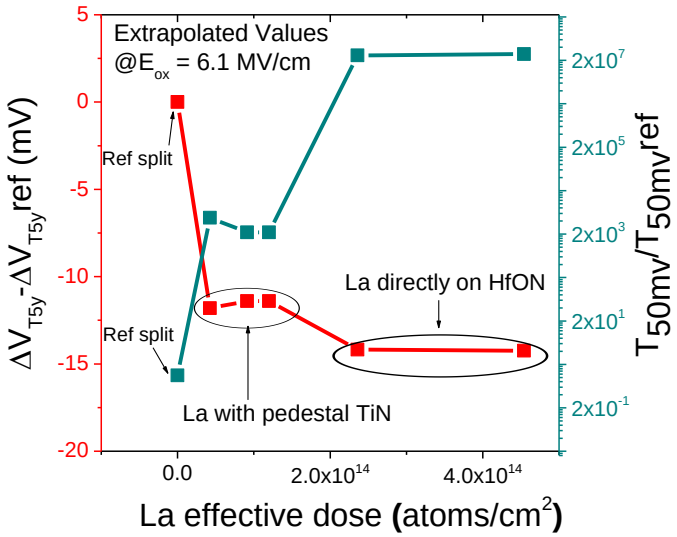


Figure 3.13: PBTI V_T shift @5years (ΔV_{T5y}) and time to failure@50mv (T_{50mv}) vs. effective La dose, at reference E_{OX} = 6.1MV/cm

If we compare the experimental PBTI results and extrapolated lifetimes at the same gate voltage, instead of the same E_{OX}, the results are different. Figure 3.14 shows PBTI V_T shift versus V_G after 1000 s of stress and ΔV_{T5y} & T_{50mv} at V_G = 0.945 V, for devices with different doses of La. The effect of La dose is not monotonous but can be explained by similar reasoning as done for NBTI results at the same V_G: La addition on NMOS devices will increase the oxide field (NMOS V_T decreases with La dose) but decreases the PBTI degradation (curing effect), so at the same V_G these two opposite effects will give rise to the PBTI trend seen in Figure 3.14 b.

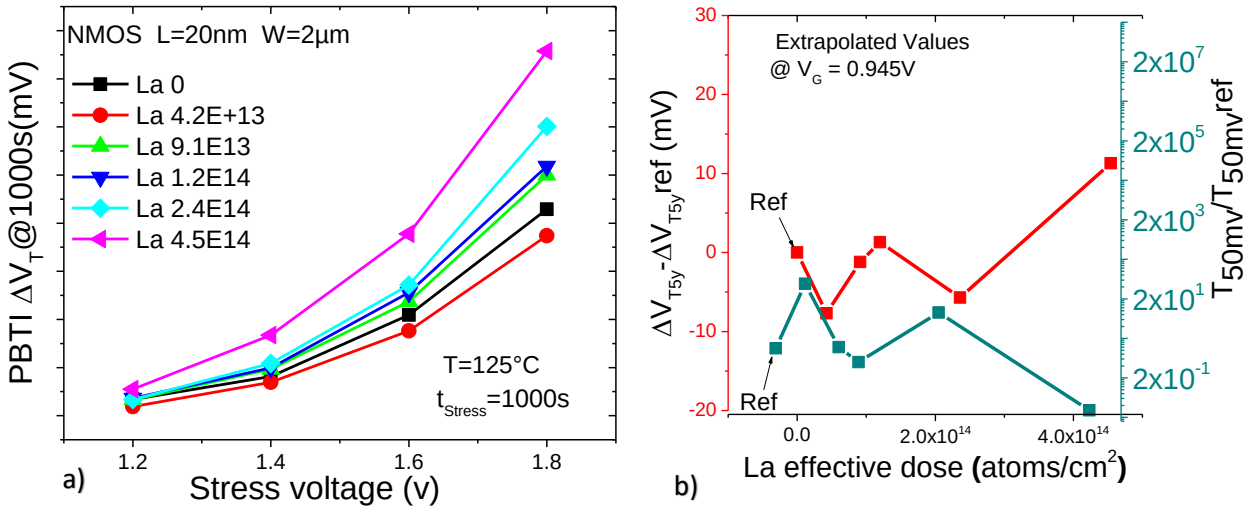


Figure 3.14: a) NBTI V_T shift vs. V_G after 1000 s for devices with different effective La doses; b) NBTI V_T shift @5years (ΔV_{T5y}) and time to failure@50mv (T_{50mv}) vs. effective La dose, at gate voltage V_G = 0.945V

Effect of Al on NBTI

NBTI tests were performed on P-MOSFET devices containing different Al doses. These doses correspond to the splits 1, 2, 3 and 5 in Table 3.1 with variations in TiN pedestal thickness (Al-0, TiN-25Å /Al-2.2Å, TiN-20Å/Al-2.2Å, TiN-10Å /Al-2.2Å), having V_T shifts shown in Figure 3.2a. The highest Al dose in this study is 4.14E14, which is comparable to the maximum La dose (4.5E14) studied earlier. However, due to roll-off effect on the effective work function at nominal EOT, Al is less effective compared to La and so the shift on V_T with Al dose is smaller [47].

Similar to devices with La, NBTI tests were performed and V_T shifts ($\Delta V_T < 0$) were measured with stress time. Figure 3.15 shows ΔV_T after 1000 s of stress versus E_{OX} for devices containing different Al doses. Slight enhancement of NBTI shift with Al dose is observed, which is much smaller than with La dose seen earlier. As discussed earlier, NBTI depends on the defects inside the interlayer and such a small dependence of NBTI with Al dose can be explained by the non-reactive nature of Al with this interlayer. Power law parameters A(V_G) at the ref E_{OX} = 4.6MV/cm and b, calculated by fitting the experimental data, are plotted in Figure 3.16. Compared to the case of device with La, fairly constant b and slight increase in A with Al dose is observed. Similar to La, we assume that dependence of b on hole trapping does not change with Al. Then, the constant

value of b with Al dose might be due to the fact that Al does not react with the interlayer and hence doesn't impact the hydrogen transport.

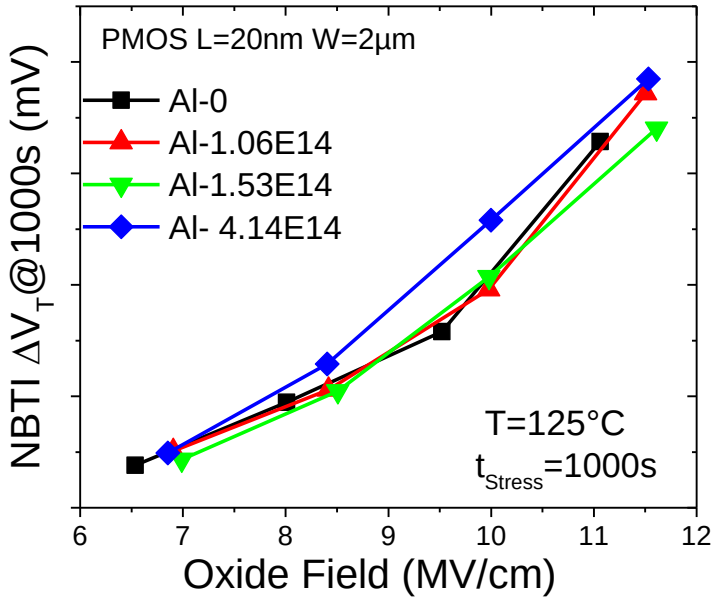


Figure 3.15: NBTI V_T shift vs. E_{OX} after 1000s for devices with different effective Al doses. Al addition causes slight enhancement of NBTI V_T degradation

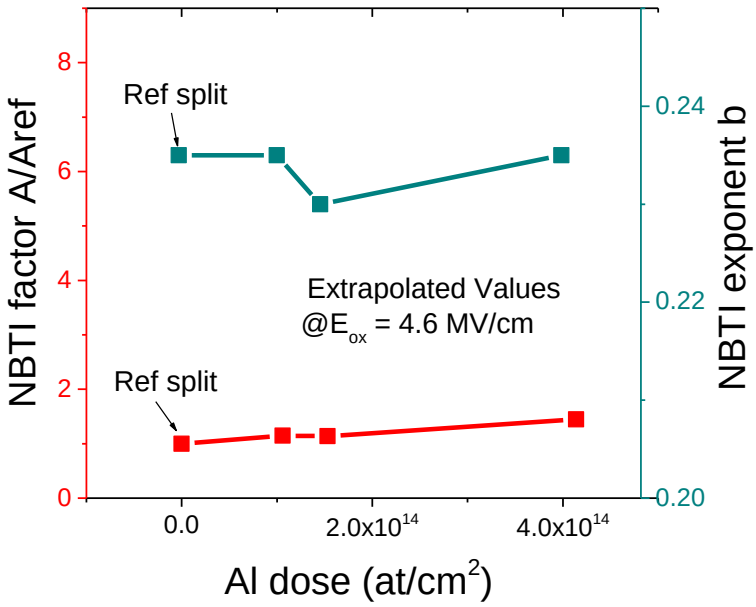


Figure 3.16: Variation of power law factor A and exponent b with Al dose

In order to investigate the difference in Al impact on NBTI at the same oxide field or at the same gate voltage, Figure 3.17 compares the ΔV_{T5y} versus Al dose at the same EOX = 4.6 MV/cm and at the same V_G = -0.945 V. Slight enhancement in NBTI is observed in both cases which is very different from La impact on NBTI for which analysis in V_G or E_{OX} were very different. This

similarity of results at the same V_G and E_{OX} for Al is due to the fact that the V_T modulation by Al addition (100 mV) is much lower than by La (380 mV).

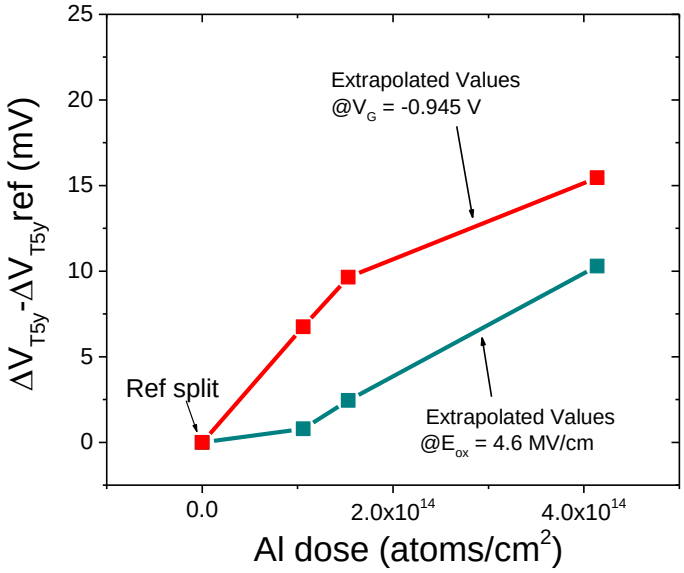


Figure 3.17: Comparison of NBTI V_T shift @5years (ΔV_{T5y}) vs. effective Al dose, at gate voltage V_G = -0.945V and reference EOX = 4.6MV/cm

Effect of Al on PBTI

NBTI tests were performed on N-MOSFET devices containing different Al doses. These doses correspond to the splits 1, 4 and 5 in Table 3.1 with variations in as-deposited Al thickness (Al-0, TiN-10Å/Al-6.6Å, TiN-10Å/Al-2.2Å), having V_T shifts shown in Figure 3.2 b. As in the previous section on NBTI, the highest Al dose in this study is 4.14E14. Figure 3.18 shows NBTI V_T shifts after 1000 s of stress versus E_{OX}.

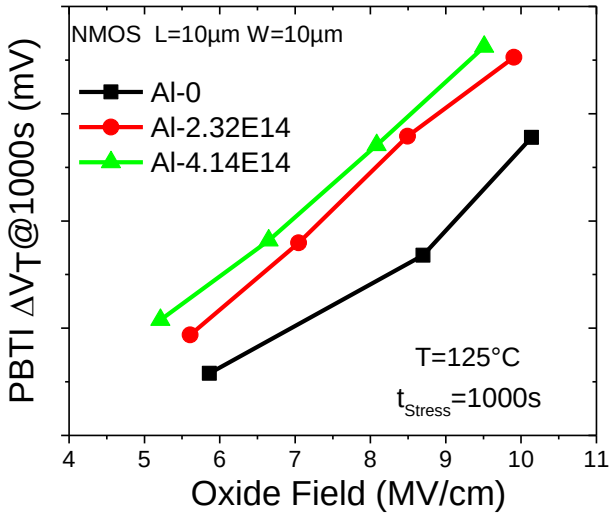


Figure 3.18: PBTI V_T shift vs. E_{OX} after 1000s for devices with different effective Al doses. Al addition causes enhancement in PBTI V_T degradation

By fitting the experimental ΔV_T versus stress time data, power law parameters $A(V_G)$ and b are determined and are used to calculate ΔV_{T5y} and T_{50mv} for various Al doses at ref E_{OX} = 4.1 MV/cm (from Table 3.2), these are shown in Figure 3.19. The effect of Al on PBTI is obvious from the analysis of these figures, showing linear increase of PBTI in a linear-log scale evidencing the drastic impact of Al dose on PBTI. As said earlier for the case of La effect on PBTI, it is caused by the defects in the high-k, the significant increase of PBTI due to Al addition might mean that Al creates such defects inside the high-k.

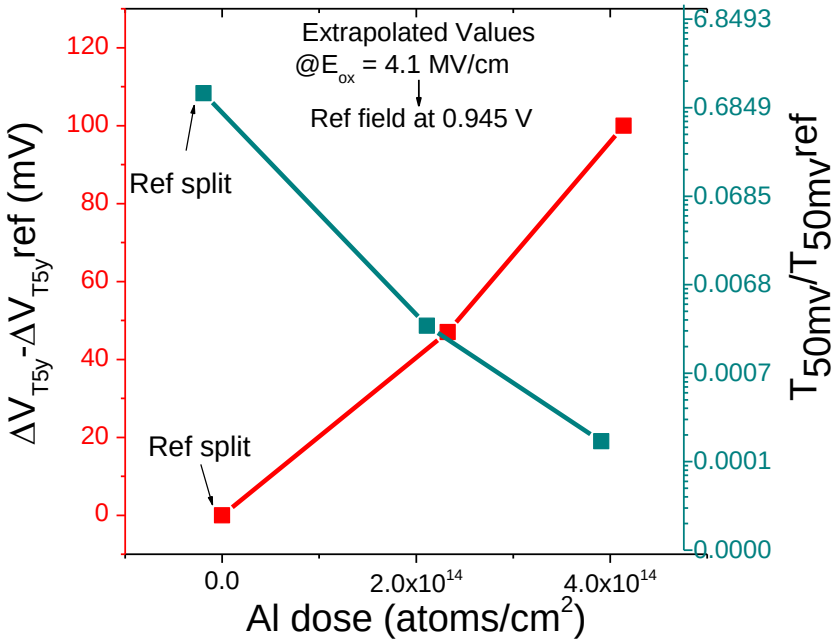


Figure 3.19: PBTI V_T shift @5years (ΔV_{T5y}) and time to failure@50mv (T_{50mv}) vs. effective Al dose, at reference E_{OX} = 4.1 MV/cm

When the results for ΔV_{T5y} are compared at the same oxide field and at the same gate voltage, PBTI due to Al addition is enhanced compared to results at the same oxide field (Figure 3.20). This is due to the increase in V_T (Figure 3.2 b) and so decrease in oxide field with Al addition (dipole formation). The difference between results at the same oxide field and at the same gate voltage is still much less then what was observed for La effect on PBTI, as V_T modulation by Al addition is lower than by La.

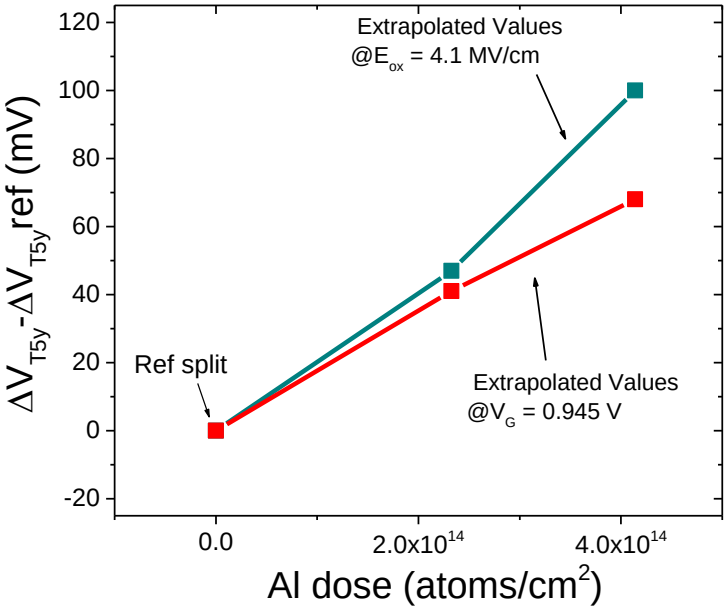


Figure 3.20: Comparison of PBTI V_T shift @5years (ΔV_{T5y}) vs. effective Al dose, at gate voltage V_G = 0.945V and reference E_{OX} = 4.1 MV/cm

Physical interpretation

In this study, we have compared the effect of Al and La addition on NBTI and PBTI by two approaches: by assuming an electric field E_{OX} dependence or by considering a gate voltage V_G dependence. Assuming an E_{OX} dependence, a monotonic trend is observed for the NBTI and PBTI factors (A, b, ΔV_{T5y} and T_{50mv}) as a function of La and Al doses in Figure 3.7, Figure 3.8, Figure 3.13, Figure 3.17, Figure 3.19 and Figure 3.20 [PK-1]. It shows that NBTI and PBTI could be affected by La and Al diffusion in gate dielectrics, leading either to defect creation or passivation. Table 3.3 summarizes these effects at the same oxide field or at the same gate voltage.

		NBTI	PBTI
La	Same Eox	Significant ↑	Slight ↓
	Same Vg	No impact	No impact
Al	Same Eox	Slight ↑	Significant ↑
	Same Vg	Slight ↑	Significant ↑

Table 3.3: Summary of the impact of La and Al on NBTI and PBTI, at a same oxide field and at a same gate voltage

La effect on NBTI shift is significant and shows a linear trend with La dose. In literature, NBTI mechanism is expected to depend on defects in the interlayer [78] and La is known to react with the IL [138]. Therefore, NBTI degradation with La can be understood as creation of precursor defects in IL during La silicate formation (Figure 3.21) that are responsible for enhancing the hole capture phenomena and so the NBTI effect. Al effect on NBTI is much lower than La at the same

dose as Al is not known to react with the IL and only remains and creates dipoles at the interface between IL and HfON (Figure 3.21). Assuming that the dependence of b on hole trapping does not change with La or Al, then exponent b can be seen as a measure of hydrogen transport through the IL which is thought to be responsible for the NBTI degradation [83]. Its decrease with La dose and constant trend with Al dose (Figure 3.8 and Figure 3.16) might mean that La silicate formation in the IL affects such diffusion.

PBTI is expected to be caused by electron trapping in the pre-existing or stress induced defects in the high-k layer. La reduces PBTI only slightly and shows a preventive effect likely due to passivation of oxygen vacancies in HfON by La [137]. Al drastically enhances the PBTI, possibly due to creation of defects in HfON rather than passivation, as shown in Figure 3.21.

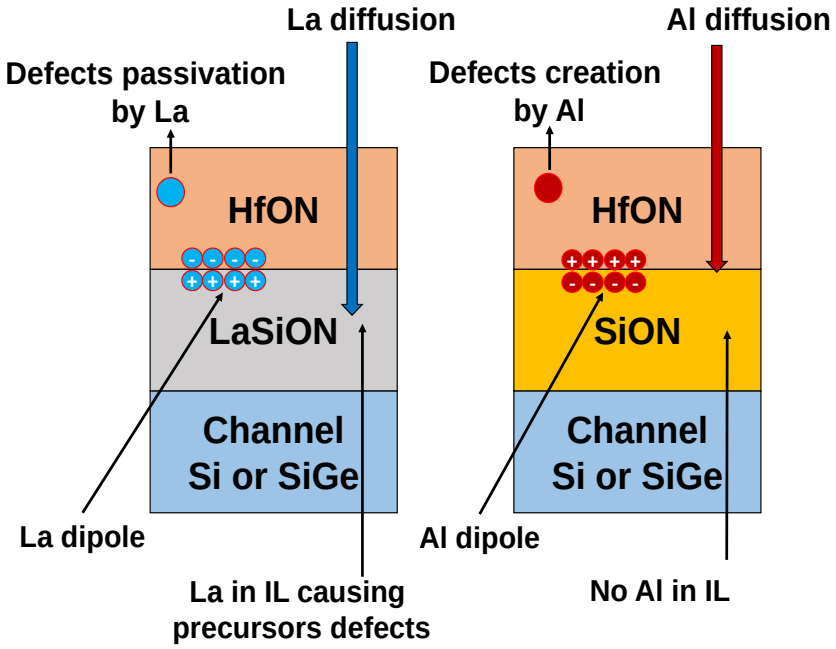


Figure 3.21: Schematic showing La and Al diffusion in the gate stack and its consequence on dipole and defects creation

When the results are compared at the same gate voltage, very different trend compared to results at the same oxide field is observed for devices with La addition. This is due to the decrease or increase in oxide field due to La addition, at the same gate voltage, in the case of NBTI or PBTI respectively (V_T increases for PMOS and decreases for NMOS). This change in electric field have an opposite impact on BTI, to the intrinsic BTI effect caused by La or Al. For the case of Al addition, the results at the same gate voltage are quite similar to the ones at the same oxide field due to smaller V_T modulation by Al compared to La.

Classically La is used in NMOS, and there seems to be no drastic degradation for PBTI with La, but it will be a problem if La is used in PMOS for future technologies. Similarly, Al is classically used in PMOS and we only observed a slight degradation on NBTI so it will not be a big issue for PMOS but will be a problem if it used in future NMOS devices.

3.2 Impact of La and Al additives on TDDB reliability

TDDB tests were performed on the same FDSOI N-MOSFET and P-MOSFET devices that have been used for BTI studies in section 3.1, and the process flow is as given in section 3.1.1. La and Al additives were introduced for V_T tuning into the HKMG stack by the sacrificial gate process. The splits for La or Al and bottom pedestal TiN thickness and their respective effective diffused doses are given in Table 3.1.

3.2.1 Electrical measurements

TDDB measurements and breakdown detection

In order to study the oxide breakdown phenomena, constant voltage stress (CVS) method is used. Therefore, a high stress voltage $V_{G,stress}$ is applied on the gate with a high temperature (125°C), while the source and drain terminals are grounded (drain current ~ 0) and the leakage current through the gate oxides is measured (Figure 3.22 a). During stress, at predefined points of time, stress voltage is interrupted and V_G is lowered to a lower value $V_{G,\,sense}$ (= ±1V) for which the gate current (I_{sense}) is measured (Figure 3.22 b).

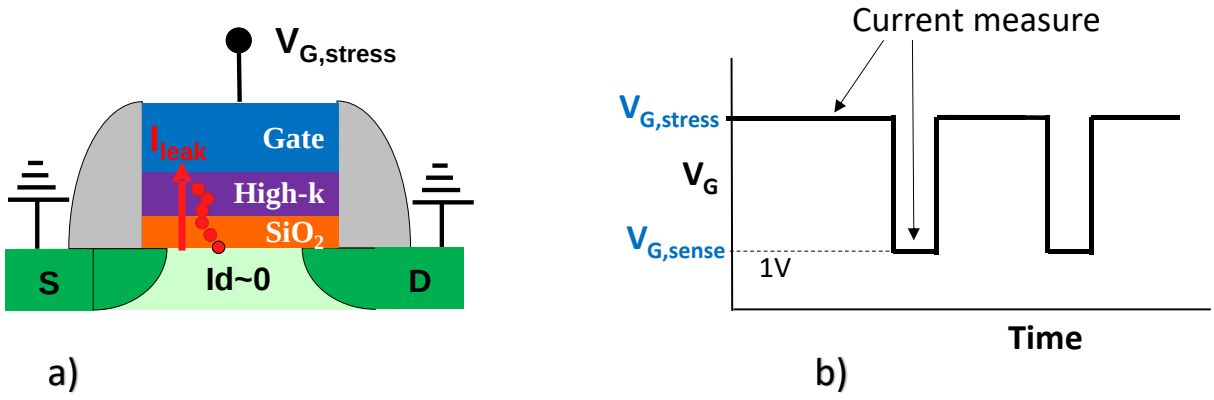


Figure 3.22: a) Device biasing by a constant gate voltage during a TDDB measurement; b) Variation of the gate bias between stress and sense voltages to measure the gate current

These measurements are performed on a large number of devices (~ 50) in order to have a good time to breakdown distribution. As shown in Figure 3.23, the gate current measured at the stress voltage (I_{stress}) and at the sense voltage (I_{sense}) first varies over time due to stress induced leakage current (SILC). After a certain time that is statistically distributed for different devices, a large abrupt jump in current (in I_{stress} and /or I_{sense}) is observed which can either be a hard breakdown (HBD) or a soft breakdown (SBD). The device is considered to be broken when it satisfies one of the following criteria described below:

1) I_{stress} increases by a factor of 100 or decreases by a factor of 10.

2) I_{sense} increases by a factor of 100 or decreases by a factor of 10.
3) $I_{sense} > I_{sense}^{max}$

where I_{sense}^{max} is the maximum exceptable value of I_{sense} until failure, its value is 100 times the value of I_{sense} before stress($I_{sense}^{max} = 100 \cdot I_{sense}^{initial}$). The stress times corresponding to these breakdown events are obtained and are called the time to breakdown.

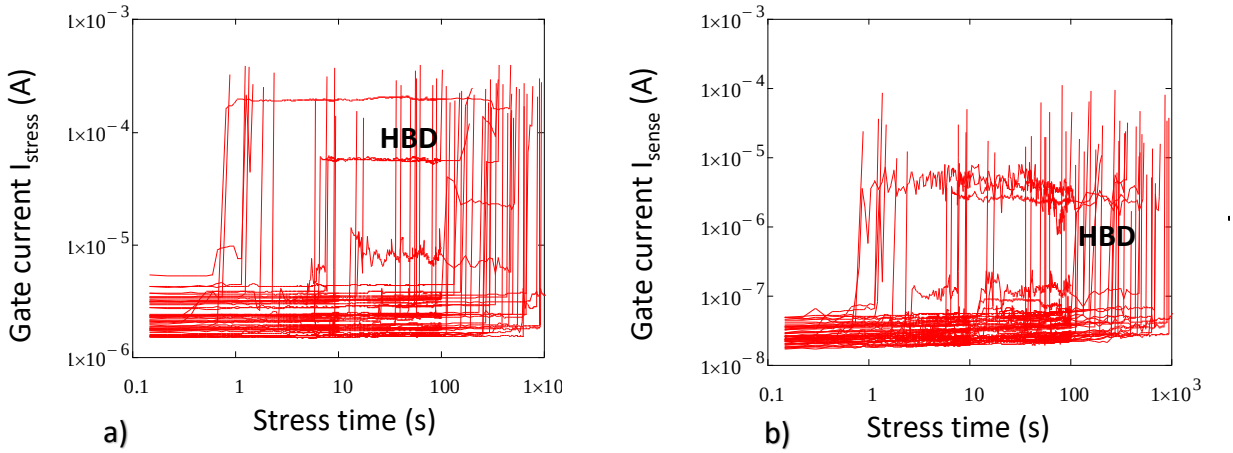


Figure 3.23: Variation of the gate current with stress time evidencing its increase due to SILC, soft and hard breakdown for PMOS devices measured at a) stress voltage or at b) sense voltage of -1V

The measurements are performed for different stress voltages and a large number of devices (about 50) are used for each voltage. The times to breakdown are statistically distributed and well described by the Weibull distribution function W (see section 1.5.2). In Figure 3.24, the cumulative time to breakdown distribution (F(t)) is plotted in a Weibull scale $W(F(t)) = \ln(-\ln(1-F))$ as defined in Eq. 1.19, for different gate stress voltages.

The distribution of time to failure fit the Weibull distribution over a large range of time. From such a fit, the slope of the fit (β) and the stress time at which 63% of devices have failed (T_{BD} or η) are obtained. T_{BD} decreases as the stress voltage is increased. β remains fairly constant for low stress voltage values but it increases for higher values of stress. Increased β values are dependent on statistics (number of devices used, size of device, correct detection of breakdown, etc.) and the model used to fit the distribution, i.e. sometimes a bimodal behavior is a better fit rather than a single model one. In this regard, T_{BD} values are more reliable and accurate (than β**)** in order to understand the breakdown mechanism. Thus, in our studies, T_{BD} will be used as the TDDB degradation parameter to compare different gate stack processes.

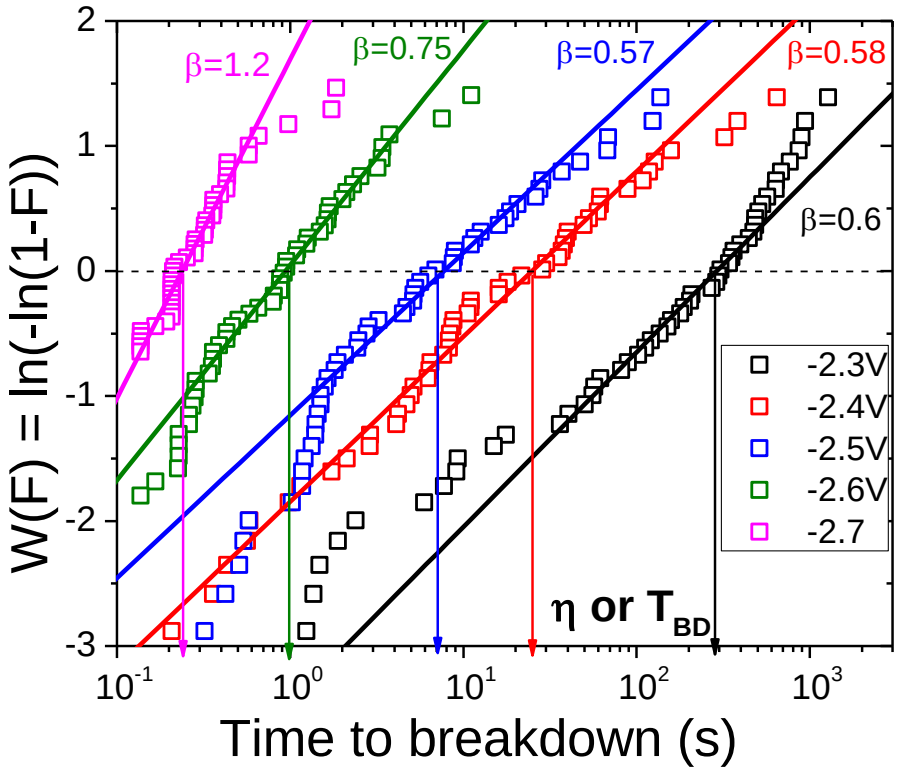


Figure 3.24: Cumulative failure distribution F(t) in a Weibull scale at different gate stress voltages

Determination of TDDB gate stress voltages

Selection of the proper range of gate stress voltages is important in order to have a sufficient number of data points (time to breakdown) that are distributed over many stress time scales. This means that the lower range of the stress voltages should be high enough so that majority of devices are broken before the end of the stress. Moreover, the upper range should be low enough so that the devices are not broken below the time resolution of the measurement.

In order to choose this proper range of voltages, several Ig (V_G) characteristics are plotted until breakdown, as shown in Figure 3.25. This is done on several devices to get a good statistics of the breakdown voltages. Then, a voltage range (shown by red crosses) is chosen to be a bit lower than the average device breakdown voltage.

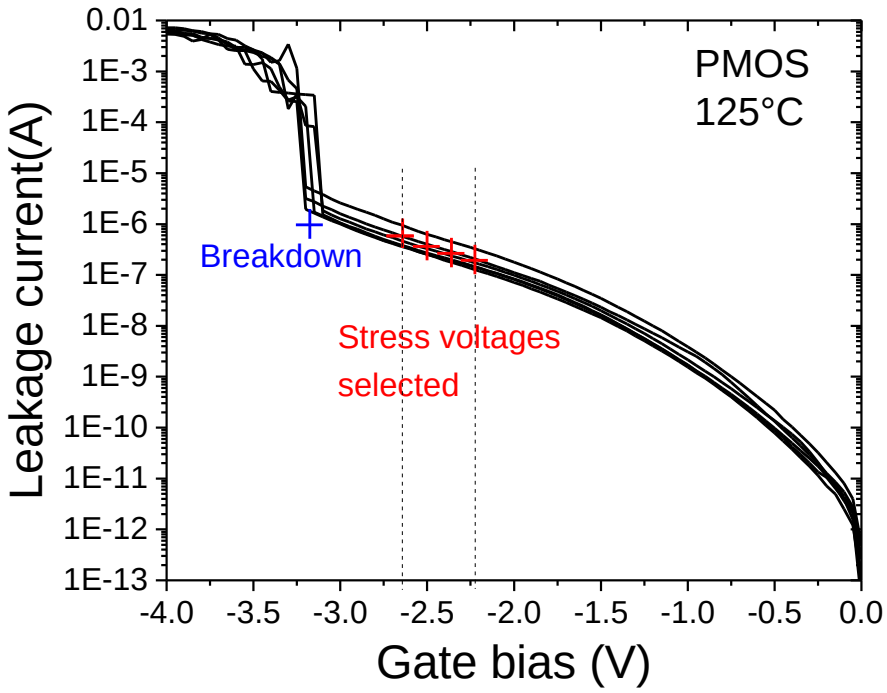


Figure 3.25: Variation of the gate current versus gate bias, showing breakdown and selection of the stress voltages for TDDB measurements

3.2.2 Results

Effect of Al on PMOS TDDB

TDDB tests were performed on P-MOSFET devices containing Al doses as used for NBTI studies in the previous section. These splits are with different pedestal TiN thickness (Al-0, TiN-25Å /Al-2.2Å, TiN-20Å/Al-2.2Å, TiN-10Å /Al-2.2Å) and corresponds to Al doses of 0, $1.06 \cdot 10^{14}$, $1.53 \cdot 10^{14}$, $4.14 \cdot 10^{14}$ at/cm^2, with V_T shifts shown in Figure 3.2 a. Stress voltages were selected, with the method illustrated in Figure 3.25, at -2.3V, -2.4V, -2.5V, -2.6V and -2.7V and 50 devices were stressed for each bias condition and Al dose. Temperature during measurements was set to 125°C. Each tested device is a set of 100 devices of width = 1 µm and length = 20 nm that are connected in parallel. This allows to provide a large effective device width (100 µm), that will increase the global tested surface and thus improving the sampling while still keeping the nominal gate width (1 µm) and length (20 nm) of individual devices.

For the case of devices with no Al dose, Figure 3.23 report the variation of current at stress and at sense voltage of -1V, for devices stressed at -2.3V, Figure 3.24 shows the Weibull plot for different gate stress biases. Figure 3.26 shows the Weibull plot of the device breakdown times for all Al doses at a same $V_{G,stress}$ of -2.4V. Some variation in the Weibull slope is observed, but more importantly the time to breakdown at 63% devices (T_{BD}) shows a remarkable increase as Al dose is increased.

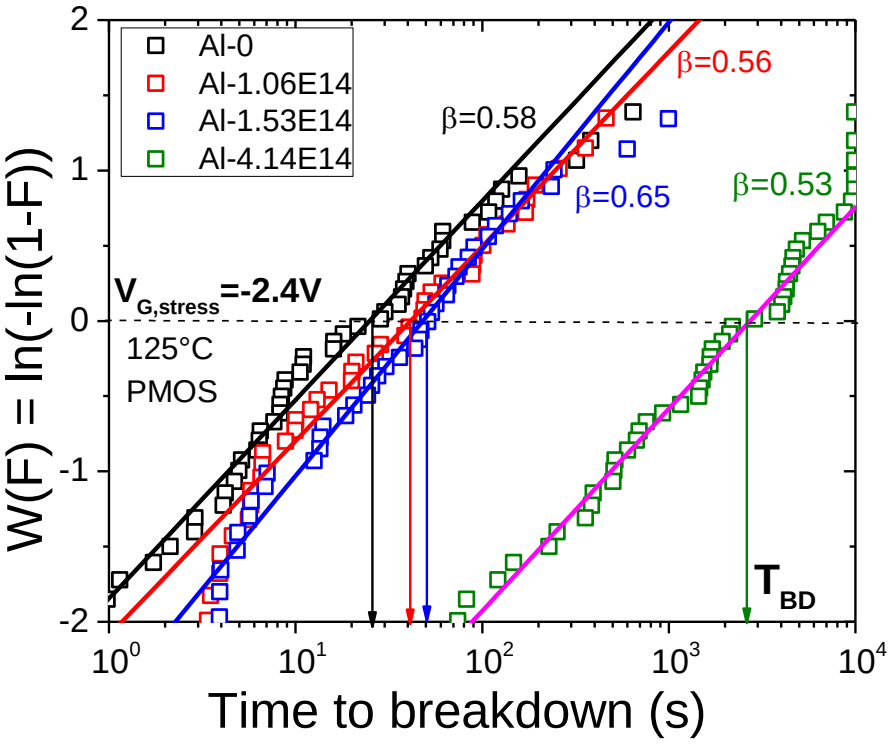


Figure 3.26: Cumulative failure distribution F(t) in a Weibull scale at gate stress voltage of -2.4V, for different Al doses

The Weibull plot analysis done above at $V_{G,stress}$ of -2.4V was done at other bias conditions and for each bias and Al dose, the T_{BD} was obtained. In Figure 3.27, T_{BD} versus stress voltage are plotted, for each Al dose. The observation done for $V_{G,stress}$ = -2.4V for which T_{BD} increases with Al dose is confirmed at every bias, with a significant increase in T_{DB} with the highest Al dose.

The T_{BD} can be extrapolated to any bias condition and depending on the kind of TDDB degradation mechanism considered (E model, AHI model, MVHR model, etc. as explained in section 1.5.2), different extrapolations can be made. In all these models, the general view is that the electrons from the gate or substrate are responsible for the gate oxide degradation and then to its breakdown. Studies have shown that below 65 nm technology and for the oxides with thicknesses used in our studies, the power law dependence with gate voltage seems to be quite accurate [97] [111] [139]. These studies show that the electrons from the gate cause oxide degradation and that can be enhanced by hydrogen and/or holes generation. The voltage power law is given as follows:

$$T_{BD} = T_{BD0} \cdot (V_G/V_{G0})^{-n} \qquad 3.4$$

where T_{BD0} is an exponential pre-factor and n is the voltage acceleration factor.

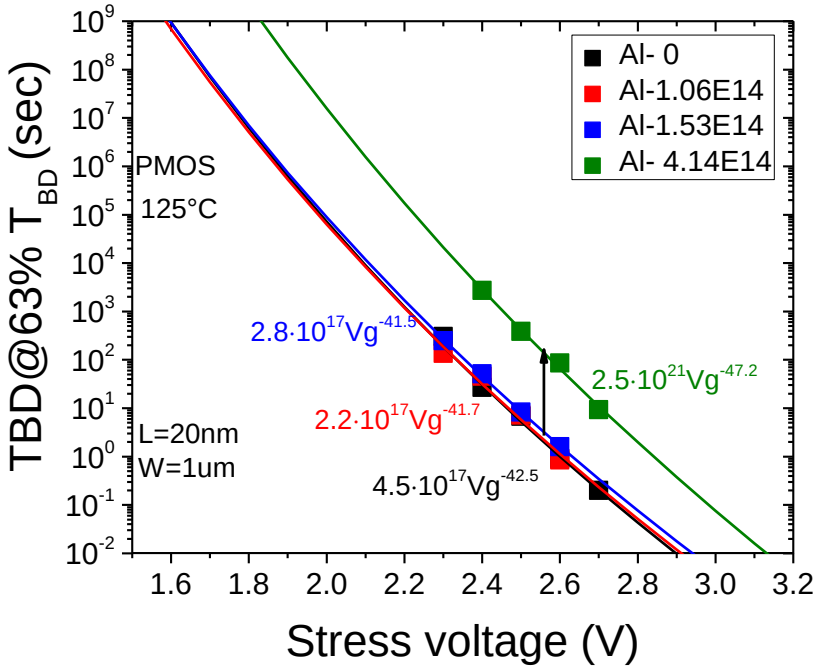


Figure 3.27: T_{BD} versus stress voltage for different Al doses and extrapolation by voltage power law

Power law extrapolations are done and shown in Figure 3.27. For the highest Al dose, T_{BD0} increases significantly and there is a slight increase in voltage accelaration factor n as well. The increase in T_{BD} with Al dose is due to a combination of increase in both these factors. Due to a higher n value for device with high Al content, the benefit of T_{BD} increase will be even higher at the nominal stress voltage (lower) values.

As mentioned earlier that breakdown of the oxide is related to the electons that crosses the oxide, and so the T_{BD} might be related to the electron current density at stress (I_{stress} or I_g) through the gate stack. In order to understand the effect of Al addition on T_{BD} from the aspect of the gate leakage current, first we have looked at the variation of leakage current with the effective Al dose in the gate stack. It is plotted in Figure 3.28, we notice that the leakage current decreases almost linearly with Al dose.

Next, in Figure 3.29, T_{BD} at different gate stress bias is plotted versus the leakage current for each Al doses. We have seen that leakage current is correlated to Al dose (Figure 3.28), but it is also expected to be a relevant parameter (Eq. 1.18). T_{BD} is seen to increase with an inverse power law of the leakage current, with a power law exponent between 1.7 and 1.9. From Figure 3.28 and Figure 3.29, it can be understood that addition of Al into the gate stack causes a decrease in the leakage current, which in turn might be responsible for lower defect creation in the gate oxide and thus to the lower breakdown time T_{BD}. In Figure 3.29, the T_{BD} versus leakage curves seems to follow a similar behavior at each bias, meaning that a universal law can be applied to estimate the T_{BD} for any particular Al dose, leakage current and stress voltage.

$$T_{BD} = C \cdot (I_g/I_0)^{-a} \cdot (V_G/V_{G0})^{-b} \qquad 3.5$$

where C is a constant having units of time and being independent of Al dose.

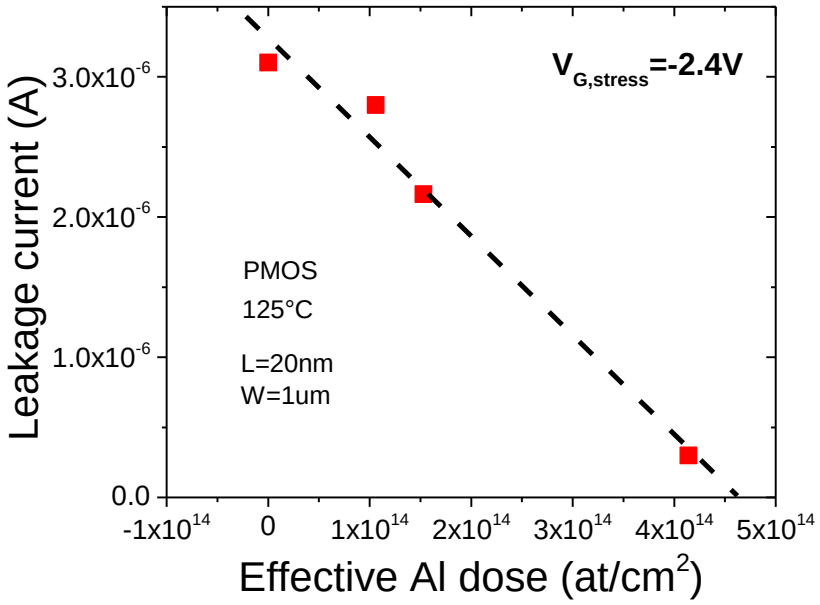


Figure 3.28: Variation of gate leakage current with Al dose at stress voltage of -2.4V

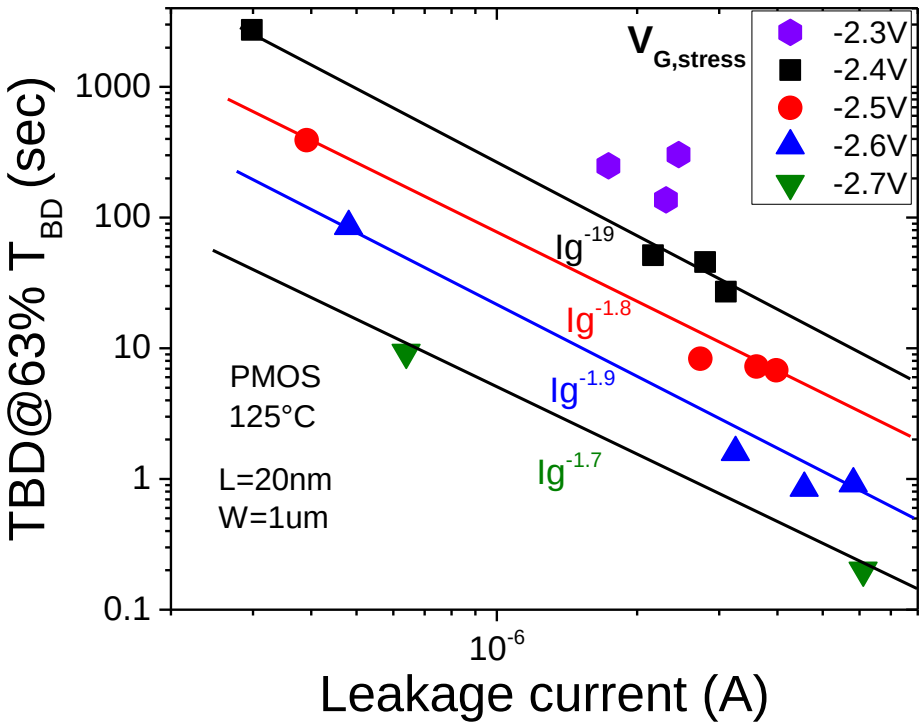


Figure 3.29: T_{BD} versus leakage current for different Al doses and its fit by a current power law

Effect of La on PMOS TDDB

TDDB tests were performed on P-MOSFET devices containing La doses as for BTI studies, without pedestal TiN (La-0, La2Å/TiN45Å and La4Å /TiN45Å) and corresponds to La doses of 0, $2.36 \cdot 10^{14}$, $4.5 \cdot 10^{14}$ at/cm^2 repectively. Stress voltages were selected at -2.2, -2.3, -2.4, -2.5V and 50 devices were stressed for each bias condition and La dose. Temperature during measurements was set to 125°C and device were 10 μm wide and 20 nm long.

Figure 3.30 shows the Weibull plot for devices with La-0 dose at different gate biases and, similar to the case with Al additive, the breakdown times fit with a Weibull distribution over a large range of time. β and T_{BD} can be calculated as explained in Figure 3.24. β value is not so consistent and increases for higher values of stress, compared to the case of Al for same stress voltage values. This might be due to a relatively lower quality of the extraction of the breakdown distribution because here we are not using multiple devices in parallel (device width = 10 um), compared to the case of Al where 100 device were used in parallel (effective device width = 100um). As said earlier, T_{BD} is the most important parameter for our studies, and from Figure 3.30 we see that its value decreases as the stress is increased. Figure 3.31 shows the Weibull plot of the device breakdown times for all La doses at $V_{G,\,stress}$ of -2.2V. Similar β is observed for different La doses. T_{BD} first decreases for first La dose but then increases for the second higher La dose, meaning that oxide degradation is not a monotonous function of the La dose.

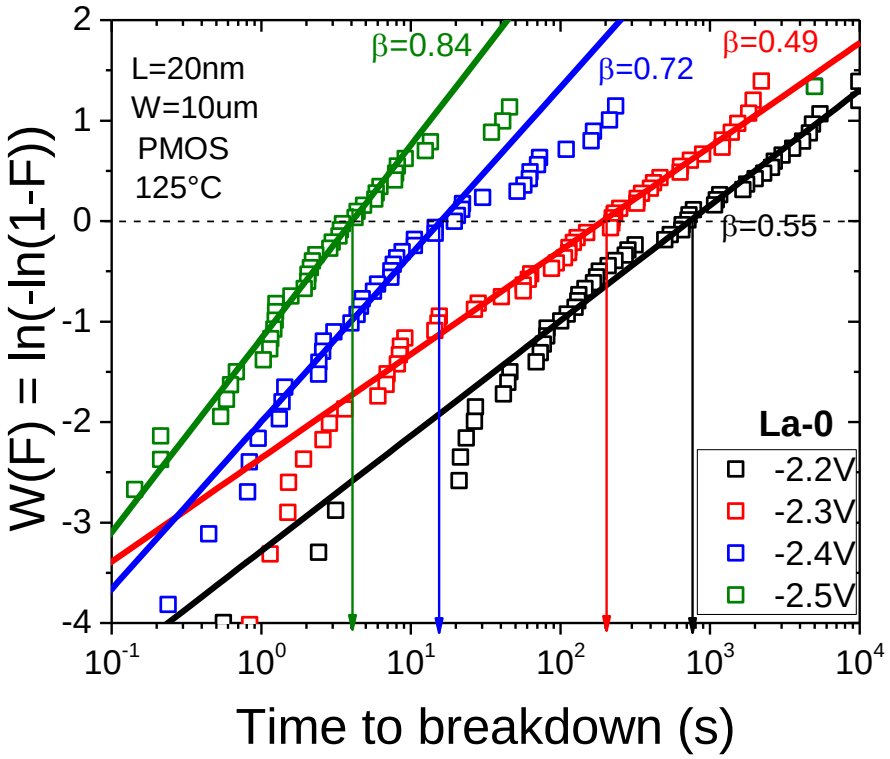


Figure 3.30: Cumulative failure distribution F(t) in a Weibull scale at different gate stress voltages, for La-0 devices

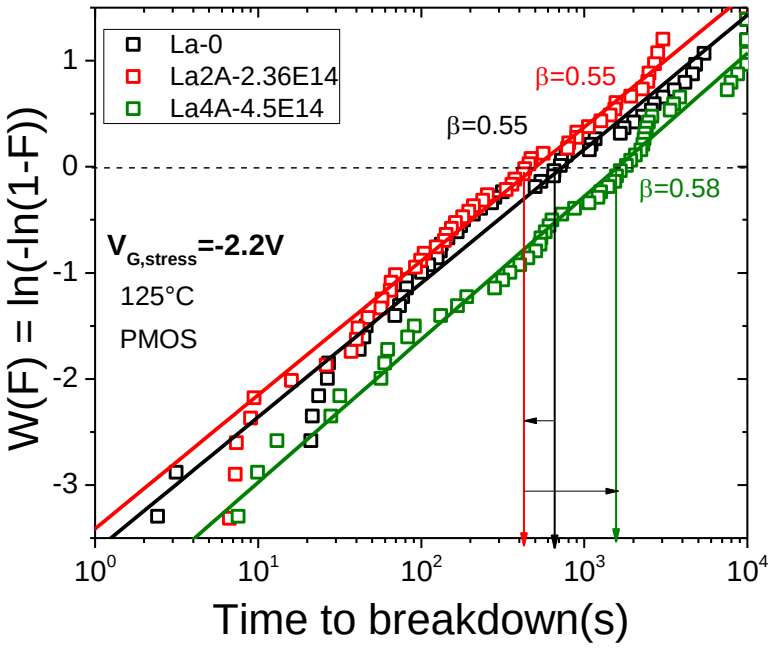


Figure 3.31: Cumulative failure distribution F(t) in a Weibull scale at gate stress voltage of -2.2V, for different La doses

From the Weibull plots for different bias conditions and La dose, T_{BD} was calculated. In Figure 3.32, T_{BD} versus stress voltage is plotted, for each La dose. In general at a same V_G, T_{BD} decreases for first La dose addition (La-2Å) but then increases for La-4Å dose addition, and this increase depends slightly on the stress bias. Power law extrapolations are presented and it can be seen that due to a higher n value for device with La-4Å dose, T_{BD} will be even higher at the nominal stress voltage values.

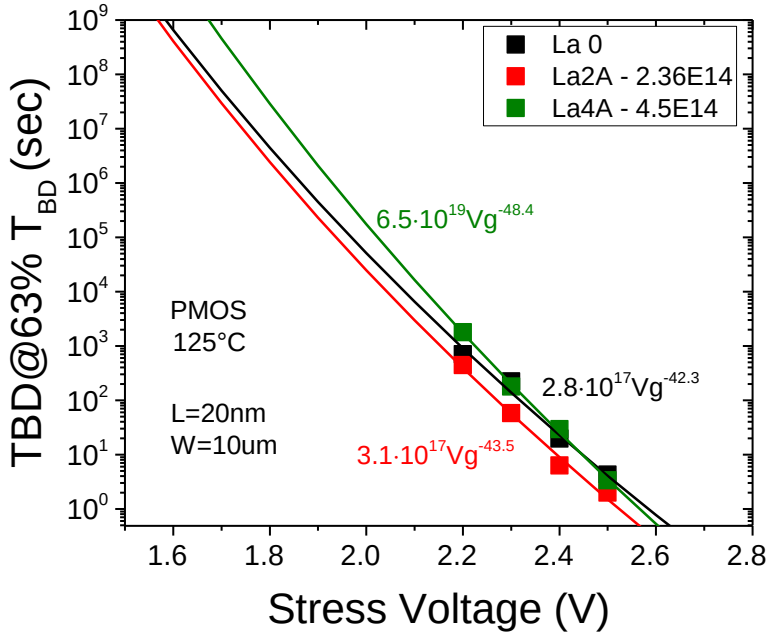


Figure 3.32: T_{BD} versus stress voltage for different La doses and extrapolation by voltage power law

Similar to the case with Al, we have considered the variation of leakage current with the effective La dose in the gate stack, in order to understand the non-monotonus effect of La dose on T_{BD}. This variation is plotted in Figure 3.33, the leakage current increases for La-2Å and then decreases for La-4Å devices, similar to T_{BD}.

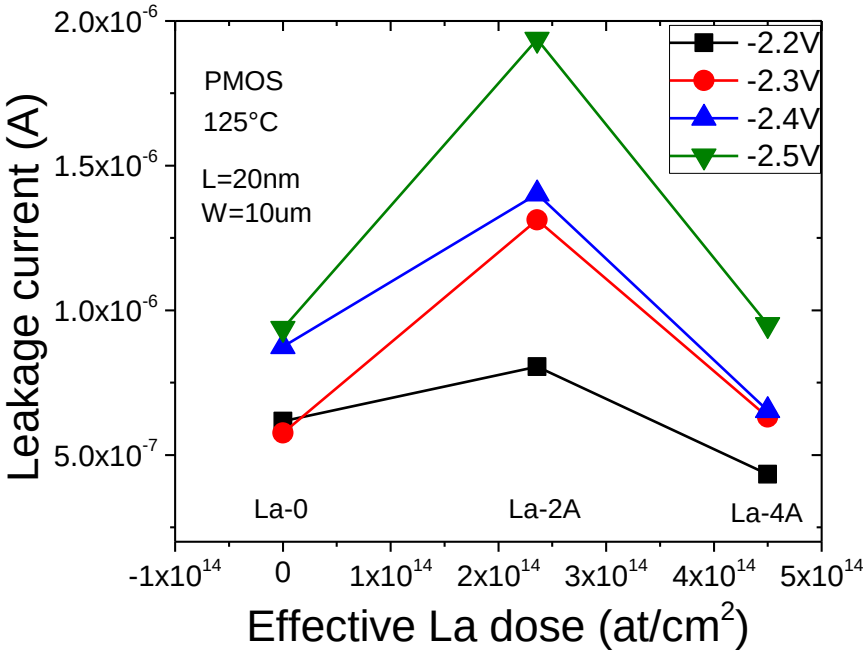


Figure 3.33: Variation of gate leakage current with La dose at different stress voltages

In Figure 3.34, T_{BD} is plotted with the leakage current (corresponding to different La doses) at different gate stress biases. T_{BD} increases with an inverse power law of the leakage current, with the power law exponent varying with the bias. From Figure 3.33 and Figure 3.34, initial addition of La into the gate stack causes an increase in leakage current, which in turn might be responsible for higher defect generation in the gate oxide and thus to the breakdown time T_{BD}. When La dose is increased to higher values, the leakage current decreases and the opposite effect takes place. As in the case of Al impact on PMOS, when the dependence of leakage current with La dose is taken into account, T_{BD} trend becomes clearer but the power law in current are less obvious with different values of exponents in Figure 3.34, compared to the case with Al where a consistent exponent value was obtained (Figure 3.29). The variation of this power law exponent might be due to an unaccounted effect that changes with the stress bias or due to a problem in extrapolation of power laws at different Ig. This can be caused by a relatively lower quality for the evaluation of breakdown distribution.

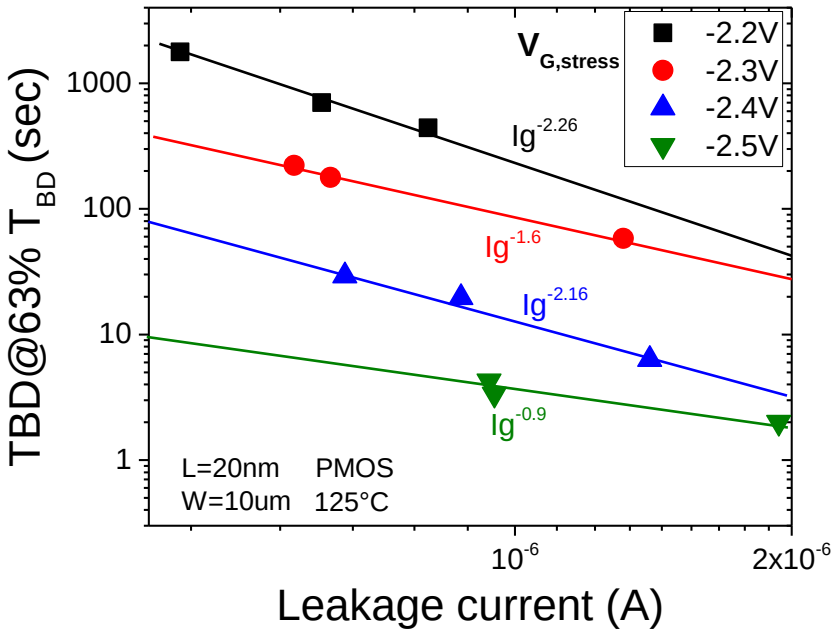


Figure 3.34: T_{BD} versus leakage current for different La doses and its fit by a current power law

Effect of La on NMOS TDDB

TDDB tests were performed on N-MOSFET devices containing the same La doses as used for PMOS TDDB and BTI studies, without any pedetal TiN and La doses of 0, $2.36 \cdot 10^{14}$, $4.5 \cdot 10^{14}$ at/cm^2. Stress voltages were selected at 2.2, 2.3, 2.4, 2.5V and 50 devices were stressed for each bias condition and La dose. Temperature during measurements was set to 125°C and device dimentions were: width = 10 um and length = 20 nm.

Figure 3.35 shows the Weibull plot for devices with La-0 dose at different gate biases, the breakdown times fit with a Weibull distribution over a large range of time. Like with PMOS, β increases as stress is increased and this might be due to a lower effective width of devices used. Figure 3.36 shows the Weibull plot of the device breakdown times for all La doses at $V_{G,stress}$ of 2.2V. Similar β are observed for different La doses. T_{BD} decreases for first La dose and then saturates for the second higher La dose.

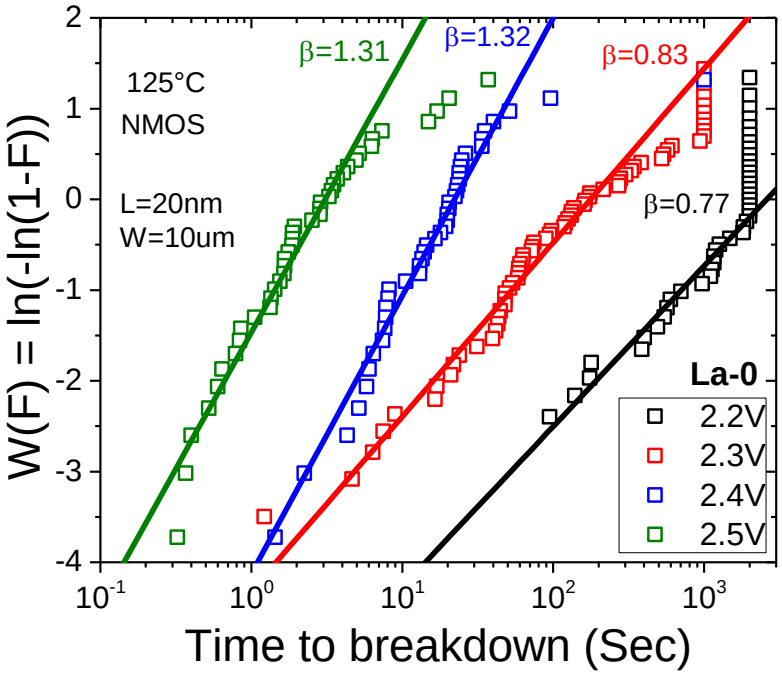


Figure 3.35: Cumulative failure distribution F(t) in a Weibull scale at different gate stress voltages, for La-0 devices

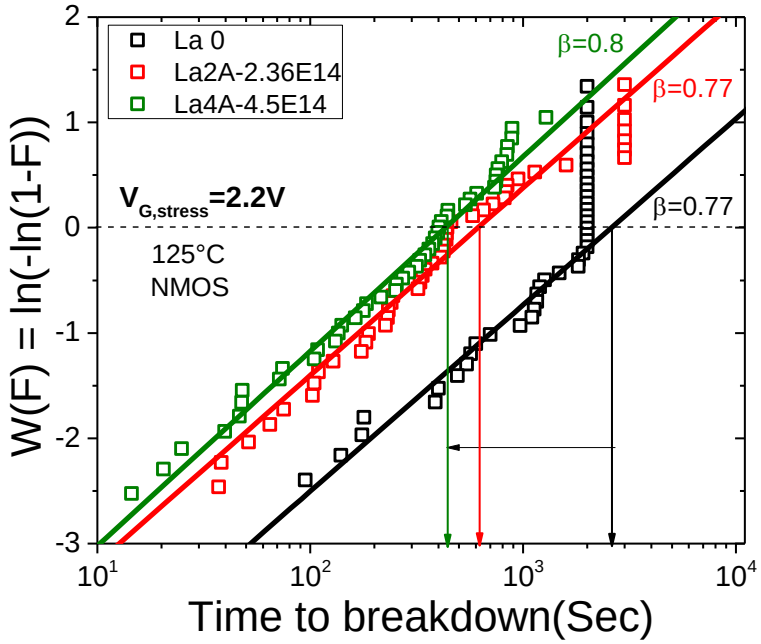


Figure 3.36: Cumulative failure distribution F(t) in a Weibull scale at gate stress voltage of 2.2V, for different La doses

From the Weibull plots for different bias conditions and La doses, T_{BD} was calculated. In Figure 3.37, T_{BD} versus stress voltage is plotted, for each La dose. In general T_{BD} decreases due to first La dose addition (La-2Å) but then saturates for La-4Å, slight variation with La dose is noticed that

depends on the stress bias. Power law extrapolations are presented but its parameters T_{BD}^{0} and n seems to change in a complicated manner with La dose.

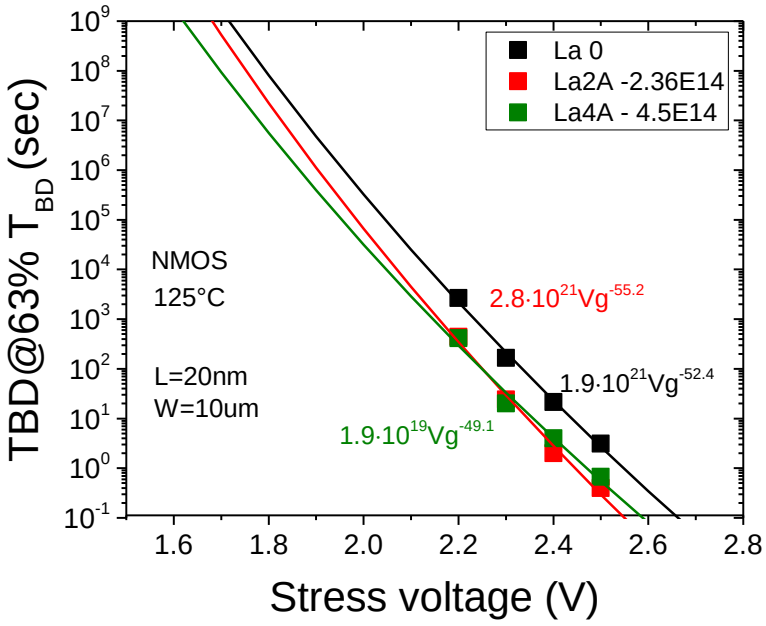


Figure 3.37: T_{BD} versus stress voltage for different La doses and extrapolation by voltage power law

As done earlier for NMOS devices with La and Al, we have analyzed the gate leakage current in order to understand T_{BD} saturation with La dose. Figure 3.38a reports on the variation of leakage current with La dose at different stress bias, it changes only slightly from La-0 to La-2Å and then decreases for La-4Å dose. Figure 3.38b shows the variation of T_{BD} with the leakage current at different stress bias. It is clear that unlike for the case of NMOS device, T_{BD} does not follow any trend with the leakage current. Leakage current remains the same from La-0 to La-2Å dose but T_{BD} decreases and when leakage current decreases from La-2Å to La-4Å, T_{BD} shows no consistent trend. This means that the oxide breakdown is not governed by the current but by some other parameter, which we will try to analyze next.

As discussed in section 1.5.2, in the case of NMOS devices under electrical stress, the electron which causes its degradation are injected from the substrate into the gate oxide. In order to understand the effect of La addition on the T_{BD} as shown in Figure 3.37, SILC measurements were performed. The devices were stressed at a bias of 2.2 V and the gate current (Ig) was measured at regular intervals of time by an Ig-V_G sweep. SILC is the normalized change in gate current Ig(t) from the initial current Ig (0), $\Delta Ig/Ig$ given in Eq. 1.23. Figure 3.39 and Figure 3.40 present the SILC spectrum ($\Delta Ig/Ig$) with gate voltage for La-0 and La-2Å respectively. As discussed earlier, the peak of a SILC spectrum can be associated with the energy level of defects in the oxide. For both cases (La-0 and La-2Å), a small peak appears at around -0.1V and a larger peak appears at around 0.5V. Studies have shown that the SILC peak at around 0.5V is usually attributed to oxide defects in the high-k, and they are considered to be of the oxygen vacancy type [118] [140][141].

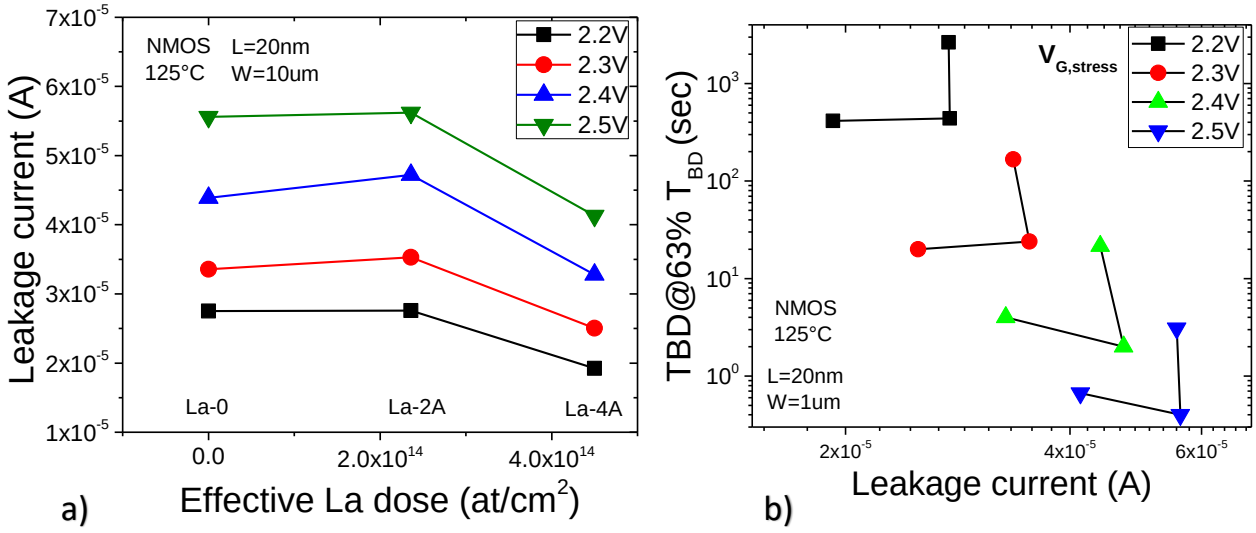


Figure 3.38: a) Variation of gate leakage current with La dose and b) T_{BD} versus leakage current for different La doses, at different stress voltages

Comparison of Figure 3.39 and Figure 3.40 shows that, the value of the SILC peaks at 0.5V for La-2Å is about twice that of La-0 device. It means that the number of stress induced defects are higher with the addition of La in the gate stack. Similar results have been obtained for La-4Å case. Figure 3.41 shows the SILC value, measured at 0.5 V, with stress time for La-0 and La-2Å devices, and it is clear that the defect generation rate is higher with La addition of 2Å. For La-4Å, this rate varies with time, but the final SILC value becomes equals to the value obtained for La-2Å. These SILC measurements evidence that even if the leakage current doesn't increase with addition of La, the rate of defect creation increases explaining a decreases in TDDB.

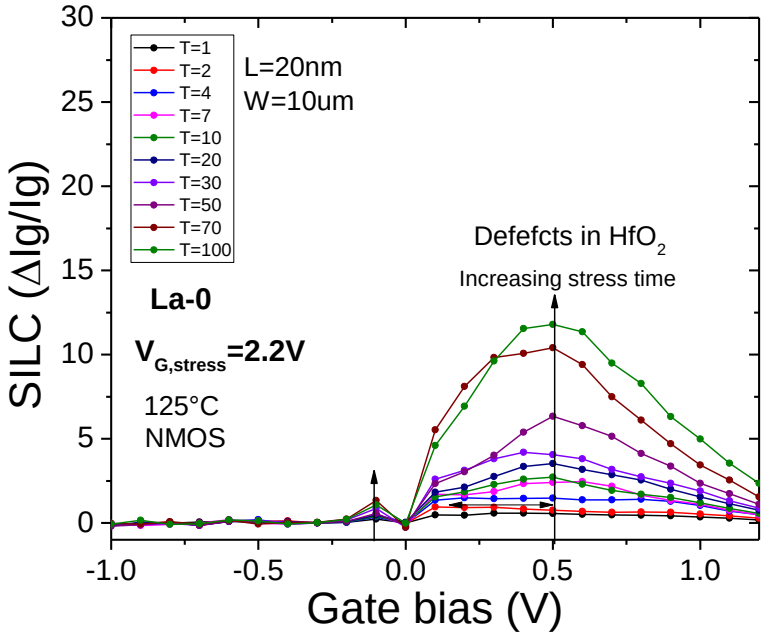


Figure 3.39: Evolution of normalized SILC (ΔIg/Ig) with time at gate stress voltage of 2.2V, for La-0 device

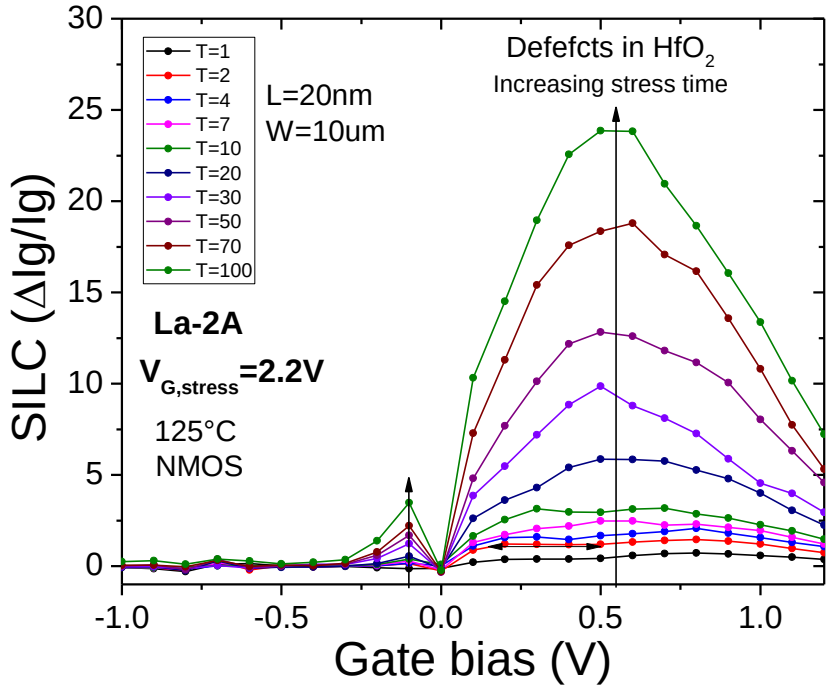


Figure 3.40: Evolution of normalized SILC (ΔIg/Ig) with time at gate stress voltage of 2.2V, for La-2Å device

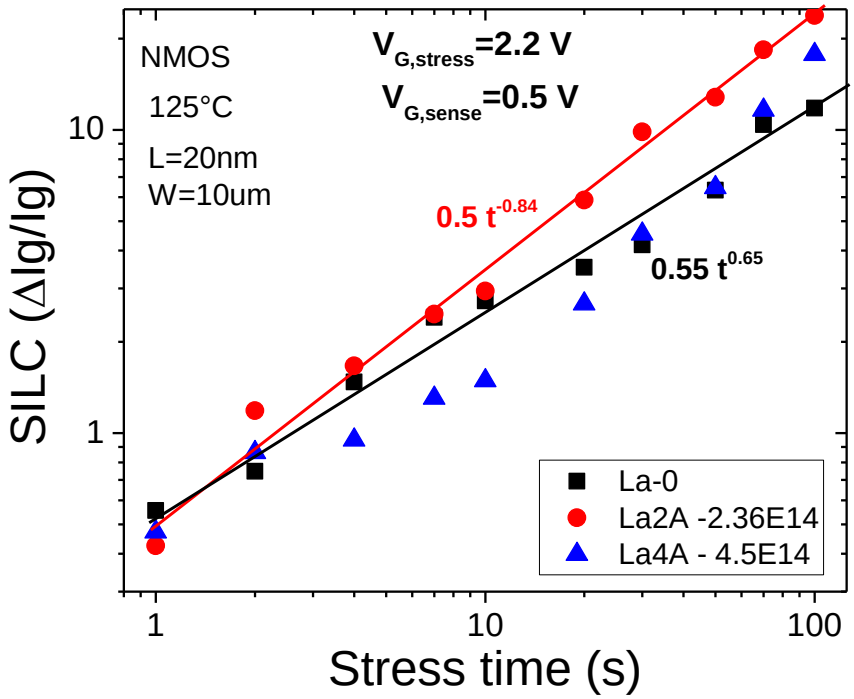


Figure 3.41: Variation of normalized SILC (ΔIg/Ig) with stress time at sense voltage of 0.5V and gate stress voltage of 2.2V, for devices with different La dose

Role of pedestal TiN on NMOS TDDB

Until this point concerning TDDB reliability, we have investigated the role of effective La dose that was introduced without using a pedestal TiN in the sacrificial gate process, i.e. La layer deposited directly on the high-k. In this section, we will investigate how the use of pedestal TiN affects the impact of La on NMOS TDDB reliaiblity. For this, TDDB measurements were carried out on devices with and without pedestal TiN, given in Table 3.1. The doses used are the same as utilized to study impact of La on PBTI. The device dimentions used for this study were much smaller then as used earlier for TDDB studies: width = 0.17 um and length = 20 nm. Thus, due to a smaller device size, the T_{BD} will be higher compared to the values reported in the previous section, for the same La dose (without a pedestal TiN).

Figure 3.42 reports the T_{BD} versus effective La dose introduced with or without using a pedestal TiN, at different stress bias. T_{BD} does not show a monotonous behaviour with La dose but rather seems to be impacted by the existence of the pedestal TiN. It first decreases (by about a decade) with the addition of La, then does not change much with La dose with a pedestal TiN. T_{BD} then increases for La dose where as deposited La was placed directly on the high-k, and then remains constant or increase slighly depending on the stress bias. This means that the interface between high-k and the sacrificial gate stack, pedestal TiN or not, is playing a role on the breakdown mechanism of the gate oxide.

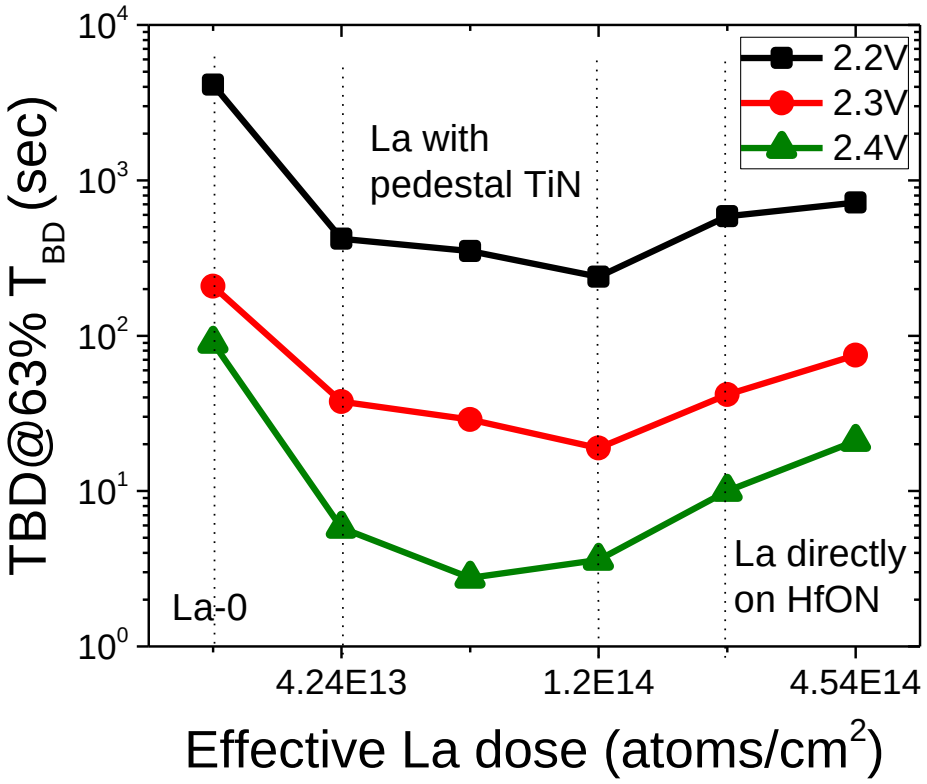


Figure 3.42: T_{BD} versus effective La dose at different stress bias, showing the role of pedestal TiN

3.3 Conclusion

The reliability effects (NBTI, PBTI and TDDB) of La and Al additives, used for threshold voltage adjustment in FDSOI gate dielectrics, have been investigated [PK-1]. BTI V_T shift follows a power law in time and increase with stress voltage. Distinction has been made between device degradation at the same oxide field and the same gate voltage, highlighting the difference in results. Comparison at the same oxide electric field shows that La causes significant enhancement of intrinsic NBTI, possibly due to defect creation by La in the interlayer SiO_2. Considering PBTI, La reduces the already low PBTI effect, likely due to passivation of oxygen vacancies in HfON by La. Comparison at the same Vg does not show a significant impact of La on NBTI and an unclear effect on PBTI, possibly due to counter balancing of La degradation effect by the reduction in oxide field caused by La dipole formation.

Effect of Al on NBTI and PBTI is quite similar at the same gate voltage and at a same oxide field due to the relatively small V_T change by Al addition. Al causes significant enhancement of PBTI, likely due to defect creation in HfON. It causes slight reduction in case of NBTI, as Al is not known to react with the interlayer and only remains and creates dipoles at the interface between IL and HfON.

Effect of La and Al addition on TDDB reliability has been studied. For all cases, the breakdown times fit the Weibull distribution over a large range of time and the average time to breakdown (T_{BD}) follows a power law in voltage. Al addition in PMOS devices increases the T_{BD}, which is caused by a decrease in the leakage current with Al dose. This current reduction is responsible for lower defect creation in the gate oxide, and thus to the higher breakdown time T_{BD}. La addition in PMOS devices decreases the T_{BD} for first La dose addition (La-2Å) but then increases it for La-4Å dose addition. This effect is also almost explained by the variation of leakage current with La dose.

La addition in NMOS devices decreases the T_{BD} due to first La dose addition (La-2Å) but then saturates for La-4Å. This trend can be explained by the SILC measurements, which have shown that defect generation is higher with La addition of 2Å but it saturates for further addition of La-4Å. Lastly, T_{BD} does not show a monotonous behaviour with La dose but rather seems to be impacted by the existence of the pedestal TiN, T_{BD} decreases with La addition and its lower with a pedestal TiN.

4. Impact of TiN process on its microstructure and electrical properties

In this chapter, the effect of TiN metal gate deposition conditions, chamber pressure and RF power, on physical and electrical properties of MOS device and TiN film itself has been studied. In section 4.1, the need for TiN microstructure modulation in the context of lowering the threshold voltage variability has been stated and then the possible RF-PVD process variations that can achieve this modulation are discussed. Section 4.2 presents the HKMG stack process and the two different process flow (or devices) that have been used in this study, to measure physical and electrical properties.

Section 4.3 presents the different physical and electrical measurements and the results obtained for variations in RF-PVD chamber pressure and RF power. In section 4.3.1 and section 4.3.2, the grain size calculated by In-plane XRD measurements and relative percentage of grain orientation calculated by out of plane XRD measurements are presented. Section 4.3.3 and 4.3.4 presents the mechanical stress and sheet resistance results respectively. In section 4.3.5, the effective workfunction (WF_{eff}) versus EOT results are presented, and the correlation between metal workfunction, dipoles and Ti/N ratio is reported.

In section 4.3.6, the effect of TiN process conditions on the wafer non-uniformity of grain size and sheet resistance are presented. Section 4.3.7 presents the impact of substrate temperature on the grain size and their relative orientations. Lastly, section 4.3.8 presents the ASTAR technique for TiN microstructure analysis and comparison of its results with XRD.

4.1 Context

As discussed in sections 1.2.2 and 1.4, introduction of a metal gate has advantages such as eliminating Fermi level pinning and phonon scattering but it also introduces the metal gate granularity (MGG) due to its polycrystalline nature. This can lead to a significant contribution in threshold voltage (V_T) variability. Moreover, metal gate microstructure can have an impact on other electrical and physical properties such as, sheet resistance [142], mechanical stress [143][144] and device effective workfunction. V_T variability due to MGG is linked to the presence of multiple grain orientations at the interface between metal gate and high-k, having different workfunctions WF_M. For TiN, used in this work, the two well-known orientations are <111> and <200> having WF_M values of 4.4 eV and 4.6 eV respectively [69].

Experimental and modelling studies have shown that there are three main technical solutions that can be implemented in order to reduce V_T variability arising from metal gate microstructure. They are presented below and shown in Figure 4.1.

- **Reducing horizontal grain size**: As stated in section 1.4, the standard deviation of the WF_M distribution is inversely proportional to the number of grains (which increases as the grain size decreases), present in a metal gate of a given area. This trend has been explained

and correlated to the metal gate microstructure (grain size and orientation) and gate area both experimentally [40] [58] [145] and by modelling [62][63][146][147].

- **Amorphous microstructure**: Studies have been conducted by adding Si or C in TiN, in order to decrease its crystallinity. Thus amorphous metal gate TiSiN [58][148] and TiCN [61] [149] have resulted in the reduction of V_T variability.

- **Unique metal orientation**: Obtaining a same and unique grain orientation on the whole gate is the most preffered choice to reduce V_T variability. Certainly, if the metal gate consists of a single crystal of a unique orientation then the WF_M will be a deterministic value instead of a statistical distribution, and the issue of V_T variability would be eliminated.

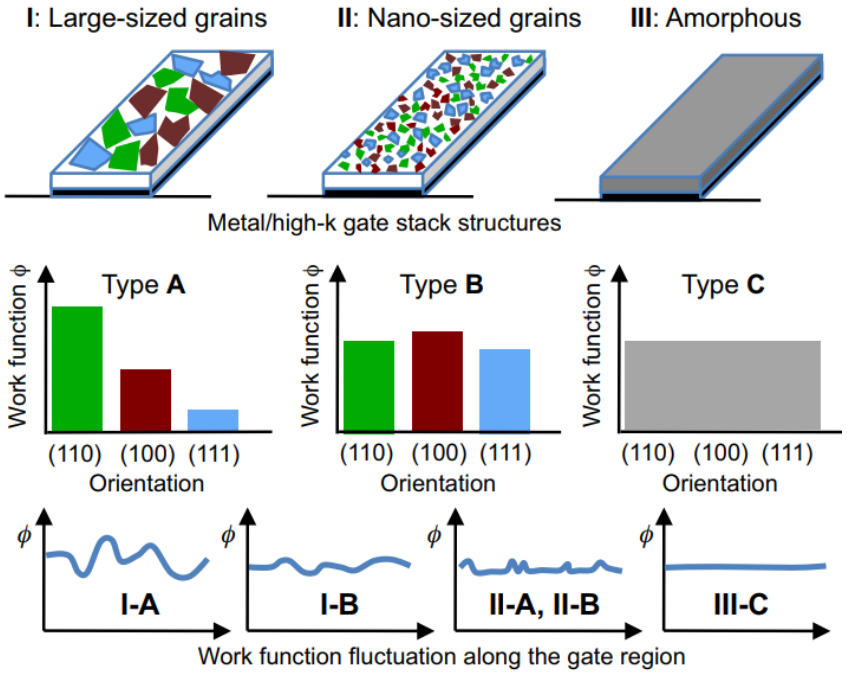


Figure 4.1: Schematic showing the work function variation, for a combination of crystal grain size (I, II, III) and orientation dependency of the work function (A, B, C) [61].

Metal thin film grain size and microstructure are determined by nucleation and growth, which can be influenced by deposition conditions during metal gate growth by sputtering, as shown in Figure 4.2 [150]. This growth process is dependent on the incident flux of sputtered metal atoms, substrate temperature, their surface energy, their mean free path, sputtering rate and the pressure of inert gases used during sputtering [150][142][151][147][152][153][154][155]. In the case of TiN metal films deposited by RF-PVD (section 1.2.2), Argon (Ar) and Nitrogen (N_2) pressure, RF power and substrate temperature can impact these parameters and hence the crystal growth process. This can lead to modulation of the preferred crystal orientation and their grain size.

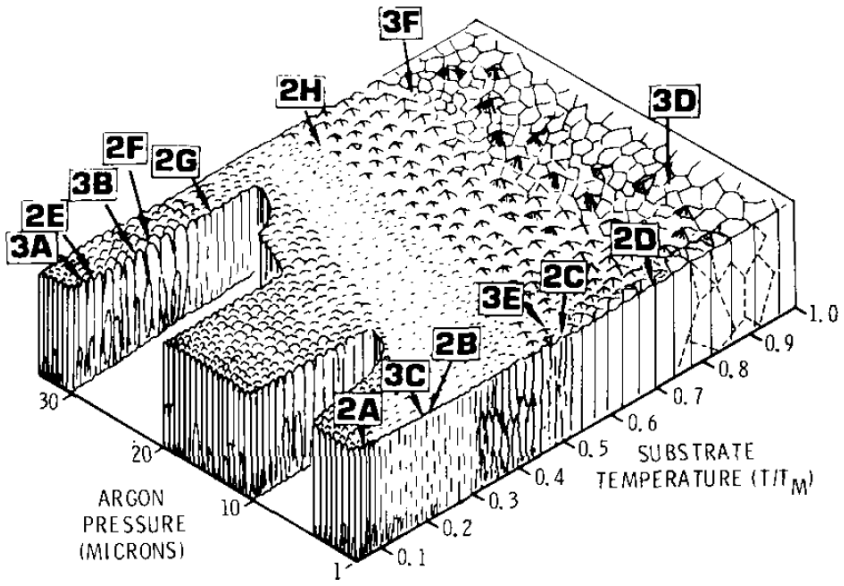


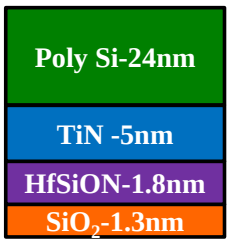
Figure 4.2: Schematic showing the dependence of sputtering gas (Ar) pressure and substrate temperature on the microstructure of metal films [150]

In this work, we have focused on identifying TiN metal gate process conditions that can reduce its grain size or provide a unique crystal orientation. Hence, the impact of the main RF-PVD parameters such as chamber pressure, RF power and substrate temperature on the preferred orientations and the size of TiN crystallites have been investigated [PK-2]. The impact of these process parameters on properties such as deposition rate, sheet resistance, film stress, and WF_{eff} of the device are also investigated [PK-2].

4.2 Device fabrication

The HKMG stack for the devices is shown in Figure 4.3 a, it consists of 1 nm thermally grown SiO_2 interfacial layer, followed by a 1.8 nm thick HfSiON High-k layer deposited by MOCVD method. Then, 5 nm TiN gate is deposited by RF-PVD with combined DC and RF generators. As presented in section 1.2.2, TiN is deposited by reactive sputtering of a pure Ti target in a nitrogen-containing ambient of Ar and N_2, and Ar and N_2 flow rates into the chamber are controlled by separate mass-flow controllers (Figure 1.13). Total chamber pressure (P) is modulated in two ways: 1) By changing Ar and N_2 flow rates, but keeping the Ar/N_2 flow ratio the same and the gate valve in open position for all deposition conditions; 2) By keeping the Ar and N_2 slow rates constant and adjusting the position of the gate valve from its mid position. For total chamber pressure below 3 mTorr, method 1) is applied and for pressure above 3 mTorr, method 2) is applied. RF power is modulated by tuning the RF generator. Chamber pressure, RF power and substrate temperature were varied over a wide range in order to investigate their impact on the preferred orientations and size of TiN crystallites. Variation of these parameters, used in this work, are shown in Figure 4.3 b.

In a normal CMOS process flow, TiN gate deposition is later followed by source drain (S-D) spike annealing at high temperature in order to activate substrate dopant diffusion. As this step can have an impact on TiN microstructure, it was included in the device gate stacks used in this work. For this, a poly-Si layer was deposited on top of TiN, followed by spike annealing at 1047 °C (Figure 4.3 a).



a)

RF power \ Ar pressure	1.6 mT	6 mT	15 mT
0 W	-10°C, 50°C	20°C	-10°C, 20°C, 50°C
600 W	50°C	20°C, -10°C	20°C
1000 W	-10°C, 20°C	50°C	-10°C, 50°C

b)

Figure 4.3: a) Schematic of the HKMG stack and b) TiN RF-PVD process splits used to study their impact on TiN microstructure

In order to study the impact of TiN deposition conditions on various properties of the TiN film itself and on the MOS devices with HKMG stack, two types of process flow were utilized.

- First is the process flow for blanket wafers that are used to study the physical properties of TiN such as its microstructure by XRD, mechanical stress and elemental composition by XRF. It contains the process steps presented above (HKMG stack + Poly-Si + S-D anneal) deposited on a p-type Si substrate, without any patterning to form devices.

- Second is the MOS capacitor flow of section 1.3.2 used to study their effective workfunction, containing the process steps presented above (HKMG stack + Poly-Si + S-D anneal) on STI SiO_2 cavities, followed by CMP patterning in order to fabricate MOS capacitors, as shown in Figure 1.21. Moreover, a bevel interlayer (section 1.3.3 and Figure 1.22) was used in order to study WF_M, charges and dipoles.

4.3 Measurements and results

Measurements were conducted on blanket wafers and MOS capacitors with HKMG gate stack and TiN process splits given in Figure 4.3. XRD measurements were done in In-plane and θ-2θ (out-of-plane) configurations, as presented in section 2.1.3, in order to determine size and preferred orientation of TiN crystallites. Sheet resistance measurements were performed by 4-probe method (section 2.1.2) and thickness measurements by X-ray-reflectometry (XRR) to calculate TiN deposition rates. Mechanical stress was calculated by measuring the wafer surface curvature before

and after the TiN deposition and applying Stoney's equation (section 2.2.4). X-ray fluorescence (XRF) measurements allowed to estimate the Ti/N ratio (section 2.2.2). WF_{eff} measurements were done on MOS capacitors with a bevel interlayer, by the method presented in section 2.1.1.

4.3.1 In-plane XRD for TiN grain size

Figure 4.4 presents a XRD spectra done in the In-plane XRD configuration. In this configuration, crystallites that are perpendicular to the sample surface are analyzed. It shows three TiN peaks that corresponds to the TiN <111>, <200> and <220> orientations, and Si peaks arrising from the Si substrate. As stated at the end of section 2.2.3, these diffraction peaks can be fitted by assuming a Gaussian or Lorentzian shape, and the FWHM is calculated. Moreover, peak broadening caused by the instrumental imperfections was removed by measuring a silicon standard powder, and then removing its FWHM value from the FWHM obtained for the TiN samples. Then, TiN grain size is calculated using the final FWHM in the Scherrer equation (Eq. 2.17).

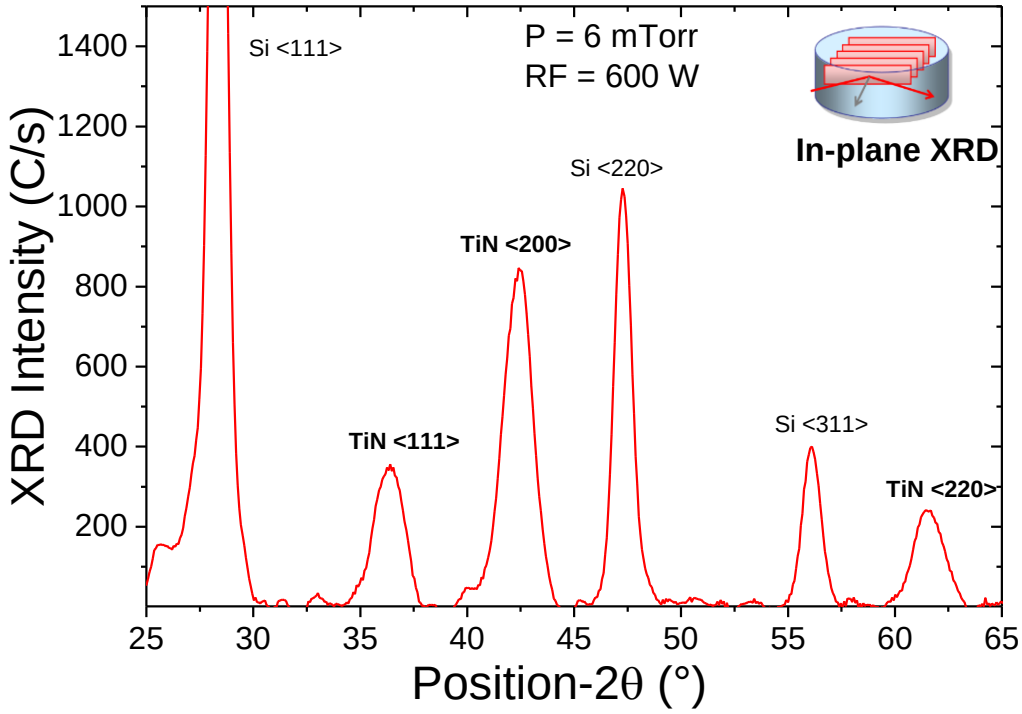


Figure 4.4: In-plane XRD spectra of the HKMG stack, showing TiN and Si peaks

Figure 4.5 shows the normalized XRD spectra for the three TiN peaks, <111>, <200> and <220> for different RF power and chamber pressure used during RF-PVD deposition of TiN. It is noticed that the FWHM decreases as the chamber pressure is decreased or RF power is inceased, except for 15 mTorr. Similarily, the FWHM was obtained for all the RF-PVD process variations in pressure and power and the grain size was calculated for each orientation. Figure 4.6 shows the grain size variation versus RF power at P = 6 mT, and it is observed that for all the crystal orientations, the grain size increases with the RF power (reverse is true for variations with

pressure). Keeping this trend in mind, the final grain size for TiN metal films was taken as an average over the grain size from all the three crystal orientations.

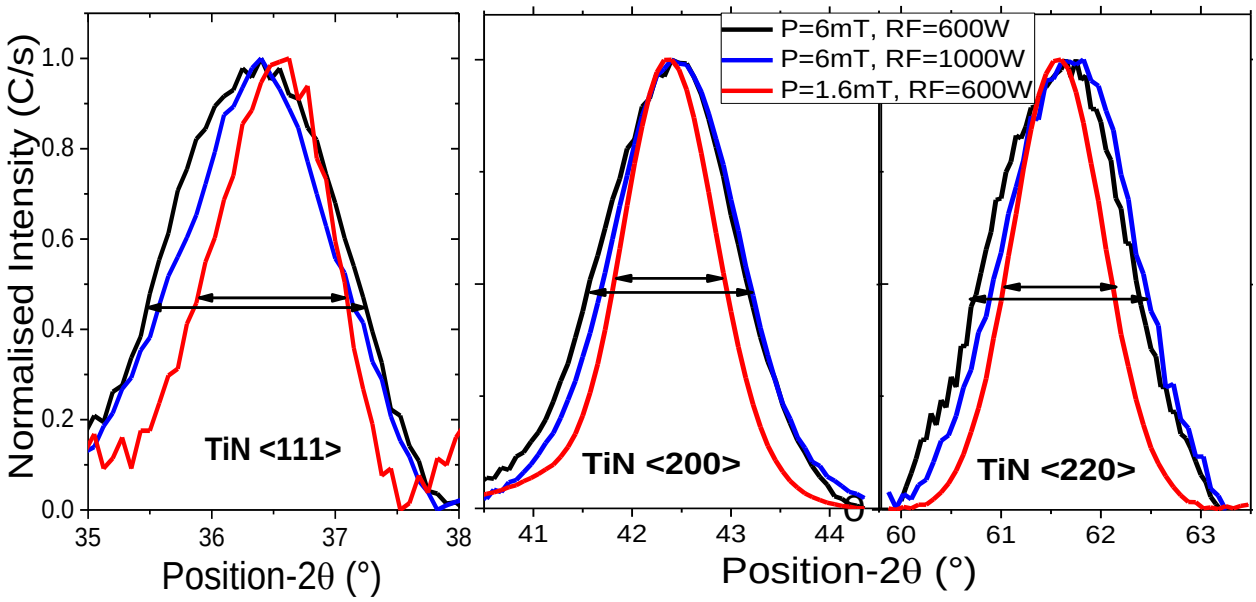


Figure 4.5: Normalized in-plane XRD spectra for TiN <111>, <200< and <220> orientations, obtained for different RF power and pressure used during TiN deposition

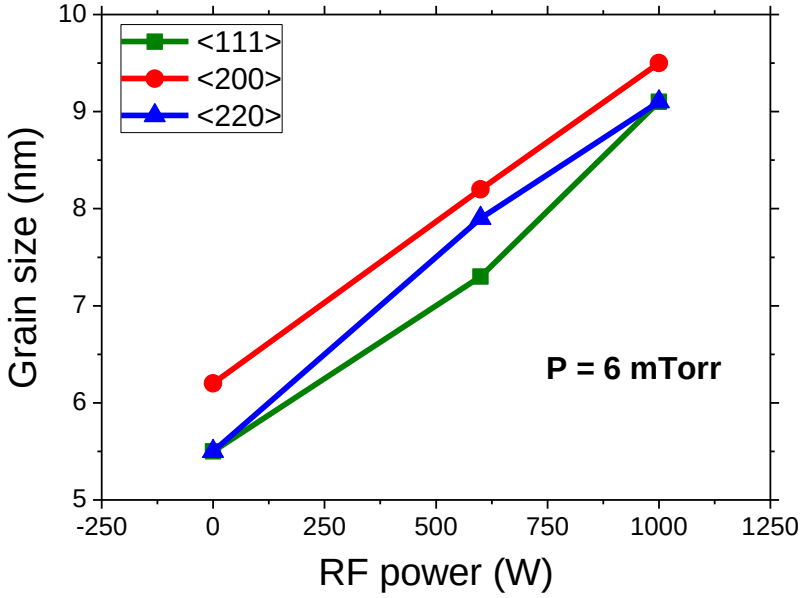


Figure 4.6: TiN grain size for different orientations as a function of the RF power

Figure 4.7 shows the variation of grain size with RF power. From Figure 4.5 and Figure 4.7, it is observed that XRD FWHM decreases and so TiN grain size increases as RF power is increased. This trend could be understood by the fact that the number and energy of plasma

electrons and ions increase with the RF power, which in turn increases the number and energy of the sputtered Ti ions [142]. The increase in sputtered Ti ions energy provides the additional momentum and energy responsible for enhancing crystal growth. Additionally, the increased number of Ar+ ions enhances crystallization due to the channeling effect [152] (detailed explanation provided in section 4.3.2). The combination of all these effects enhances crystal growth and so increases the grain size.

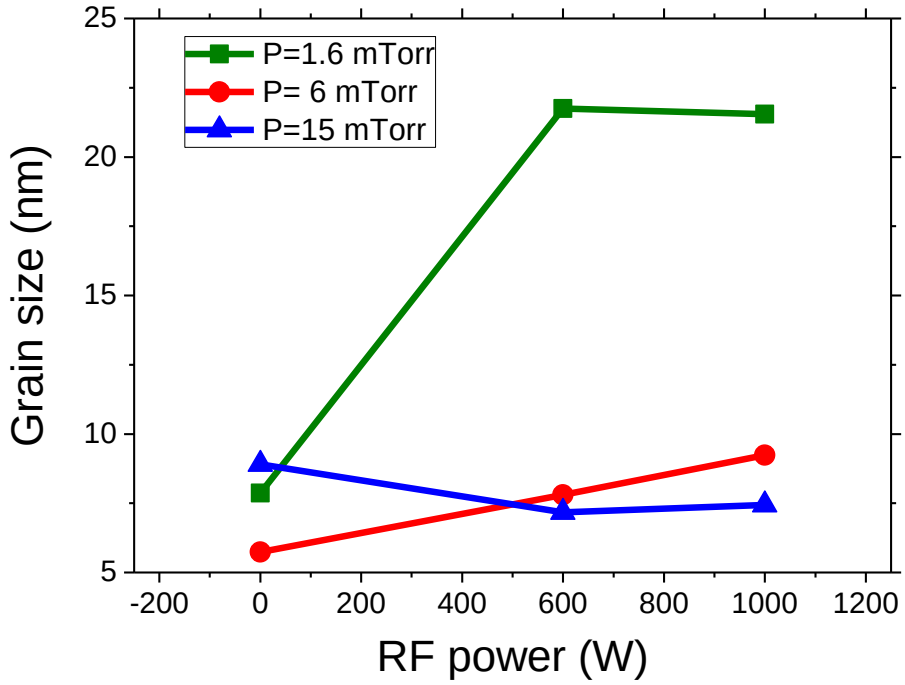


Figure 4.7: TiN grain size as a function of the RF-PVD RF power at different chamber pressure

Figure 4.8 shows the variation of grain size with chamber pressure. From Figure 4.5 and Figure 4.8, it is observed that XRD FWHM decreases and so TiN grain size increases as pressure is decreased, except for RF = 0W. This trend could be understood by the combined effect of increase in the mean free path of plasma electrons, creating more Ar+ ions for sputtering, and increase in the mean free path of the sputtered Ti species as the pressure decreases [142]. Following the same explanations given above for variations with RF power, the increase in sputtered Ti energy and channeling effect caused by the increased number of Ar+ ions are responsible for enhancing crystal growth with decrease in the chamber pressure.

Figure 4.8and Figure 4.9 shows that grai size as low as 5.2 nm can be obtained. As discussed earlier, this lowering of TiN grain size can have an impact on the VT variability. Moreover, wet or dry etching processes used during gate patterning steps of CMOS, can get impacted by the grain size of the TiN layer. In general, for wet etching of TiN, etching rate decreases as the film crystallinity is increased. Lastly, apart from TiN layers are also used as diffusion barrier over the source drain regions, the grain size can be important in this application,

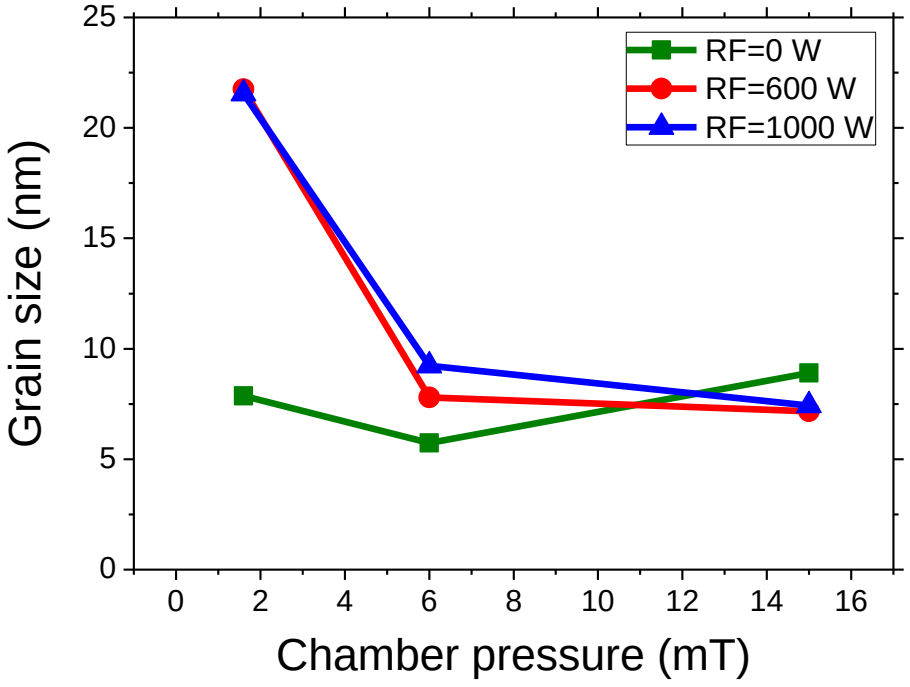


Figure 4.8: TiN grain size as a function of the RF-PVD chamber pressure at different RF powers

Increase in the number and energy of sputtered Ti atoms, by increasing RF power or decreasing chamber pressure, will increase the deposition rate as well, as reported in Figure 4.9.

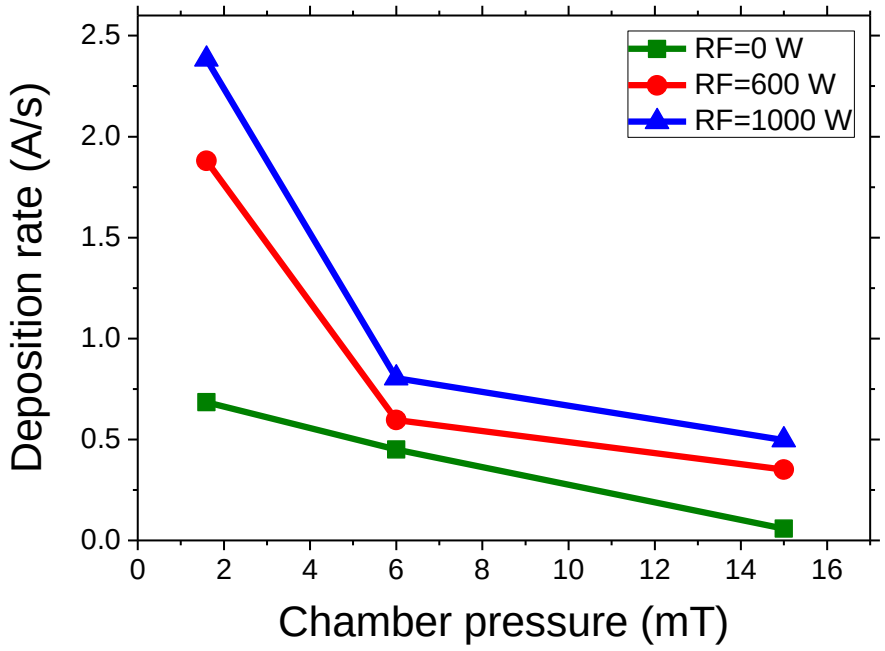


Figure 4.9: TiN deposition rate as a function of the deposition pressure at different RF powers

4.3.2 θ-2θ XRD for TiN grain orientation

Figure 4.10 presents an XRD spectra done in the θ-2θ XRD configuration. It shows the two most intense XRD peaks that corresponds to the TiN <111> and <200> configurations. Note that other orientations are not observed. In this configuration, crystallites that are parallel to the surface of the sample are analyzed. These crystal orientations are the ones that will determine the work function of the TiN metal gate and so might impact the WF_{eff} and/or the V_T variability. As stated at the end of section 2.2.3, these diffraction peaks can be fitted by assuming a Gaussian or Lorentzian shape, and the relative amounts of different crystallographic orientations present in TiN films can be obtained by their area ratio. The ratio of area under the peaks for different orientations, corrected by the relative intensities due to this structural factor (F_{hkl}) has been taking into account to calculate the relative amount of different grains. Note that as only TiN <111> and <200> orientations were observed, their sum was forced to 100 %, to calculate the relative amounts of these grain orientations.

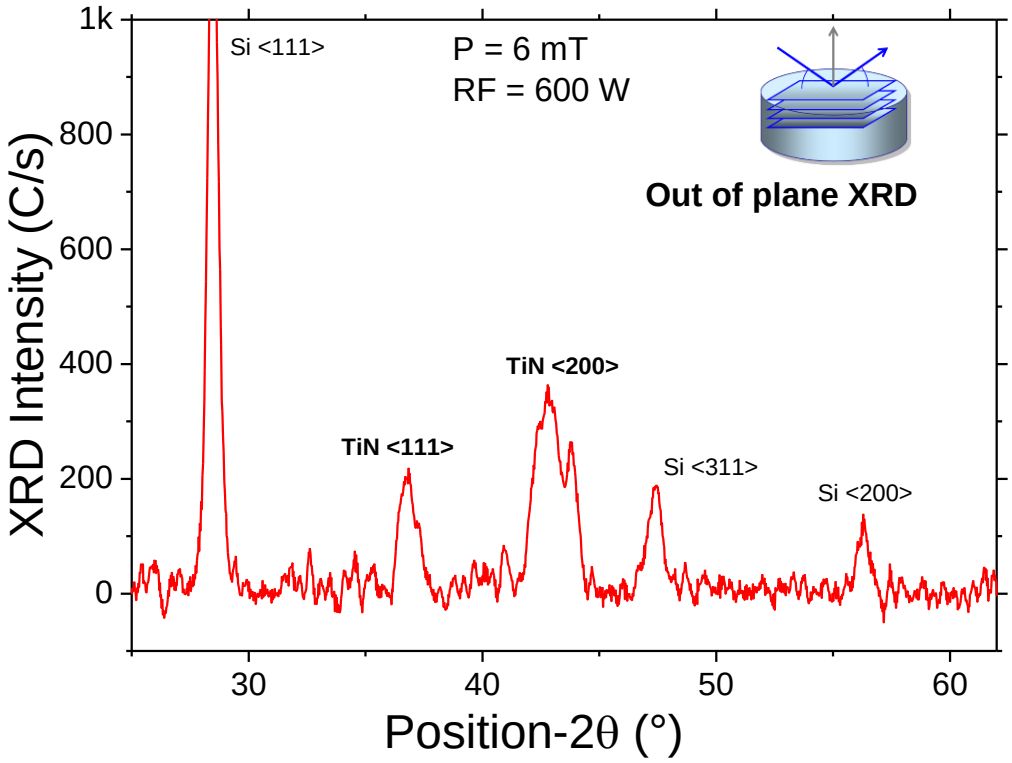


Figure 4.10: Out of plane XRD spectra of the HKMG stack, showing TiN <111> and <200> and Si peaks

Figure 4.11 shows the XRD spectra for the TiN peaks <111> and <200> for different RF powers and chamber pressures used during RF-PVD deposition of TiN. It is observed that as the RF power increases or pressure decreases, the peak XRD intensity and the area under the peak of <200> orientation increases and <111> decreases.

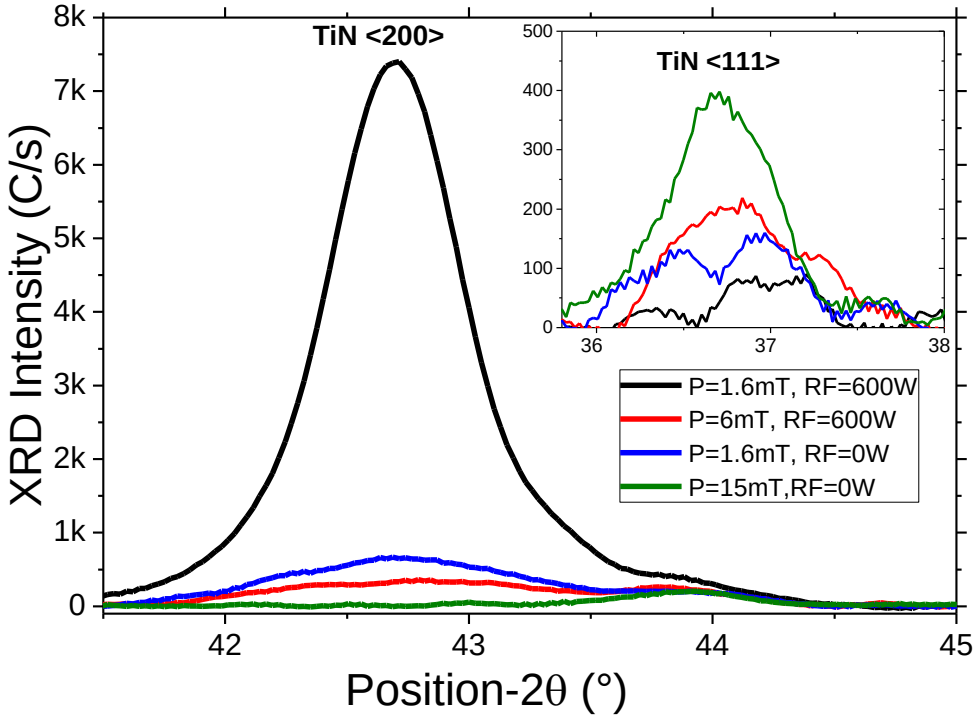


Figure 4.11: Out of plane XRD spectra for different RF power and pressure used during TiN deposition

Figure 4.12 and Figure 4.13 shows the calculated percentage variation of TiN <200> crystallites with chamber pressure and RF power. As the RF power increases and pressure decreases, the energy of sputtered Ti atoms and number of Ar+ ions increase. The trends in Figure 4.12 and Figure 4.13 could be understood by a combination of the following three effects:

- <200> orientation is the closest packed in the TiN crystal and so has a lower surface energy than <111>. This is due to a higher coordination number (bonds) in <200> surface and hence it will have a lower surface tension compared to <111>. The increased energy of Ti atoms thus favors the preferred orientation <200>, having the lowest surface energies [153].

- The channeling effect [152]: If some Ti+ or Ar+ ions reach the substrate, they may travel inside the already formed TiN film to a distance that depend on the TiN region's crystal orientation. The RF-PVD equipment, described in Figure 1.13, is designed in a way to minimize Ar+ ions reaching the substrate, thus the channeling effect might only be cause by Ti+ ions. In an FCC crystal (such as TiN), ion penetration in <200> crystals is more than in <111>, and erosion of crystals by ions has the reverse order. Thus, ions penetrating the <200> crystals penetrate deeper and so dissipate energy over a large area and hence causes less damage and atomic erosion. Thus, an increased number of Ti+ ions will re-sputter the <111> crystals more than the <200> in the TiN film.

- Focusing of impacts: During Ti+ ions impact on the depositing TiN films, the irradiation affected area is also enlarged by the transfer of impact along the most closely packed atomic planes (<200>). This enlargement dissipates the incoming Ti+ ions energy over a large are and hence reduced the damage. As <200> crystals are packed closely, compared to <111>, this might further increase the % of <200> among the TiN film [152].

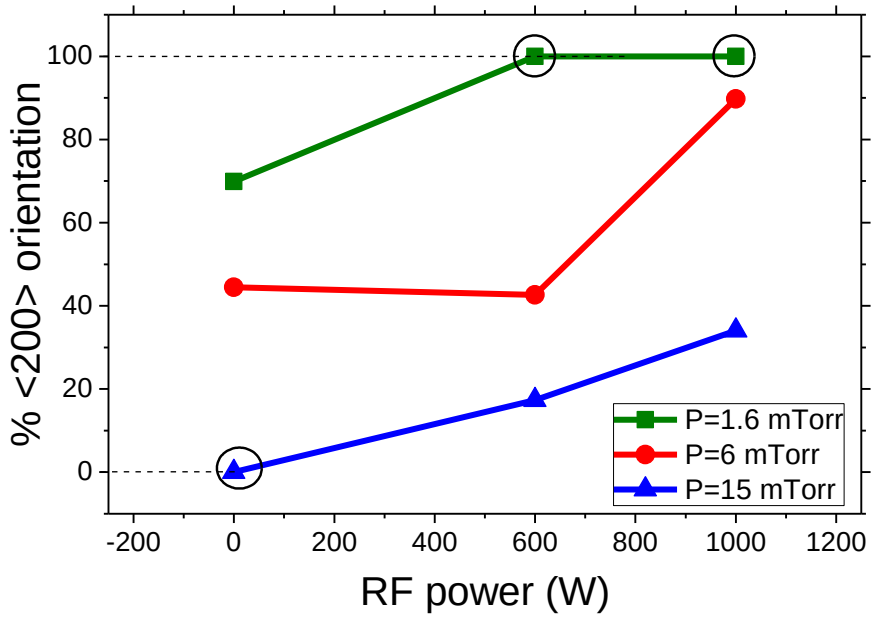


Figure 4.12: Percentage of <200> TiN crystallites orientation as a function of the RF-PVD RF power at different chamber pressure (%<111> = 100 - %<200>)

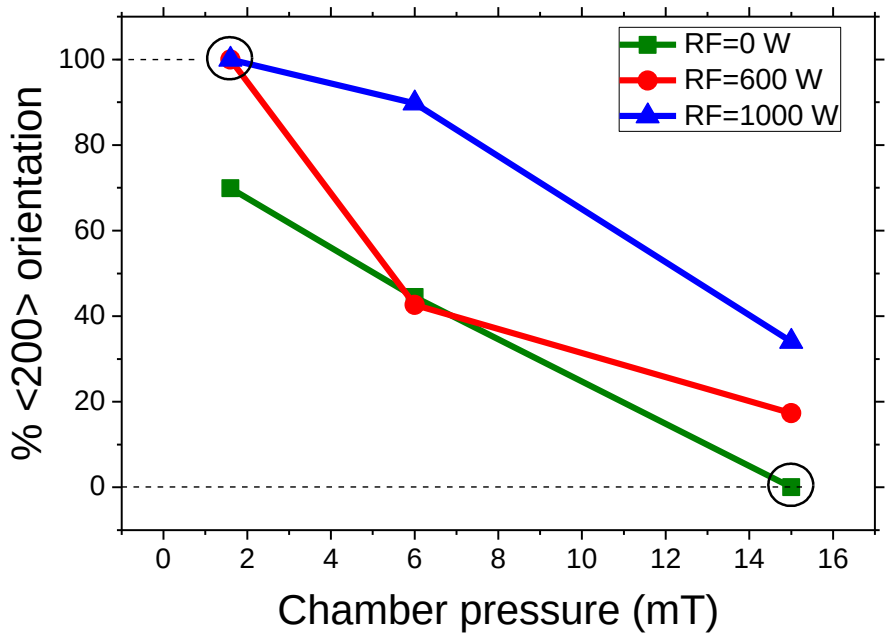


Figure 4.13: Percentage of <200> TiN crystallites orientation as a function of the RF-PVD chamber pressure at different RF power (%<111> = 100 - %<200>)

The most interesting conditions are the ones where either the <111> or the <200> orientation exists alone (black circles in Figure 4.12 and Figure 4.13), demonstrating here that a unique grain orientation can be achieved for TiN films by the modulation of RF-PVD chamber pressure and RF power [PK-2] [PK-6]. These 3 conditions (P = 1, RF = 600 W; P = 1, RF = 1000 W; P = 15, RF = 0 W) are very important as they satisfy one of the conditions that might significantly reduce or even eliminate MOSFET V_T variability caused by TiN metal gate microstructure or MGG.

It is important to highlight the relationship between crystal orientations obtained from the in-plane and out-of-plane configurations. This relationship depends on the texture of grains present in the TiN film. For the TiN process used in our studies, the grains are present in a fiber texture, meaning that the grains obtained from in-plane measurements are randomly oriented around one or few specific crystal planes parallel to the sample surface. The out-of-plane measurements gives the fiber texture axis for the in-plane measurements. As we only see <111> and <200> orientations in the out-of-plane measurements, it means that the in-plane grains are randomly oriented around these two directions. The general relationship, for a fiber texture, between out of plane and in-plane obtained orientations is given in Table 4.1.

Out of plane orientation	In-plane grain orientation
<111>	<220>
<200>	<200>, <220>
<220>	<200>, <220>, <111>

Table 4.1: Relationship between different orientations obtained from out-of-plane and in-plane XRD measurements

In our results, we obtained a <111> orientation signal from in-plane measurements which is supposed to be only from <220> out-of-plane orientation (from Table 4.1), but this orientation <220> is not observed in Figure 4.12. This can be due to two reasons: 1) either there is another crystal orientation parallel to the surface, which is not <220>, but is not observed in the out of plane XRD spectra (for instance <112>) due to h,k,l indices not being all odd or all even; 2) in-plane <111> orientation is from out of plane <220>, but it is not observed in the out-of-plane XRD spectra due to it being buried in the background signal.

Table 4.1 shows that the measured in-plane orientations does not correspond directly to the out of plane orientation (or effective orientation). The relationship is complex and a specific in-plane orientation signal is a mixture of signals from multiple real orientations. So the horizontal grain size should be calculated as the average over all the orientations, which justifies taking average of the grain size from all these crystal orientations (as done in section 4.3.1 for in-plane measurements) and we must accept to not finding any other orientation from in-plane analysis.

4.3.3 Mechanical stress

Stress added by TiN film deposition on Si substrates was calculated by the Stoney's equation (section 2.2.4 and Eq. 2.18 and 2.19). For these measurements, the radius of curvature of the wafers were measured before and after TiN deposition and put into the Stoney's equation. Figure 4.14 shows the variation of mechanical stress with chamber pressure and RF power, it is clear that the stress is always compressive and this stress decreases as the pressure is increased or RF power is decreased. This stress behavior with pressure and RF power is quite the opposite to the trend observed for the grain size in Figure 4.7 and Figure 4.8, percentage of <200> orientation in Figure 4.12 and Figure 4.13, and deposition rate in Figure 4.9. This relationship can be understood as follows: a pressure decrease or RF power increase results in bigger crystals of preferred orientation and thus reduces the voids in the film [153]. Moreover, the increase in number and energy of sputtered Ti atoms could cause the atomic peening effect [143], where the film is bombarded by these atomic species and hence forms a dense microstructure. Both these effects enhances the compressive stress.

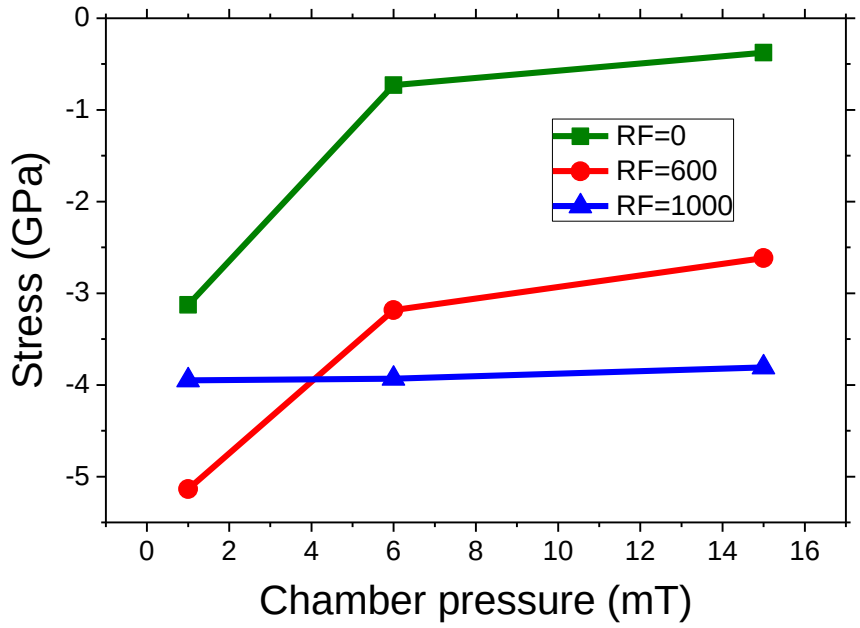


Figure 4.14: Mechanical stress as a function of chamber pressure at different RF power

Lowering mechanical stress of deposited TiN films is important for correct CMOS gate patterning, its adhesion to high-k layer below and poly-Si layer on the top, and in order to reduce defects and voids. Significant stress variation is observed for such range of RF power and pressure, and significantly low stress values can be achieved for low RF and high pressure conditions.

4.3.4 Resistivity of TiN films

In order to determine the impact of TiN deposition conditions (or microstructure) on its electrical conductance, sheet resistance measurements were done by the four probe method as presented in section 2.1.2. Figure 4.15 shows the variation of resistivity (sheet resistance × thickness) with the grain size obtained from in-plane XRD measurements of the TiN films, for different pressure and RF power. It is observed that the resistivity changes only slightly at larger grain size and then follows an inverse power law as the grain size is decreased. As the grain size is reduced, the grain boundaries increase and thus charge carrier scattering is increased which causes increase in the resistivity of the TiN film. A wide range of resistivity values can be obtained (150-550 μΩ.cm). This will affect electrical biasing and gate access resistance of the device, and so the performance of the CMOS.

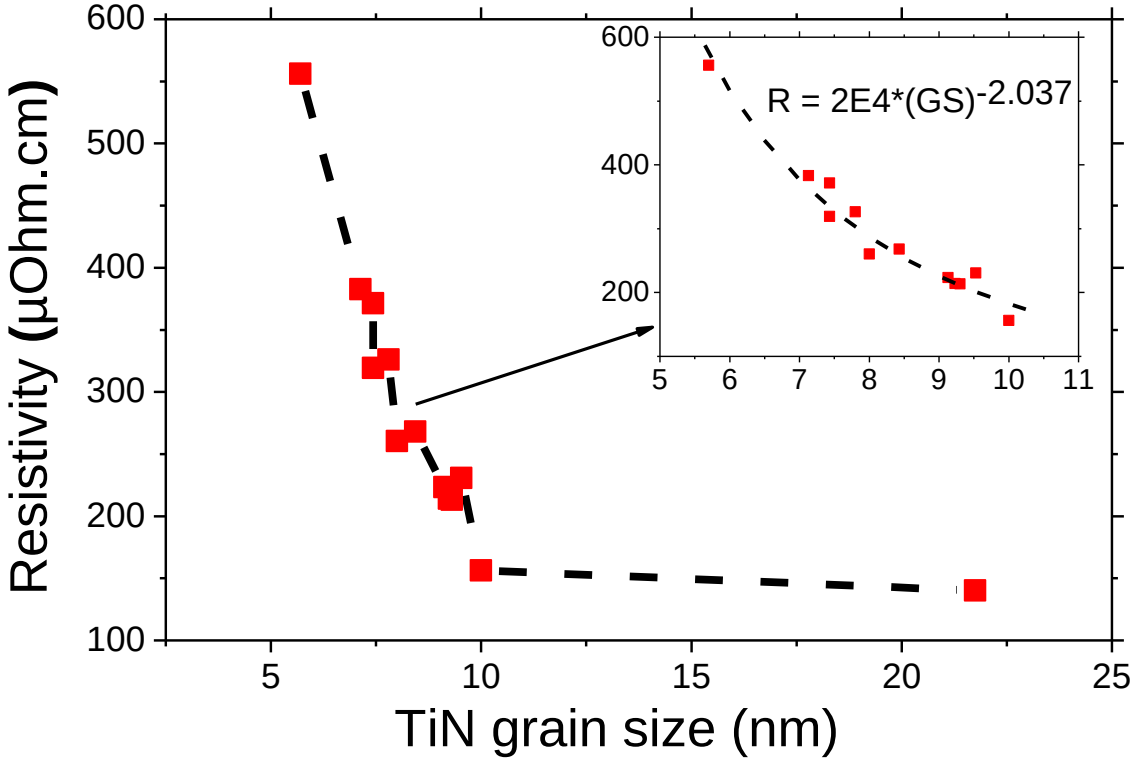


Figure 4.15: TiN film resistivity as a function of its grain size

4.3.5 Effective workfunction measurements

In order to determine the impact of TiN deposition conditions on MOS electrical parameters, EOT and Flatband voltage (V_{FB}) were extracted by fitting experimental C-V measurements with Poisson Schrödinger quantum simulations on MOS capacitor, fabricated with the process flow presented in section 1.3.2 and 4.2. The HKMG stack is as given in Figure 4.3a and consists of TiN/HfSiON/beveled-SiO_2. CV measurements were conducted on devices present on different dies along the wafer radius, thus modulating the EOT. From Eq. 1.11, V_{FB} is extracted and given as follows:

$$qV_{FB} = WF_M + \delta - Q_{Si/SiO2}(EOT/\varepsilon_{ox}) - WF_{Si} \quad ; \quad WF_{Si} = q\Phi_{Si} \tag{4.1}$$

where δ denotes the sum of dipoles at the SiO_2/high-k interface, $Q_{Si/SiO2}$ are the interface charges at this Si/SiO_2 and WF_{Si} is the silicon work function. V_{FB} is extracted from C-V measurements and the WF_{eff} is calculated from Eq. 2.9 and given again in the equation below:

$$WF_{eff} = qV_{FB} + WF_{Si} \tag{1.12}$$

As stated in section 2.11 and from Eq. 4.1 and 4.2, the extrapolation of WF_{eff} to EOT = 0 allows therefore the assessment of only $WF_M + \delta$ [17], which can be modulated by a combined effect of:

- Change in WF_M which is expected to depend on the relative quantity of <200> (WF_M = 4.6) or <111> (WF_M = 4.4).

- Change in WF_M or δ by other effects, such as TiN compositional change or dipole formation at the interfaces.

Figure 4.16 shows the WF_{eff} versus EOT plots, for devices with variations in TiN RF power and chamber pressure during its deposition, it also reports on the $WF_M + \delta$ (WF_{eff} at EOT = 0). From Figure 4.12 and Figure 4.13, percentage of <200> crystal orientation decreases and the one of <111> increases as the pressure is increased or RF power is decreased. This modulation of relative crystal orientation quantity should decrease the WF_M. From Figure 4.16 an opposite trend is observed for $WF_M + \delta$, where an increase of pressure from 1.6 mT to 15 mT or a decrease of RF power from 600W to 0W increases the $WF_m + \delta$. Moreover, Figure 4.12 shows that (for P = 6mTorr) from RF = 0W to 600 W the % of <200> remains constant (and so should WF_M) and then increase for RF =1000W but in Figure 4.17a the WF_{eff} vs EOT at P = 6 mTorr and different RF power does not follow this trend. On the other hand, at P = 15mT the % of <200> increases with RF power and so does the $WF_M + \delta$ (in Figure 4.17b). Thus it is clear that the $WF_M + \delta$ does not follow the variations of WF_M (or relative % of grains), except for some conditions (Figure 4.17 b). These trends evidences that another physical mechanism is at work here, other than relative amount of grain orientation of <200> or <111>, and this physical mechanism impacts either the WF_M or δ.

In order to understand the cause of this unexpected WF_{eff} shift, Ti/N ratio inside the TiN gate was measured, which can be modulated by pressure and RF power. Ti/N was measured by the XRF technique (section 2.2.2) and its variation with RF power and chamber pressure is shown in Figure 4.18, it mostly decreases with chamber pressure except for RF = 0W.

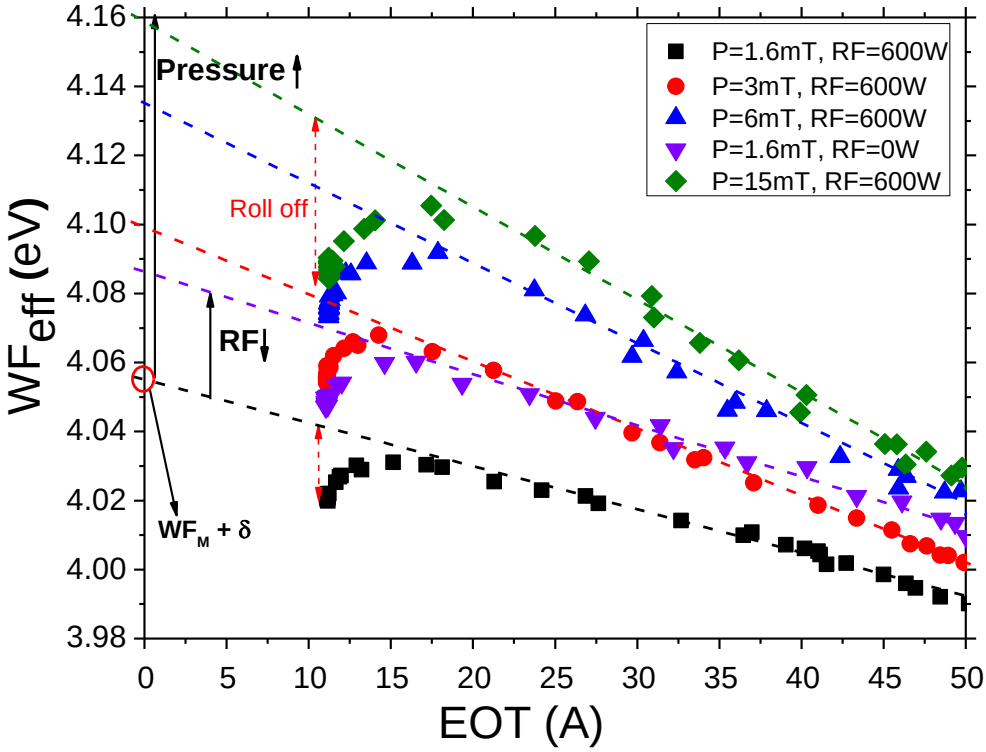


Figure 4.16: WF_{eff} vs EOT plots for different RF power and chamber pressure used during TiN deposition

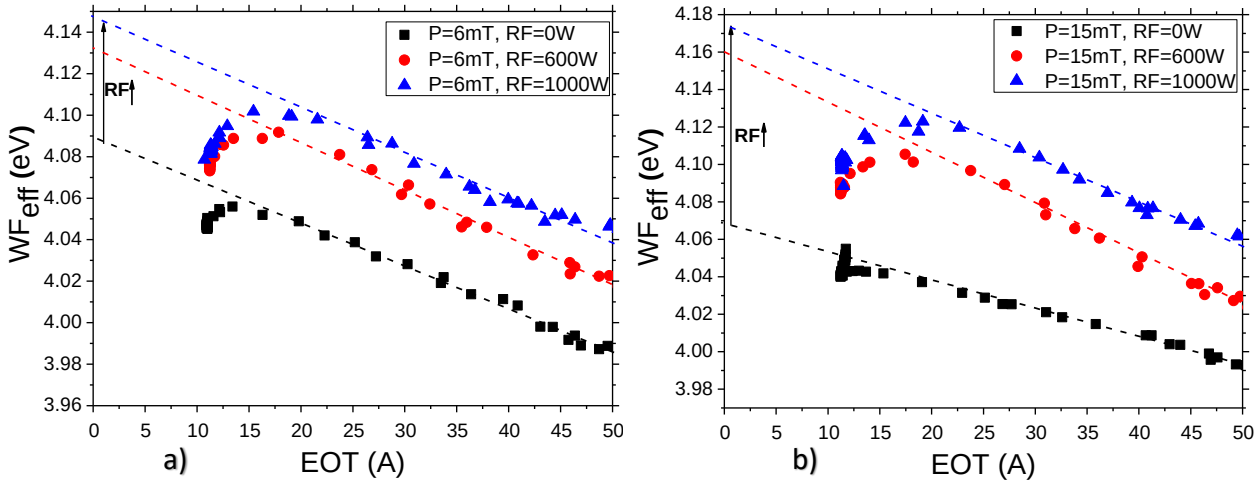


Figure 4.17: WF_{eff} vs EOT plots for different RF power at chamber pressure a) 6 mT and b) 15mT

Figure 4.19 shows WF_{eff} at EOT=0 (WF_M+ δ) versus the Ti/N ratio calculated for different RF-PVD process conditions, and indeed a good correlation between the two is observed. This correlation does not seem to depend on the WF_M change with Ti/N because the observed dependence with increasing Ti/N ratio is not really linear, and on the contrary it seems to saturate at a higher value than expected for pure Ti (0.3V below TiN, WF_M for Ti = 4.33 eV). This

dependence can be explained by the suppression of dipole δ at SiO_2/high-k interface by Ti rich gates. Indeed, Ti rich metal gate can de-oxidize the SiO_2/high-k interface and therefore, would have a lower $WF_M+\delta$ level. In the case of N rich metal gate the dipole at SiO_2/high-k interface may increase δ [37][156], and so $WF_M+\delta$. A sign of such a phenomenon can be seen in Figure 4.16 where the roll off value at thin EOT is increasing as the WF_{eff} increases. Indeed, roll off at thin EOT has been related to an evolution of dipole at thin EOT [17]. Thus, a change in the deposition conditions modulates the Ti/N ratio in the TiN gate, which is responsible for the modulation of dipole δ at the SiO_2/High-k interface and of the WF_{eff} at EOT =0.

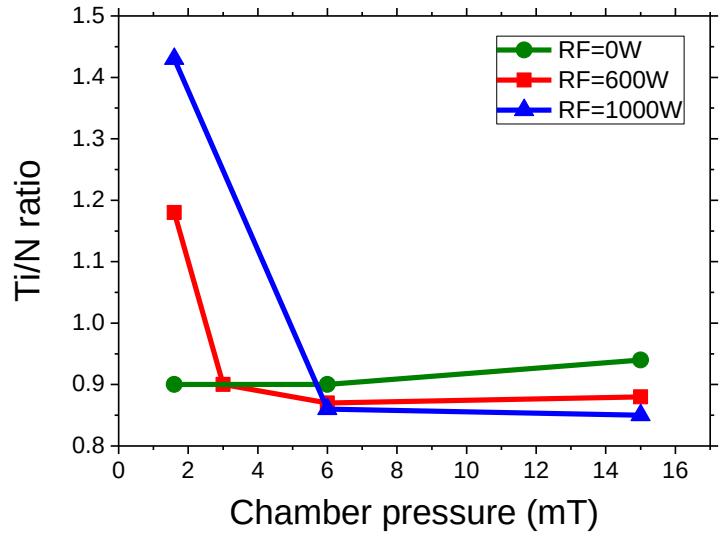


Figure 4.18: Ti/N ratio in the TiN gate as a function of its chamber pressure at different RF power

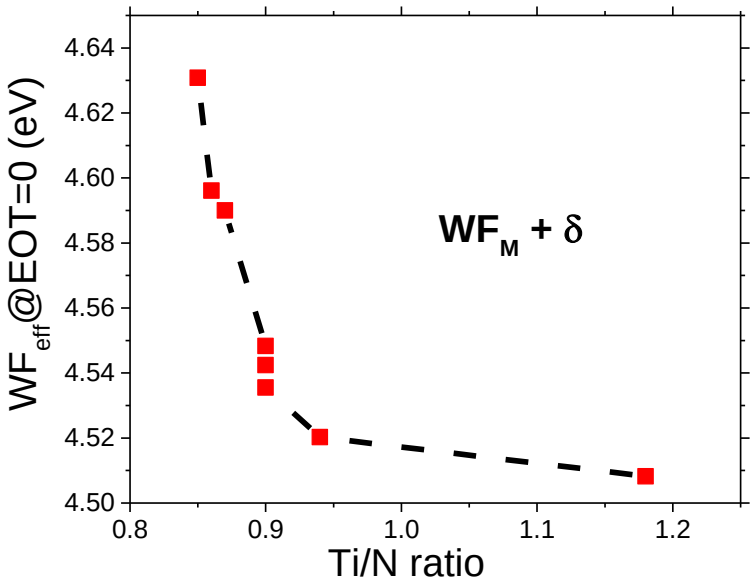


Figure 4.19: Variation of WF_{eff} at EOT = 0 (WFM + δ) with the Ti/N ratio

It is important to highlight that even though the grain orientation effect on WF_M is not observed here, due to a stronger effect of Ti/N ratio on SiO_2/high-k dipole δ, this does not imply that WF_M or V_T variability does not change with grain orientation. Further analysis by X-ray photoelectron spectroscopy (XPS) is required, by which it is possible to measure WF_M of TiN independently of δ.

4.3.6 Grain size variation with wafer radius

In order to check the non-uniformity of the grain size over the wafer and its dependence with the RF power and chamber pressure, in-plane XRD was performed at the center and the edge of a 300 mm diameter wafer. The ratio of the grain size at the edges and the center (R_{GS}), representing its non-uniformity, versus chamber pressure is shown in Figure 4.20 for different RF power. R_{GS} increases with chamber pressure and decreases with the RF power, and reaches a maximum value of 1.27.

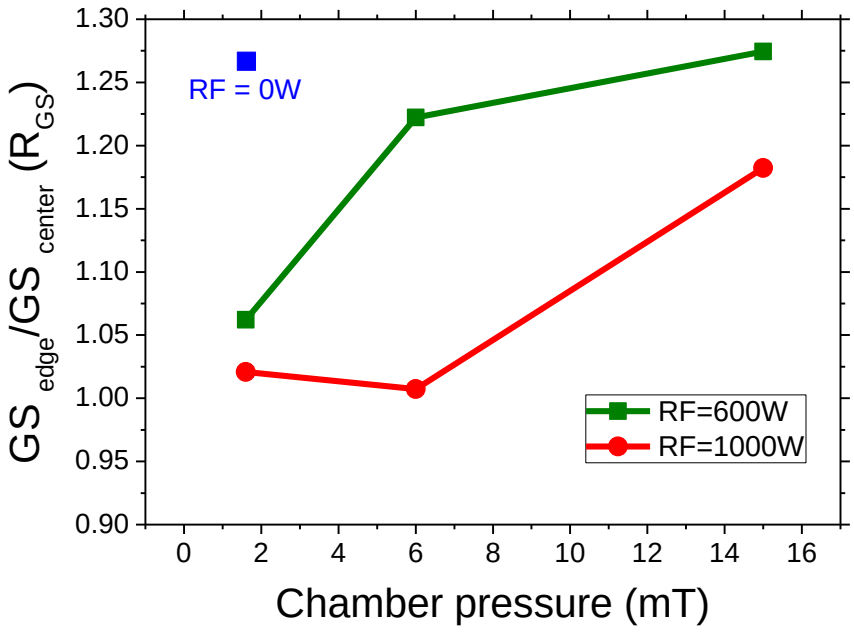


Figure 4.20: Ratio of TiN grain size at the center and edge (R_{GS}), as a function of the RF-PVD chamber pressure, at different RF power

In order to determine the impact of TiN deposition conditions on the non-uniformity of electrical conductance, sheet resistance measurements were done by the 4-probe method at various points along the wafers radius. Figure 4.21 shows the variation of resistivity (sheet resistance × thickness) with wafer radius for TiN films deposited at different RF power and chamber pressure. In Figure 4.15, the power law correlation between resistivity and grain size was demonstrated. The resistivity trends in Figure 4.21 can be understood in the same way. In Figure 4.21a, at RF = 600W the increase in resistivity from wafer edge to the center (non-uniformity) is augmented as the

pressure is increased. This can be understood by and correlated to the increase in R_{GS} with chamber pressure at RF = 600W (green line). In Figure 4.21b, at P = 6mT the increase in resistivity from wafer edge to the center declines as the RF power is increased, and can be correlated to the decrease in R_{GS} with RF power at P = 6mT.

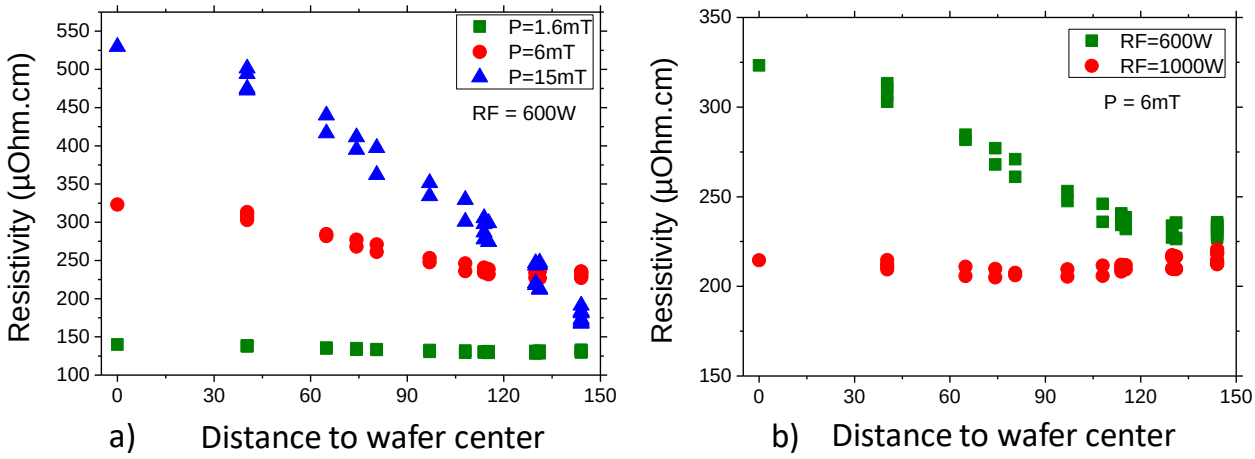


Figure 4.21: Variation of TiN film resistivity with wafer radius for a) RF power = 600W and different chamber pressure; b) P = 1.6mT and different RF power.

As discussed at the end of section 4.3.1, a modulation in TiN grain size cn have an effect on the V_T variability, gate patterning and diffusion barrier properties of the TiN film. Moreover, modulation in the sheet resistance of TiN can affect the electrical behaviour of CMOS devices. Generally, a non-uniformity of around 5% is acceptable in CMOS processing, and we see that TiN process conditions can induce a non-uniformity of about 25%, depending on the pressure and RF power used (Figure 4.20). Fortunately, the most interesting TiN process conditions found in this work to achieve unique grain orientation, i.e. P = 1.6 mT and RF = 600 W and 1000 W, have a grain size and Rs non-uniformity within the 5% limit.

4.3.7 Effect of substrate temperature

In order to determine the impact of substrate temperature during TiN deposition on its microstructure, in-plane and out of plane XRD was done to obtain the grain size and percentage of relative grain orientations. Substrate temperature was varied from -10°C to 50°C. The % change in grain size and relative amount of <200> orientation, with temperature variation, is reported in Figure 4.22a and b respectively, for different RF power and chamber pressure. The temperature variations is mentioned for each RF-PVD condition.

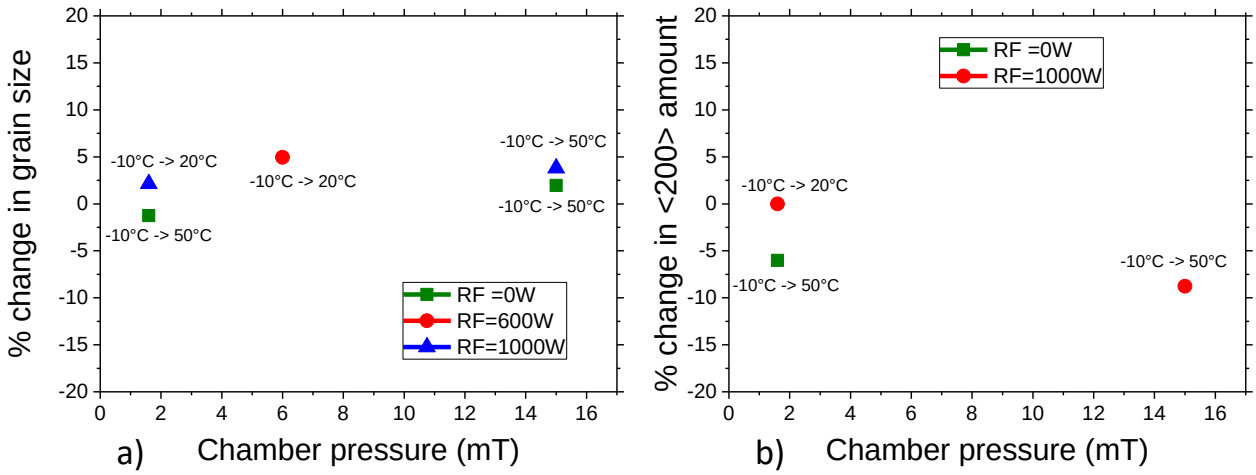


Figure 4.22: % change in a) grain orientation and b) amount of <200> orientation due to substrate temperature, with chamber pressure at different RF power.

From Figure 4.22, no obvious impact of substrate temperature is seen on the grain size (< 5%) or on the relative amount of <200> orientation, and the slight variations could be the result of reaching the accuracy limit of grain size measurements. This lack of a significant temperature impact can be due to the chamber hardware limitation, on which we could change the temperature only by 60°C and this range might not be sufficient to see an effect on TiN microstructure.

4.3.8 TiN grain analysis by ASTAR technique

ASTAR is a commercial system from NanoMEGAS [157], based on the Automated Crystal Orientation Mapping (ACOM) technique, used to obtain automated crystallographic and orientation mapping. It is based on a high energy electron beam in a TEM equipment where an electron beam is scanned over the sample area of interest. Electron diffraction pattern consisting of spots is formed in relation with the crystal orientation of the scanned area. Information on the crystal orientation is obtained by cross correlation of the diffraction pattern (spots) with simulated templates [158][159]. This working principle is shown in Figure 4.23.

In this section, we will present the capability of the ASTAR technique to map the crystal orientation of TiN metal layer. Later, the grain size and relative quantity of grains obtained from ASTAR will be compared to the ones obtained from in-plane and out-of-plane XRD measurements. ASTAR and XRD measurements were perfomed on blanket wafers with very thick TiN of 40 nm, deposited by the reference TiN RF-PVD conditions, over a Si substrare and then followed by a poly-Si layer deposition . A thick TiN layer is required to build a thick-enough lamella (or sample) which can be analysed by ASTAR, as it will increase the diffracted signal. ASTAR is performed in two orientations depending on the direction of the electron beam, these will provide grain orientation information about TiN top surface or the cross section as shown in Figure 4.24.

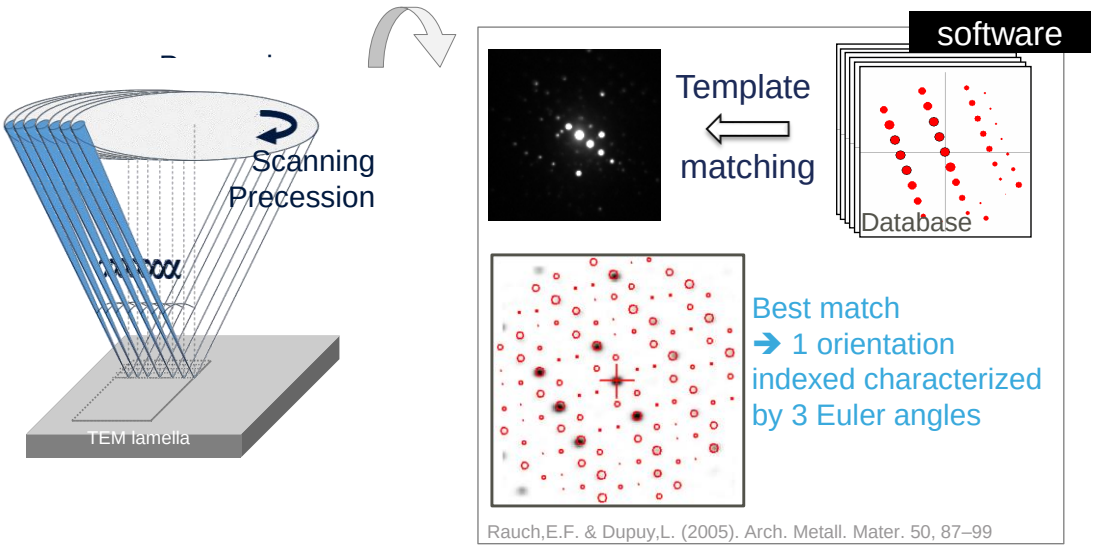


Figure 4.23: Working principle of the ASTAR technique [158]

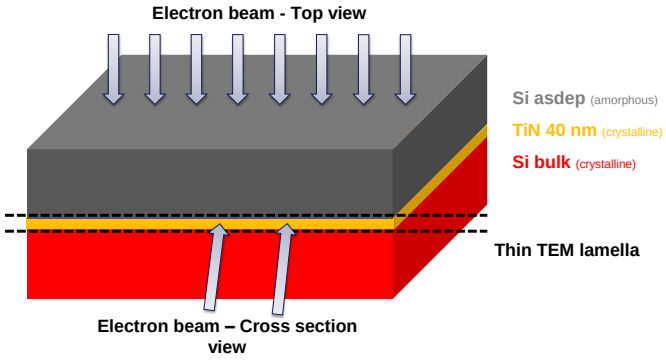


Figure 4.24: Schematic showing the sample used for ASTAR measurements and different analysis orientations

Grain orientation maps of TiN obtained from ASTAR are shown in Figure 4.25, both in the top view and the cross section view. It shows that many orientations are present in the TiN layer and the major ones are <111> (blue), <200> (red) and <220> (green). The other orientations observed in smaller amounts are <112>, <102>, <212>, <111> and <113>. Figure 4.25c shows the top view of a 70x30 nm^2 region of TiN gate, which is close to the size of a FDSOI transistor used. It shows that many grain orientations with different size are present within this area, which highlights the importance of MGG on V_T variability. From the top view image of the different grain orientations, the grain size and relative amount of grains can be obtained, simply by the area ratio of the grains. These are shown in Table 4.2, which also shows the values obtained from XRD in-plane and out-of-plane measurements.

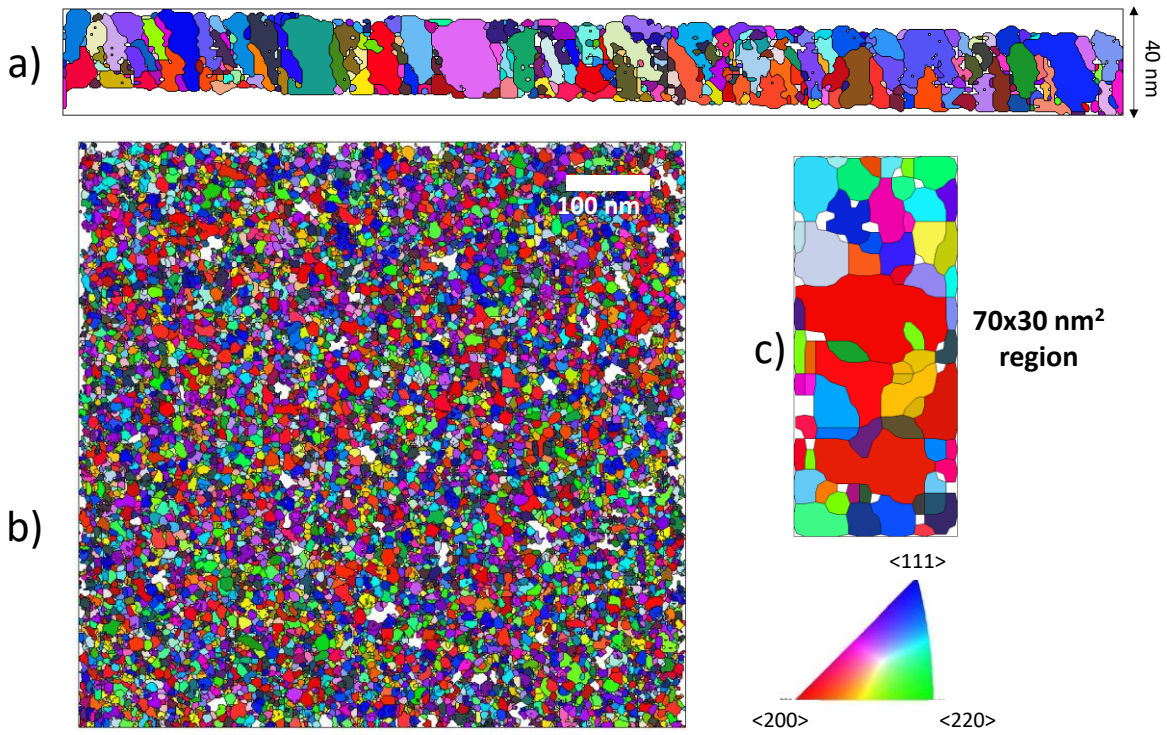


Figure 4.25: Grain orientation mapping of TiN layer in a) cross section view, b) top view and c) top view of an area of 70x30 nm^2 region

		Mean	<111>	<200>	<220>
ASTAR	Volume %		20%	21%	16.7%
	Vol. % (%111 + %200 =100)		49%	51%	
	ASTAR grain size	7.3 nm	7.3 nm	8.5 nm	7.4 nm
XRD	In-plane Grain size	7.8 nm	7.6 nm	8.4 nm	7.5 nm
	Out-of-plane Volume%		54 %	46 %	0

Table 4.2: Grain size and % of TiN crystallites orientation obtained from ASTAR and XRD

Note that as <112>, <102>, <212>, <111> and <113> have an even-odd combination of h,k,l indices, they will not be observed in an XRD spectra (section 2.2.3). For this reason ASTAR and XRD are compared in Table 4.2 only for <111>, <200> and <220> orientations. From Table 4.2, it is clear that the grain size obtained by ASTAR and in-plane XRD are similar. The relative percentage of grains of <111> and <200> can be compared between the two techniques by applying the condition: % 111 + % 200 =100 (as <220> is not observed in out-of-plane XRD), and these relative % for <111> and <200> orientations obtained from ASTAR and XRD are also quite close to each other. Therefore, we can conclude that ASTAR can be used to analyze the microstructure of TiN metal gate and gives additional information, i.e. an actual grain orientation mapping of the sample area. Further, there is a good matching of results obtained from ASTAR and XRD techniques.

4.4 Conclusion

In this chapter, the effect of TiN metal gate RF-PVD deposition conditions (chamber pressure RF power and substrate temperature) on the grain orientation, grain size, mechanical stress, resistivity, effective workfunction (WF_{eff}) and wafer-level non-uniformity has been investigated [PK-2]. The grain size and orientation is significantly modulated by varying RF power and chamber pressure during TiN RF-PVD process. Grain size as low as 5.2 nm has been obtained. Moreover, a unique TiN crystal orientation have been obtained for the first time, for few optimized process conditions [PK-2] [PK-6], which might have an impact on reducing threshold voltage variability.

Deposition rate is modulated significantly, knowledge of which is very important for a good thickness control of ultrathin TiN films. Mechanical stress varies over a large range and for some conditions of RF power and pressure, significantly low stress values are obtained that might improve gate patterning and adhesion of TiN films. Sheet resistance also varies over a large range and for the unique grain orientation conditions, sheet resistance value of 150 μΩ.cm is achieved that might improve device performance.

As the RF power and pressure are varied, MOS WF_{eff} is modulated but the trend is opposite to the expected modulation of the intrinsic TiN metal gate workfunction with grain orientation. On the contrary, WF_{eff} variation is well correlated to the Ti/N ratio, suggesting an effect related to dipole at SiO_2/high-k interface.

No obvious impact of substrate temperature between -10°C and 50°C is observed on the grain size or their relative orientation. Sheet resistance decreases along the wafer radius and has been correlated to the increase of grain size, and this non-uniformity is <5% for the unique grain orientation conditions. ASTAR technique is capable to map and image all crystal orientations present in the TiN metal layer with an excellent accuracy, which was not possible with XRD. The results on the grain size and relative amount of crystal orientations are quite similar to as obtained from XRD measurements.

5. Study of La and Al diffusion in HKMG stack

In this chapter, the diffusion of La and Al additives, introduced inside the HKMG stack for WF_{eff} engineering, is investigated. The main studied parameters are the temperature and time of the drive in anneal (DIA) step in the sacrificial gate process. Section 5.1 introduces the mechanisms and the fundamentals of the diffusion process. In section 5.2, the sacrifical gate process and the use of two different process flows (or devices), that have been used in this study to measure diffused dose and effective workfunction are presented.

Section 5.3.1 presents the X-Ray Fluorescence (XRF) results for diffused dose of La or Al in the gate stack depending on the DIA time and temperature. First, a methodology has been defined to correct the diffused dose from the no-anneal steps. Then, this section highlights the main differences between La and Al additive diffusion and different high-k. Section 5.3.2 first presents the modulation of effective workfunction or additive dipole by the DIA conditions. Then, the modulation of dipole and diffused dose by DIA conditions are correlated, and the results obtained on the two additives and different high-k have been compared.

5.1 Introduction to diffusion

Diffusion is the transport of a material into another one by their respective atomic motion. It takes place by statistical net motion from regions of high concentration to that of low concentration, i.e. by concentration gradient or in other context by a gradient in chemical potential. Diffusion is a thermally activated process with a strong dependence on temperature. Diffusion is used extensively in integrated circuit (IC) and micro-electro-mechanical systems (MEMS) manufacturing, with the use of high temperatures in the range of 700°C - 1200°C. Diffusion is used to introduce impurities (doping) into the silicon substrate, to control the Fermi level and so the electrical properties of the device such as resistors, diodes, BJTs, MOSFETs (P-type or N-type). It is also involved in the thermal oxidation of silicon to form the interlayer SiO_2. It takes place during chemical vapor deposition process used to deposit many material during IC processing, such as poly silicon, SiO_2, silicon nitride and metals such as tungsten, molybdenum, tantalum, titanium, etc. Doping by diffusion is very useful but sometimes unintentional dopants such as gold, copper or nickel can ruin solid-state electronic devices.

During diffusion the atoms move from one site to the next, and this movement can be of different types depending on the type of atoms involved and the defects present in the host material. The four type of possible diffusion mechanisms are shown in Figure 5.1. The two basic types, based on where the diffusing atoms occupy specific sites, are interstitial diffusion and substitutional diffusion. Other mechanisms can take place with the combination of interstitial and substitutional diffusion. A brief summary of these is given below:

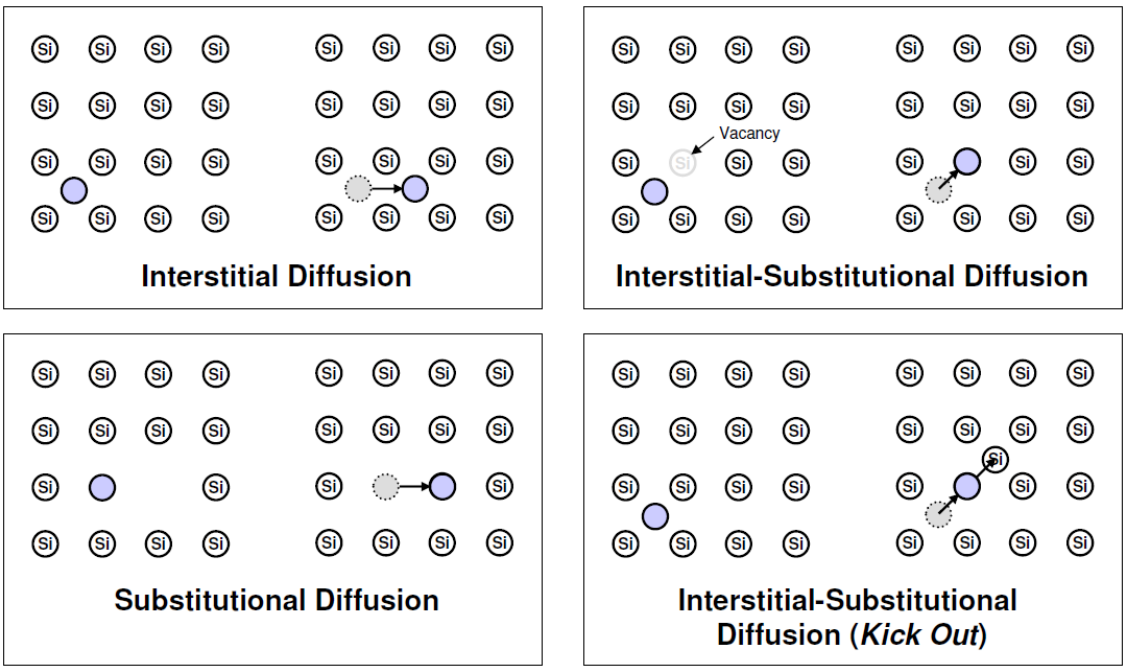


Figure 5.1: Different possible diffusion mechanisms in a solid material [160]

- **Interstitial diffusion**: In this form of diffusion, the impurity atoms in one interstitial site jumps to the neighboring site. Such a movement or jump of the interstitial atom involves a considerable distortion of the lattice, and this mechanism is probable when the interstitial atom is smaller than the atoms of the host material. So, small atoms like carbon, hydrogen, nitrogen, oxygen and some metals are able to diffuse though this mechanism. It is characterized by fast diffusivity as there are many interstitial sites.

- **Substitutional diffusion**: This diffusion is said to take place by substitutional or vacancy mechanism, when an atom at a normal site jumps into an neighboring unoccupied site (vacancy). Since the concentration of vacancies in a material are low, compared to interstitial sites, substitutional diffusion is much slower than interstitial diffusion. Substitutional elements (like dopants) are characterized by substitutional diffusion.

Diffusion is described by Fick's diffusion laws. Fick's first law describes how the flux (or flow) of dopant depends on the doping gradient (Figure 5.2a). It is described by the following equation:

$$J = -D\frac{dN}{dx} \tag{5.1}$$

where J is the flux of diffusing atoms inside the host material, D is the diffusion coefficient, N is the dopant concentration and dN/dx is its gradient along diffusing direction x. This is also known as the steady-state diffusion as the J (or dN/dx) does not change with time, as depicted in Figure 5.2a. Here N_{HI} and N_{LO} are respectively the constant concentrations at the entry and exit surfaces of the host material.

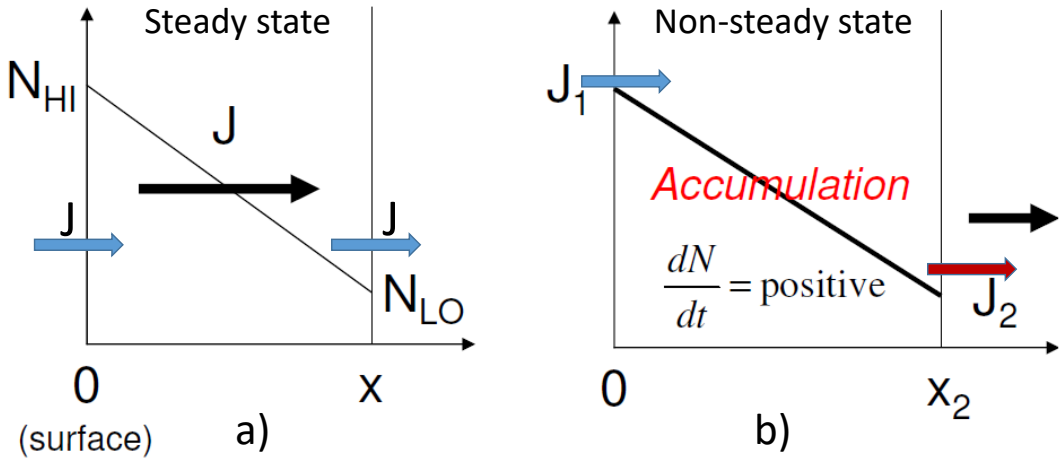


Figure 5.2: Concentration and flux of dopants in a) Steady state diffusion and b) Non-steady diffusion [160]

In practical situations, the concentration profile dN/dx, and so the flux J, changes with time. A flux being dependent on time will result in an accumulation of dopants inside the host material, as depicted in Figure 5.2b. Fick's second law deals with the time dependence of concentration and flux, it is also termed as nonsteady-state diffusion (depicted in Figure 5.2b) and given by Eq. 5.2.

$$dN/dt = - D \cdot (d^2N/dx^2) \qquad 5.2$$

Eq. 5.2 can have at least two solutions for N(x,t) depending on the different boundary conditions, as described below:

1) If the concentration of dopants is constant at the surface, i.e, $N(x=0,t) = N_o$, then the variation of dopant concentration with time is given by Eq. 5.3.

$$N(x,t) = N_0 \cdot \mathrm{erfc}(x/(4Dt)^{1/2}) \qquad 5.3$$

where $(Dt)^{1/2}$ is called the diffusion length and erfc is the complementary error function which is derived by integrating the normal probability function, as follows:

$$\mathrm{erfc}\,(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t \cdot t} \cdot dt \qquad 5.4$$

Total diffused dose Q in a time t is the integral of the dopant distribution N(x,t) from the surface to infinity.

$$Q(t) = \int_0^{\infty} N(x,t)\,dx = 2 \cdot No \cdot \left(\frac{Dt}{\pi}\right)^{0.5} \qquad 5.5$$

In this case the dopant concentration and the total dose varies with distance and time as shown in Figure 5.3.

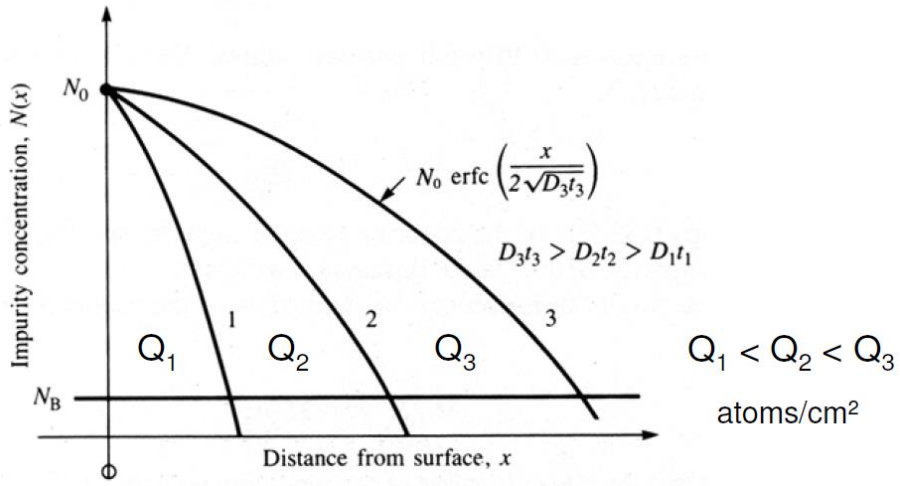


Figure 5.3: Variation of dopant concentration and total dose with distance inside the host material, at different times with fixed surface concentration [160]

2) If we assume that the total dose Q is always fixed, i.e. the integral of N(x,t) = Q, then N(x,t) is given by Eq. 5.6 and the profile is shown in Figure 5.4.

$$N(x,t) = \frac{Q}{\sqrt{\pi Dt}} \cdot e^{-\left(\frac{x.x}{4Dt}\right)} \qquad 5.6$$

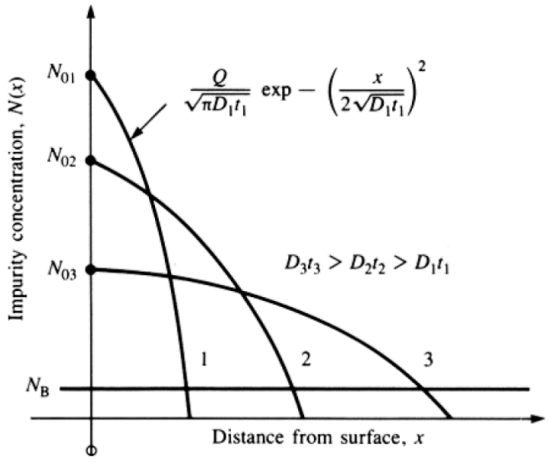


Figure 5.4: Variation of dopant concentration with distance inside the host material, at different times with fixed total dose [160]

Diffusion coefficient D is defined by Eq. 5.7

$$D = D_0 \cdot e^{-(Ea/KbT)} \quad 5.7$$

where D_0 is the temperature independent pre-exponential term, Ea is the activation energy for diffusion, Kb is the Boltzmann constant (1.38×10^{-23} J/K) and T is the temperature during diffusion (K). In order to diffuse, the dopant atoms jumps from one host material site to the other by squeezing through its neighbor atoms. The energy needed for this transition is the activation energy Ea, as shown in Figure 5.5. This makes the dopant diffusion a thermally activated process.

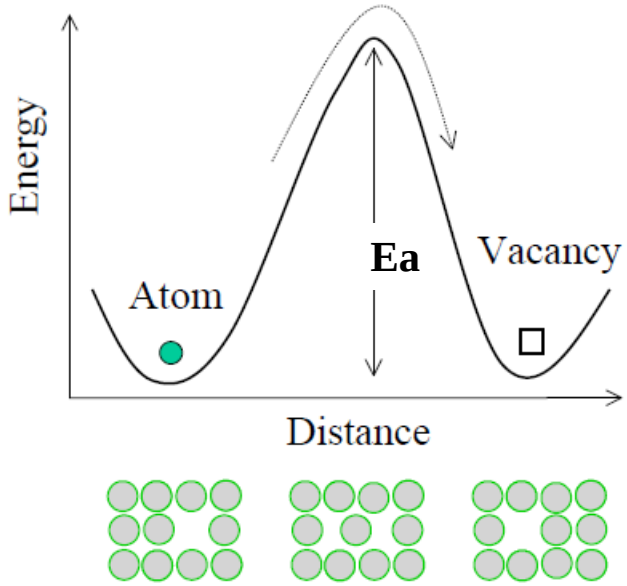


Figure 5.5: Schematic representation of diffusion of an atom from one site to the other, and the activation energy needed for this transition [161]

The average thermal energy of an atom is KT, which is much smaller than the activation energy Ea. Thus a large fluctuation in energy is required in order for the atom to jump from one site to the other. The probability of such fluctuation (or the frequency of jumps) is described by the Arrhenius equation described below:

$$R = R_0 \cdot e^{-(Ea/KbT)} \quad 5.8$$

where R_0 is proportional to the frequency of atomic vibrations. Ea depends on both the nature of the host material and the diffusing atoms. This way the diffusion coefficient in Eq. 5.6 follows the Arrhenius dependence on temperature. The second part of the diffusion coefficient, D_0, is different for every pair of diffusing and host species. In literature the typical values range from 10^{-11} for slow diffusing species to 10^{-5} for fast diffusing species.

Diffusion also depends on the microstructure of the host material. It is faster in poly crystalline materials compared to single crystal materials, as diffusion along grain boundaries is faster due to large spacing between atoms. In general, diffusion is fast for open crystal systems, lower melting point materials, materials with non-covalent bonding, smaller diffusion atoms, for cations and for lower density materials.

Secondary ion mass spectrometry (SIMS) measurements are able to measure the dopant concentration profile inside various materials. From previous studies based on time-of-flight SIMS (ToF-SIMS) measurements on diffusion of additives in HKMG of STMicroelectronics, the dopant profile at the surface was found to be constant [162]. As we use the same process for the HKMG stack, we thus assume that that our dopant profile will be as described in Figure 5.3 and the total dose as described by Eq. 5.5. Thus inserting the diffusion coefficient from Eq. 5.7 into Eq. 5.5 for total diffused dose Q, we get:

$$Q(t,T) = 2 \cdot No \cdot (D_0/\pi)^{0.5} \cdot (t)^{0.5} \cdot (e)^{-Ea/2KbT} \quad 5.9$$

$$Q(T) = K \cdot (t)^{0.5} \cdot (e)^{-Ea/2KbT} \quad 5.10$$

Therefore, the activation energy Ea and K (containing information on the pre-exponential D_0) can be estimated by plotting ln(Q) versus 1/T. In section 1.2.4, we presented the introduction of La and Al additives into the HKMG stack by the sacrifical gate process used, for effective workfunction WF_{eff} engineering. FDSOI devices require a fine tuning of the WF_{eff} in order to reach its specific requirements. In the next sections, we will investigate the modulation of the diffusion of La and Al in the gate stack, by varying the process conditions of the drive in anneal (DIA) step, i.e. the time and temperature of anneal.

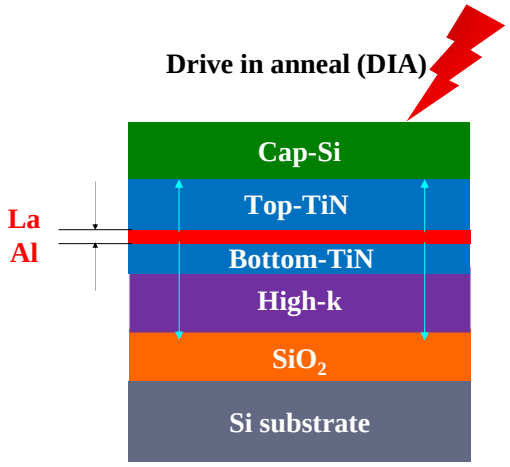
5.2 Device fabrication

The sacrificial gate structure used to introduce La and Al in the gate stack is shown in Figure 5.6a. It consists of 1 nm thermally grown SiO_2 interfacial layer, followed by a 1.8 nm thick HfSiON or HfON High-k layer deposited by MOCVD or ALD respectively. Then, the sacrificial gate stack consisting of La or Al layers sandwiched between a bottom pedestal TiN (0 or 1 nm) and the top capping TiN (4.5 nm) were deposited by RF-PVD. Next, silicon is deposited as a capping layer on the top, followed by a thermal treatment under N_2 atmosphere, at the choosen DIA temperature and time, in order to activate the diffusion of additives into the high-k/SiO_2 stack (Figure 1.18a). Figure 5.6b shows the DIA splits on temperature and time that were used in this work to study La and Al diffusion. Then the Poly Si and TiN (sacrificial gate) were removed by wet etching (Figure 1.18b). Now, two types of process flow were followed depending on the type of measurements conducted on the device thus fabricated:

- First is the process flow for blanket wafer used to study diffused additive dose in section 5.3.1. The process flow described above was done on Si substrates until the wet removal of the sacrificial gate, followed by XRF measurements to obtain the diffused additives dose. The process is the same as presented in Figure 1.18 for sacrificial gate, except that the final TiN and poly Si were not deposited (step 3). After the sacrificial gate removal by

wet etching, Ti signal was measured by XRF to make sure that the sacrificial gate has been removed completely and that only the La or Al dose diffused inside the gate stack will be measured. Indeed we confirm that, on all the splits the measured Ti signal was negligible, meaning that sacrificial gate was completely removed.

- Second is the process flow for the MOS capacitors, that was introduced in section 1.3.2 and also used in section 4.3.2, to study their effective workfunction. After the wet removal of the sacrifical gate, final TiN and poly-Si were deposited and S-D anneal was done on STI SiO_2 cavities, followed by CMP patterning in order to fabricate MOS capacitors, as shown in Figure 1.21. A bevel interlayer (section 1.3.3 and Figure 1.22) was used in order to study dipoles formed by La or Al at the SiO_2/high-k interface through the analysis of flatband voltage (V_{FB}) versus EOT plots.



DIA splits

Temperature (°C)	Time (s)
700	0
760	5
820	10
900	30
	45
	60
	110

b)

Figure 5.6: a) Schematic of the HKMG stack and b) DIA process splits used to study their impact on diffused dose and effective workfunction

5.3 Results

In this section, we will first present the results on diffused dose of La measured by XRF or Al measured by XPS in the gate stack, with DIA time and temperature. Next, the results on effective workfunction and dipole modulation by DIA conditions will be presented. Finally, the correlation between diffused additive dose and the WF_{eff} or dipoles formed by these additives will be made.

5.3.1 Modulation of diffused additive dose by DIA conditions

Diffusion of La in HfON/SiO_2 stack

Blanket wafer process, containing the HKMG stack as shown in Figure 5.6a, with HfON as the high-k and 4Å La deposited directly on it (no pedestal TiN used) was used. Various DIA time and temperature were used to have a wide range of diffused La dose: with 10 s at temperatures = 700,

760, 820 and 900ºC, or time = 0, 10, 30, 60, and 110 s at 760°C. A time of 0 sec (no-anneal) means that the entire sacrificial gate process was done except the DIA step. Indeed, with this skip-anneal condition the gate stack will contain La dose diffused by steps other than the DIA, particularly the poly-Si deposition step done at 525°C for about 1000s. The total dose calculated by XRF will be a combination of dose diffused during thermal processes before DIA and the DIA itself. Thus, information on the no-anneal dose will be helpful to correct the additive dose diffused, and thus obtain a correct diffusion model that enables the extraction of reliable activation energy values. We indeed found a reasonable amount of diffused dose ($4.36 \cdot 10^{13}$ at/cm^2) even without the DIA step. Thus, we must define a method to correct the diffused dose of other splits with different DIA conditions, it is presented below:

In Eq. 5.9 and 5.10, we defined the total dopant dose (Q) diffused in a material at a certain annealing temperature (T) and for a duration (t). In our case, Q is a combination of the no-anneal and the DIA step. Thus for each temperature, we can define an effective time $t_{0'T}$ that gives the same no-anneal dose, if it was done at the DIA temperature T (instead of at 525°C for 1000 s). Thus Eq. 5.9 can be written for the dose diffused during the no-anneal step (Q_0) as:

$$Q_0 = K \cdot (t_{0'T})^{0.5} (e)^{-\frac{Ea}{2KbT}} \quad 5.11$$

The diffusion equation can be written for the total diffused dose Q as:

$$Q(t, t_{0'T}, T) = K \cdot (t + t_{0'T})^{0.5} (e)^{-Ea/2KbT} \quad 5.12$$

where T and t are the temperature and time of the DIA step alone. Thus for each DIA temperature and diffused dose, a corresponding value of $t_{0'T}$ can be calculated from Eq. 5.11 and Eq. 5.12 by:

$$Q(t, t_{0'T})/Q_0 = \frac{(t + t_{0'T})^{0.5}}{(t_{0'T})^{0.5}} \quad 5.13$$

$$t_{0'T} = \frac{t}{(\frac{Q}{Q_0})^2 - 1} \quad 5.14$$

As the total effective diffusion time ($t+t_{0'T}$) will be different at each temperature, the actual dose Q cannot be compared directly to calculate the activation energy. In order to compare the diffused dose at different temperatures to calculate Ea and K, a dose normalized to a given time is defined as:

$$Q (T)^* = \frac{Q}{(1+t0/t)^{0.5}} = K \cdot (t)^{0.5} \cdot (e)^{-Ea/2KbT} \quad 5.15$$

where t = 10 s. Q(T)* is calculated for the different temperatures, and it is the dose that we should have with only the DIA time (10 s). The raw dose Q and the corrected dose Q* are plotted versus the inverse of the temperature (1/T) in Figure 5.7, for a DIA time of t = 10 s. We see that the diffused dose changes with temperature in both cases and ln(Q) varies linearly with the inverse

of the temperature. This behavior is what was predicted in Eq. 5.9, meaning that the dose follows the Arrhenius law, i.e. exponential behavior of dose in temperature. The value of the diffused dose, K and the activation energy changes after correcting the data with the no-anneal dose. This means that La diffusion during process steps, other than DIA, is not negligible and must be taken into account. The dose decreases especially for lower temperature and Ea increases after correction, as we take into account the effect of the no-anneal step.

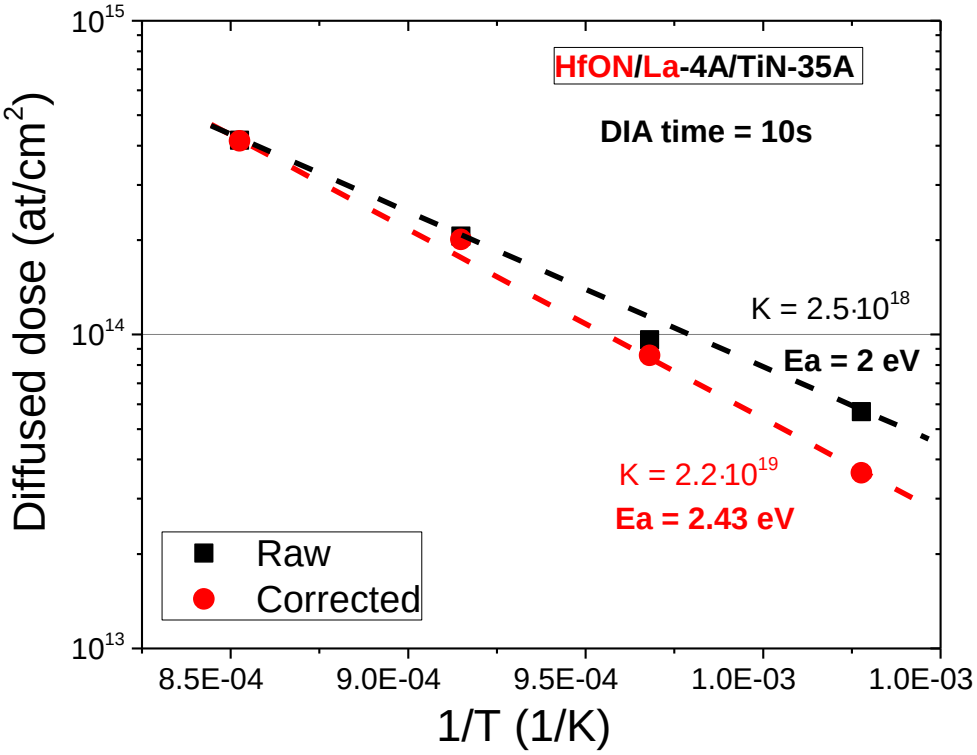


Figure 5.7: Variation of diffused La dose with the inverse of DIA temperature (1/T), for raw diffused values and corrected by the no-anneal dose

Diffused dose Q was measured for different DIA time t = 0, 10, 30, 60, 110 s, at a temperature of 760°C. Figure 5.8 shows the variation of this diffused dose versus DIA time, and similar to the trend in Figure 5.7, the effect of no-anneal dose is seen with some diffused dose with zero DIA time. Moreover, the ideal diffused dose can be calculated from the model presented above, by inserting the extracted values of K and Ea, and value of $t_{0,760}$ obtained for T =760°C in Eq. 5.15. Figure 5.8 compares the raw dose with the modelled dose, and the raw dose follows the modelled dose at small DIA times but then deviates from it. The raw dose seems to follow a power law in time, but with a lower exponent (0.4) compared to the ideal or modelled dose curve (0.5 in Eq. 5.9). This deviation from the ideal behavior can be due to the fact that the ideal behavior is expected for diffusion in an infinitely thick material, but our gate stack is only 3 nm thin and quite complicated with a bilayer structure (high-k/SiO_2). Moreover, La reactions with the high-k and SiO_2 layers [138] [163], which might be slowing down the diffusion process inside the gate stack, especially at the interface between them. This phenomena has been previously observed by Z. Essa, C. Gaumer et al. [162].

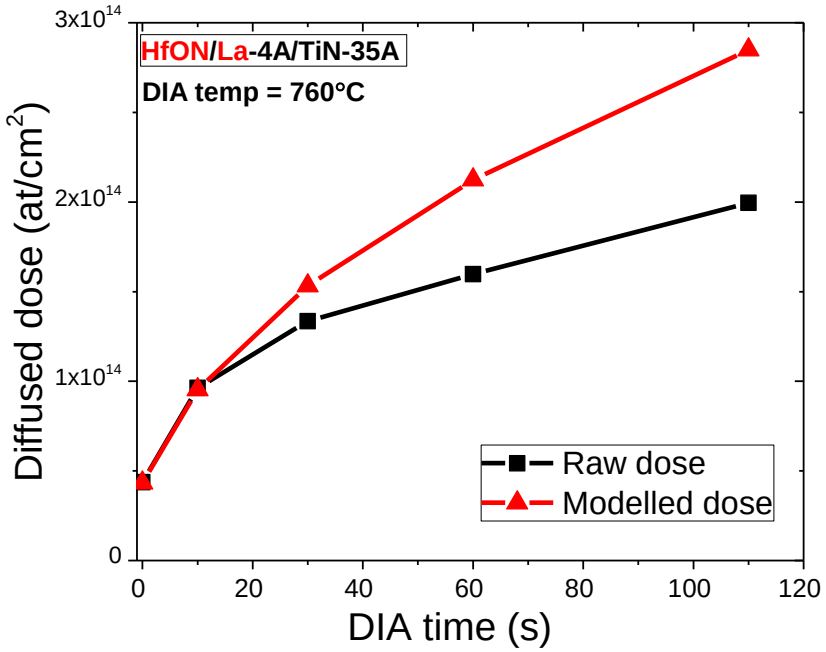


Figure 5.8: Variation of the raw diffused dose and the modelled dose versus DIA time

Diffusion of La in HfSiON/SiO_2 stack

For these studies, HfSiON was used as the high-k in the gate stack of Figure 5.6a, and La 6Å layer was deposited on top on it. DIA was performed for 10s at temperatures of 700, 760, 820 and 900ºC, and DIA duration was 0, 5, 10, 30, 60, 110 s at 900°C. As a reminder, DIA time of 0 s is where annealing was skipped. As done with HfON in the previous section, t0 is calculated for each temperature, leading to the calculation of corrected dose Q(T)*. Figure 5.9a shows, the variation of diffused dose versus the inverse of DIA temperature, with or without correction with the no-anneal dose (raw-Q and corrected-Q*). As with Figure 5.7, the diffused dose varies significantly with temperature in an exponential manner. The value of the diffused dose and activation energy changes by correcting the diffused dose with the no-anneal dose, highlighting that some La diffuses inside the gate stack even without the DIA. However, this change due to the correction is a bit lower than observed on HfON high-k. Compared to La diffusion in HfON, activation energy is higher in HfSiON and the diffused dose is lower.

In Figure 5.9b, diffused dose and the ideal diffused dose are plotted with the DIA. The two curves are quite the same, showing that diffusion of La inside HfSiON follows the ideal diffusion model predicted by the diffusion equation in Eq. 5.8. Moreover, the dose diffused in the no-anneal steps (t=0) is very low at this high temperature.

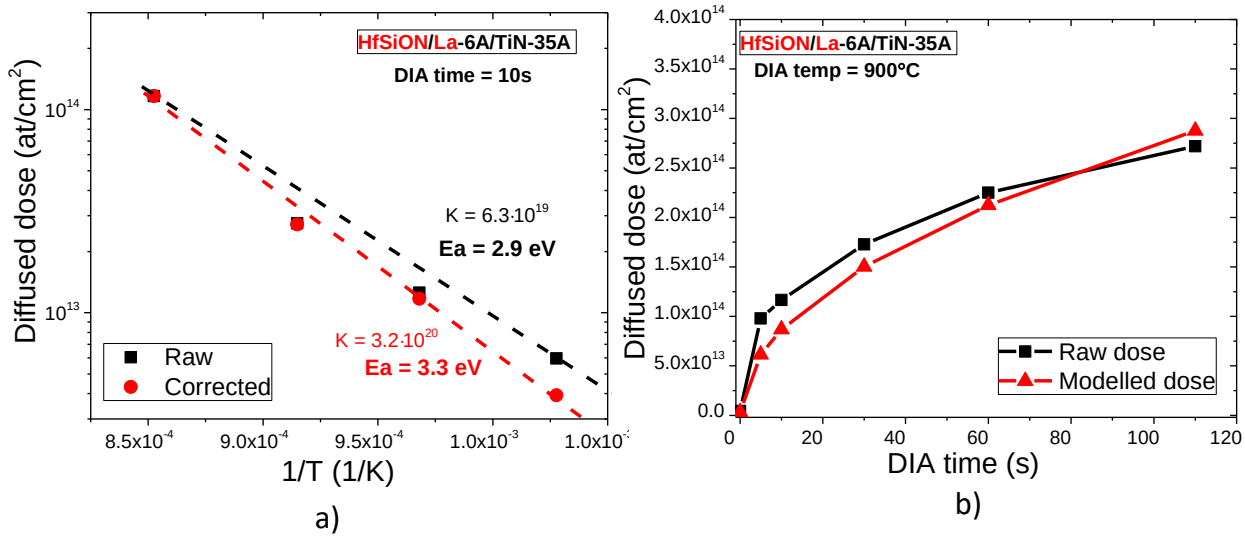


Figure 5.9: a) Variation of diffused La dose with inverse of DIA temperature and b) Variation of the raw diffused dose and the modelled dose versus DIA time

Diffusion of Al in HfON/SiO_2 stack

In order to study Al diffusion, Al 2.2Å layer was deposited either directly on top of HfON or above a pedestal TiN, as described in Figure 5.6a, to investigate the effect of DIA conditions and of the pedestal TiN on Al diffusion in the gate stack. DIA temperature was varied from 700 to 900ºC at 10s, and DIA times of 0, 10, 30, 45 and 60 s at 760°C were used. Figure 5.10a shows the variation of diffused Al dose with inverse of DIA temperature and DIA time. Again, there is a difference between the raw (Q) and corrected (Q*) values of dose and activation energy (from 1.4 to 1.76 eV), but this correction seems slightly less important than it was for La diffusion in HfON due to a relatively smaller value of $t_{0,750}$. Diffused dose is lower and the activation energy of 1.76 eV obtained is smaller compared to La diffusion on HfON (Figure 5.7).

Figure 5.10b plots the raw diffused dose and the modelled dose with the DIA time. We see that a significant dose is diffused even at zero DIA time. The dose begins to saturate if compared to the modelled dose. This does not seem to be due to the exhaustion of the as-deposited Al layer, because the dose obtained in Figure 5.10a with the highest temperature is higher than the saturation value with DIA time. This might be explained by the reaction of Al with the HfSiON/SiO_2 layers, similar to the case of La, or due to the formation of AlN clusters at high Al doses [135]. Both these phenomena can stop Al diffusion into the high-k, and hence Al diffusion does not follow the ideal diffusion equation in Eq. 5.9. Lastly, diffused Al dose is higher and Ea is lower compared to La in HfON.

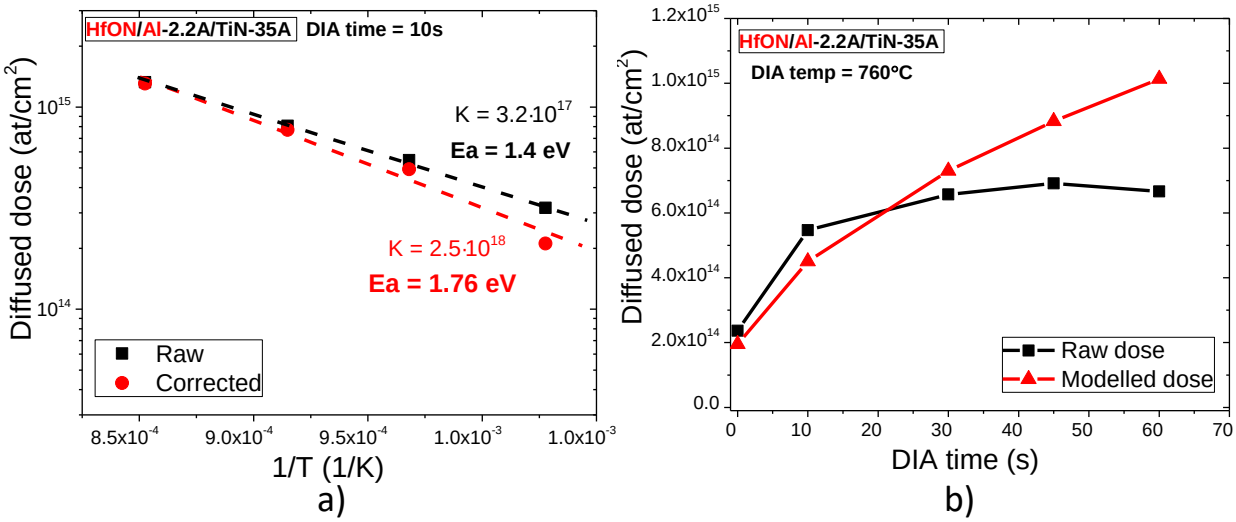


Figure 5.10: Variation of diffused Al dose with inverse of DIA temperature and b) variation of the raw diffused dose and the modelled dose versus DIA time

Next, we will study the impact of the pedestal TiN on the diffusion of Al. For this, we will compare the diffused Al dose on splits with or without the pedestal TiN (see Figure 5.6a). Figure 5.11 compares the temperature behavior of the raw Al dose Al for the two splits. Here we could not correct the raw values to obtain Q*, due to the lack of no-anneal data for the pedestal TiN case. Moreover, even if the no-anneal data was available, the model presented earlier will not be valid due to the presence of the pedestal TiN (a bilayer of TiN+oxide). Thus, the activation energy comparison is not reliable because the model presented above is not valid here.

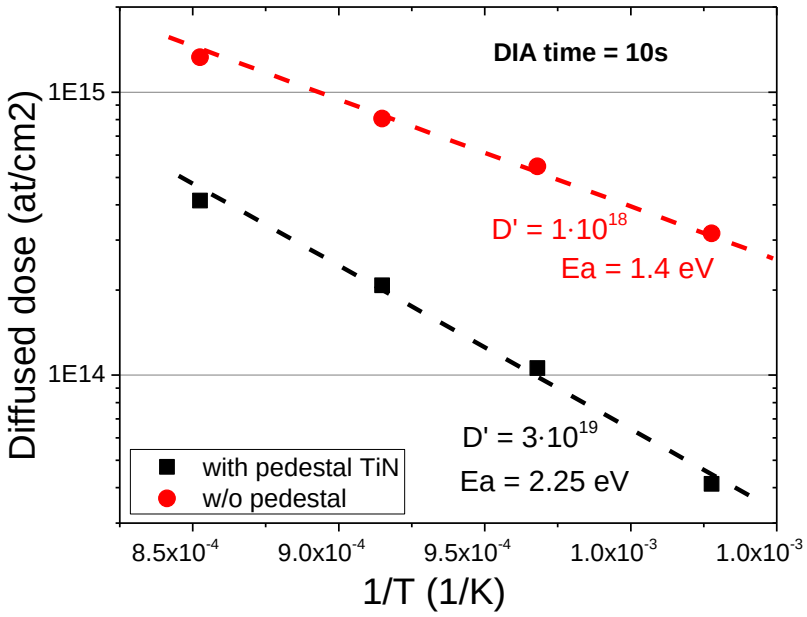


Figure 5.11: Variation of diffused Al raw dose with the inverse of DIA temperature (1/T), with or without a pedestal TiN on HfON

Nevertheless, compared to diffusion without any pedestal TiN, diffusion through a pedestal TiN is significantly lower even at the highest temperature and this difference increases as the temperature is lowered. This shows that the pedestal TiN layer acts like a very efficient diffusion barrier for Al. The lower diffusion with pedestal can be explained by the lower concentration of dopants (N_0 in Eq.5.9) at the surface of the high-k, due to the TiN pedestal layer.

Diffusion of Al in HfSiON/SiO_2 stack

For this study, a 2.2Å Al layer was deposited directly on the high-k and DIA was done for temperatures between 700-900 ºC for 10s, and times between 0 - 110 s at 820°C. Figure 5.12 shows the variation of dose with inverse of DIA temperature and time, and comparing the raw (Q) and corrected (Q*) curves shows that some Al diffuses even before the DIA step and the activation energy changes from 1.41 to 1.92. Moreover, the variation of the raw dose with DIA time does not follow the modelled dose behavior and saturates. Diffused dose at the highest temperature is higher than this saturation value, so this behavior again might be due to the reaction of Al with HfSiON/SiO_2 stack or due to formation of AlN clusters at high Al dose [135].

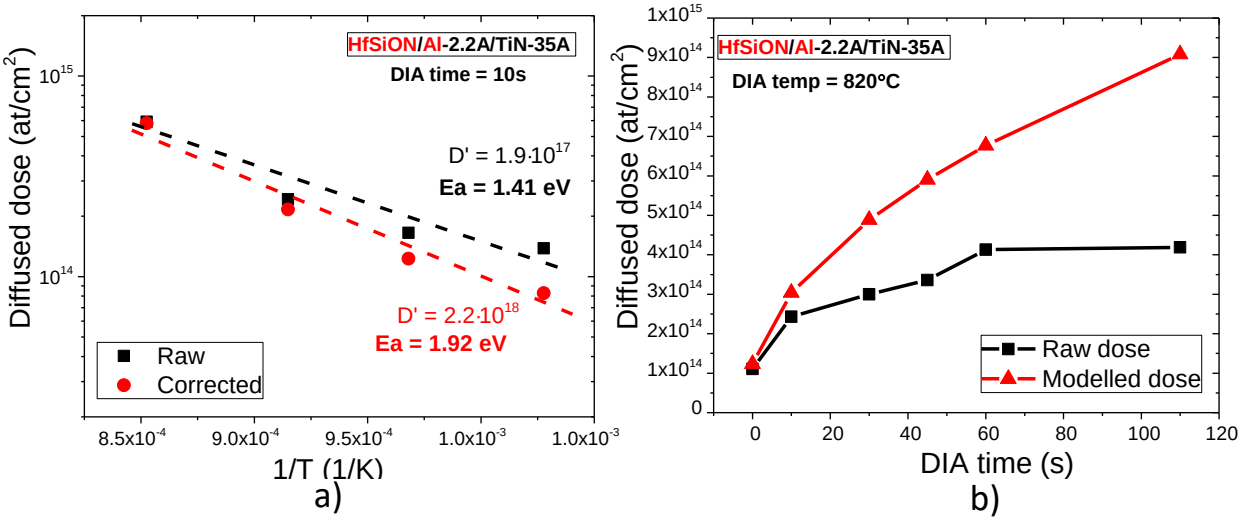


Figure 5.12: Variation of diffused Al dose with inverse of DIA temperature and b) Variation of the raw diffused dose and the modelled dose versus DIA time

Summary

Study of the diffusion of Al and La additives with DIA temperature and time has been conducted on different high-k materials. Some of the as deposited La and Al diffuses into the gate stack in other steps before the DIA, probably in the poly-Si deposition step. A model is proposed to correct the dose, by taking into account the contribution of the no-anneal steps. In all the cases, the diffused dose has an exponential (Arrhenius) dependence with temperature, as predicted by diffusion Eq. 5.9 and the corrected dose allows to calculate the value of K (estimation of $N_0 \cdot (D_0)^{1/2}$) in Eq. 5.9) and activation energy Ea, as shown in Figure 5.13. The activation energy for La is higher in HfSiON compared to HfON, meaning that La diffuses faster in HfON and so the diffused dose

is higher. For Al, activation energy is lower and so the diffused dose is higher in HfON compared to HfSiON, but this difference is smaller than obtained for La. This difference between HfON and HfSiON might be due to the low volumetric density and higher number of defects in HfON, compared to HfSiON, which are the precursor sites for diffusion. In HfON and HfSiON, diffused dose of Al is higher as compared to La with a lower activation energy, which might be explained by the fact that the size of Al atoms are smaller than La. Due to a smaller size, Al atoms are able to easily jump from one site to another in the high-k. For all cases, the value of K increase monotonously with the activation energy.

Diffused dose has a power law dependence in DIA time, where the power law exponent is smaller than as predicted in Eq. 5.9. Moreover, for Al, the diffused dose even saturates after a certain DIA time. The power law exponent slightly improves when the dose is corrected with the no-anneal dose, but still remains different from the ideal value of 0.5. This low exponent or saturation does not seem to be due to exhaustion of the as deposited La or Al layers, but rather due to reactions of La and Al with the high-k/SiO_2 stack [138] [162][163] or due to the complexity and ultrathin oxide layers. For Al, formation of AlN clusters might also be contributing to the dose saturation with time. Therefore it seems that the diffusion mechanism is thermally activated and a kinetic reaction mechanism between La or Al and the gate oxides, that can explain the diffusion saturation with DIA time.

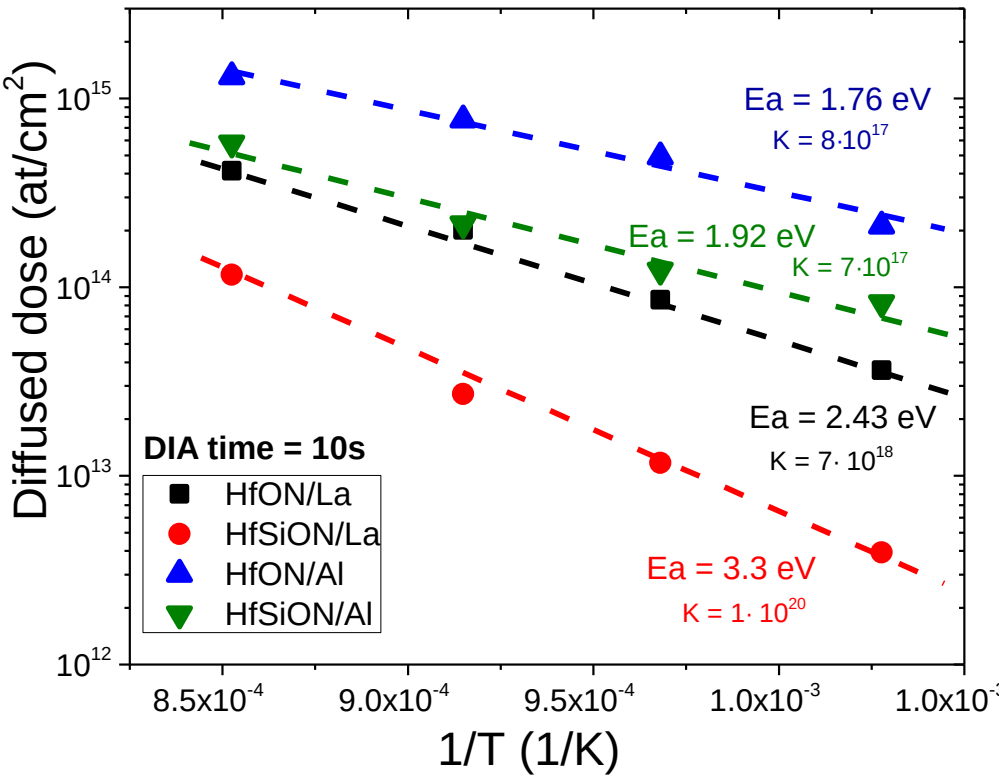


Figure 5.13: Variation of corrected dose of La and Al dose with the inverse of DIA temperature (1/T), on different high-k

5.3.2 Modulation of effective workfunction by DIA conditions

The process to introduce La or Al in the gate stack for workfunction enginnering, by forming dipoles at the SiO_2/high-k interface (δ), was presented in section 1.24. In the previous section we have investigated the effect of the DIA conditions on the total diffused dose inside the HKMG gate stack. Modulation of the effective workfunction by these additives will depend on their dose present at the SiO_2/high-k interface, which might be related to the total dose in a complicated manner. In this section, we will investigate the impact of the DIA conditions, same as used in the previous section, on the effective workfunction and the dipole formed by the additives in the gate stack. Finally, a correlation will be made between the δ and the total diffused dose.

For these studies, sacrificial gate stack with MOS capacitor process flow was used as described in section 5.2. La or Al were deposited on top of the HfON or HfSiON and DIA was done for several splits on temperature and time. C(V) measurements were conducted on dies along the wafer radius with a varying EOT, and flatband voltage V_{FB} and WF_{eff} were obtained from fitting them (as done in section 4.3.5 in Eq. 4.1 and 4.2). Figure 5.14 and Figure 5.15 shows the WF_{eff} versus EOT plots for devices, with La on HfON/SiO_2, having variations in the DIA temperature and time respectively.

In Figure 5.14, the slope of the plot (or fixed charges) for 820°C is quite different due to a different Si doping compared to other splits, but it will change the value of the La or Al dipole. As the extrapolation of WF_{eff} to EOT = 0 is equal to $WF_M + \delta$, the value of δ can be estimated from these plots because WF_M is not expected to change with La or Al diffusion. We see that the WF_{eff} and δ decreases as DIA temperature and time is increased. This can be understood by the trends on Figure 5.7 and Figure 5.8 where La dose is increasing with temperature and time, and we know that WF_{eff} decreases as La is added and forms dipoles inside the gate stack.

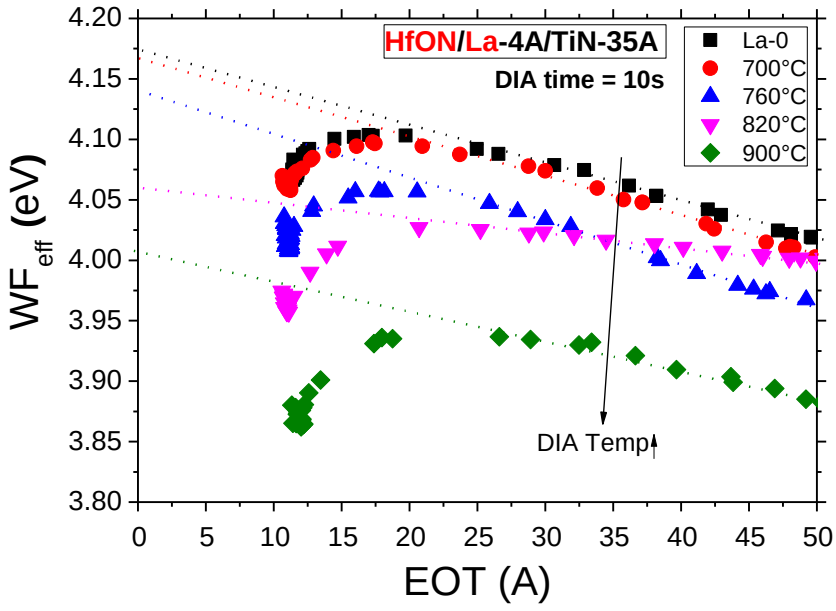


Figure 5.14: WF_{eff} vs EOT plots for devices with variation in DIA temperature, showing the effect of La addition

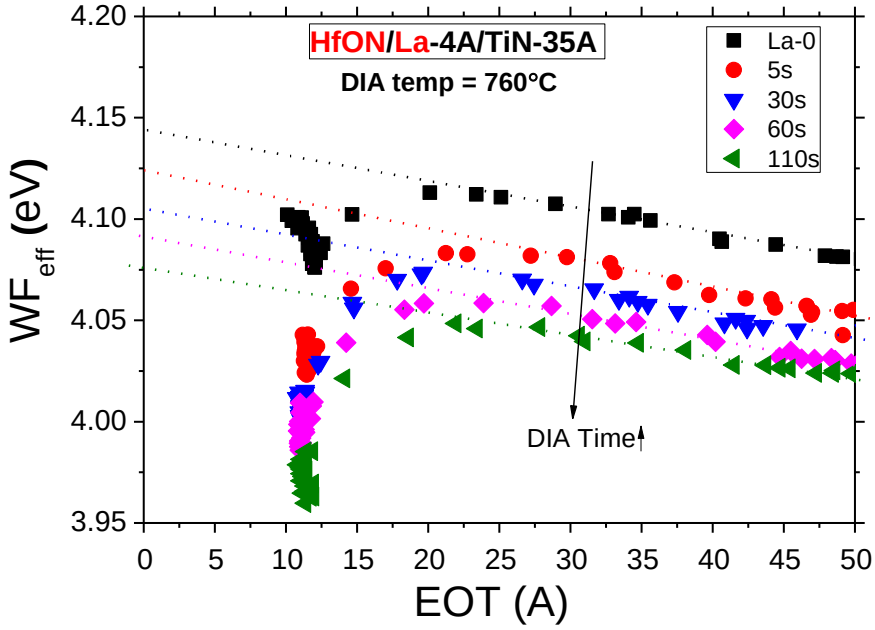


Figure 5.15: WF_{eff} vs EOT plots for devices with variation in DIA time, showing the effect of La addition

Similarily, Figure 5.16 and Figure 5.17 presents the results for Al diffusion on HfON, with DIA temperature and time. WF_{eff} and δ increases with temperature and time and can be understood by increases of Al dose with temperature and time in Figure 5.10. This trend is the opposite of what was observed for La because the electrical impact of Al dipole at the SiO_2/high-k interface is opposite to that of La. In Figure 5.16 and Figure 5.17, the results of no-anneal split are also plotted and we see that the workfunction change without any anneal is significant, that is similar to as observed for diffused Al dose.

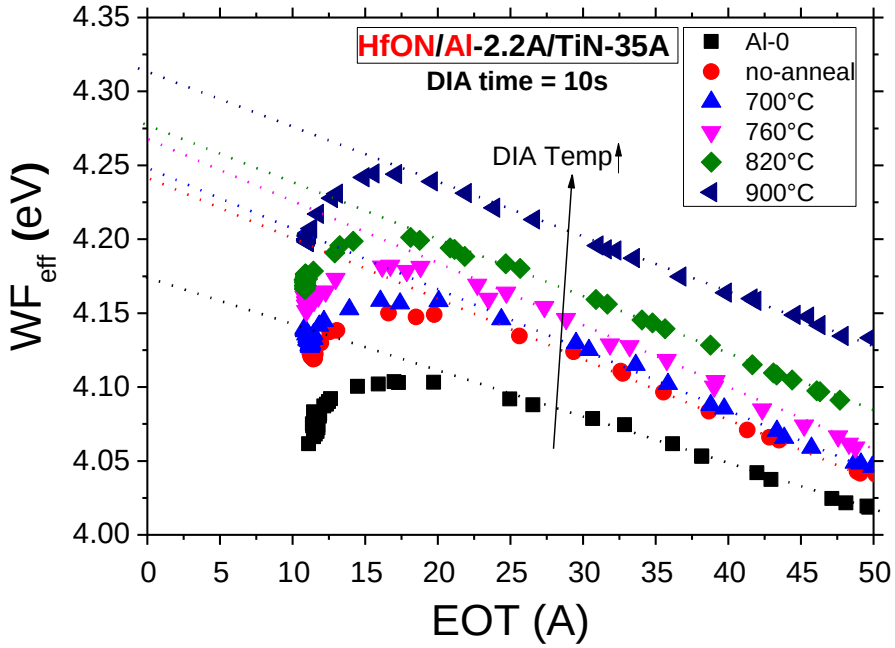


Figure 5.16: WF_{eff} vs EOT plots for devices with variation in DIA temperature, showing the effect of Al additio

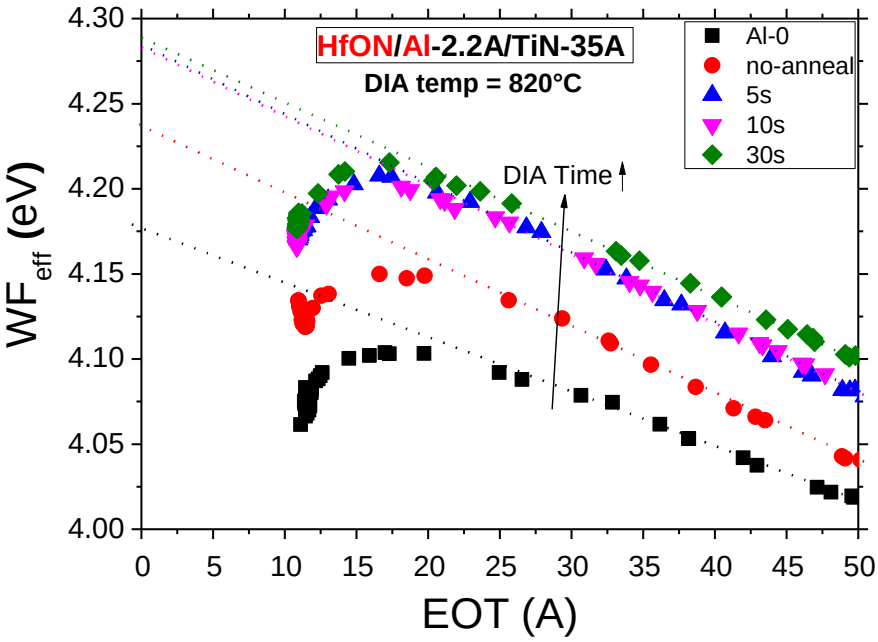


Figure 5.17: WF_{eff} vs EOT plots for devices with variation in DIA time, showing the effect of Al addition

From such plots of WF_{eff} vs EOT, the $WF_M + \delta$ (WF_{eff} @ EOT = 0) is obtained for devices without additives and devices with the diffusion of La and Al on HfON and HfSiON, with variations in DIA temperature and time. For the devices with La or Al introduced in the gate stack, their $WF_M + \delta$ value is subtracted from the device wthout any additives. This difference gives the δ value only due to the addition of La or Al at the SiO_2/high-k interface.

It is important to estimate the variation of electrical strength of additive dipoles and to compare the dipole between Al and La on different high-k, normalized by the diffused dose. In order to obtain this information, variation of the diffused dose from Figure 5.13 and that of dipole modulation δ with temperature and time are compared in Figure 5.18 and Figure 5.19, for Al and La on different high-k. In these figures, the diffused dose modulation achieved with changing DIA time (green) and DIA temperature (red) are shown. The black line represents the linear behaviour that was previosuly obtained by changing the as deposited thickness of La and Al [40]. In the case of Al, the δ shift follows a linear behaviour for small doses but then deviates and begins to saturate for higher doses of Al introduced in the gate stack. This behaviour on HfON is similar irrespective of dose introduction by time or temperature variation, but on HfSiON it is different. On HfSiON, dipole shift also shows a saturating behaviour for higher doses. When compared at the same diffused dose, the change in δ for HfSiON is a bit smaller compared to that on HfON, thus indicating that Al dipole on HfON is slightly stronger.

In the case of La, the δ shift follows the linear behaviour in a better way compared to Al, except for the case of lowest diffused dose by temperature. This trend of δ versus dose is quite the same for dose introduced by DIA time or temperature variation. δ shift on HfON increases monotonously with no saturation but for HfSiON the δ shift saturates at higher values of doses, similar to the case of Al introduction on HfSiON. When compared at the same diffused dose, the change in δ for HfSiON and HfON is almost the same at the same dose, meaning that dipoles on both the high-k

materials are equally strong. Comparison between La and Al results in Figure 5.18 and Figure 5.19 shows that the δ shift at the same diffused dose is higher for La compared to Al on both HfON and HfSiON. This confirms, thre results obtained earlier [40], that dipole formed by La has a stronger electrical impact, compared to that of Al. Moreover, the roll-off phenomena for La is stonger compared to Al, as seen in Figure 5.14 and Figure 5.16.

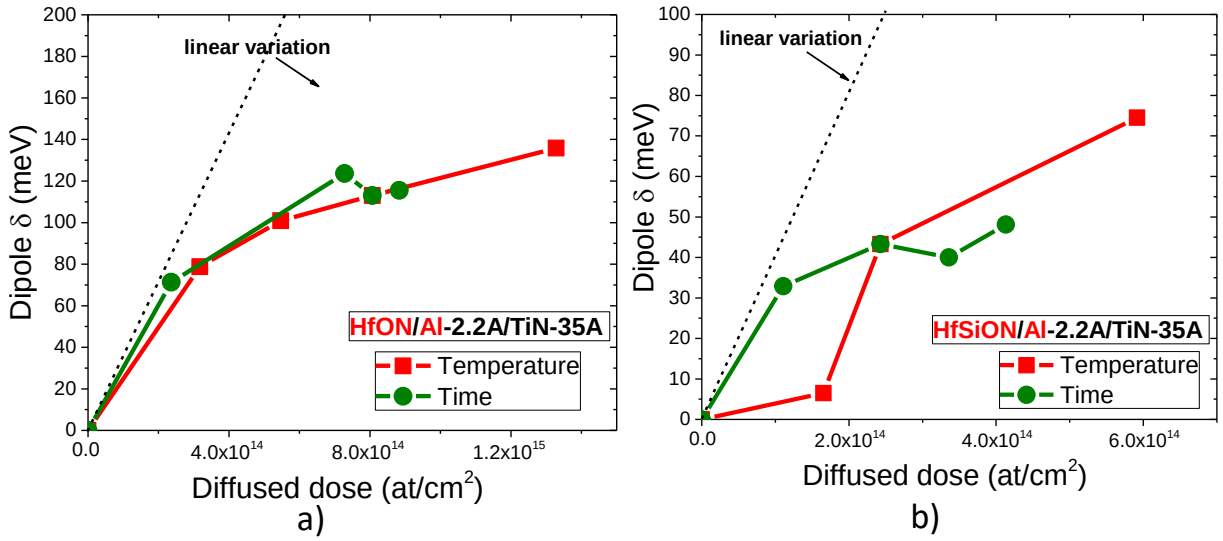


Figure 5.18: Variation of dipole at the SiO_2/high-k interface versus the diffused Al dose on a) HfON and b) HfSiON

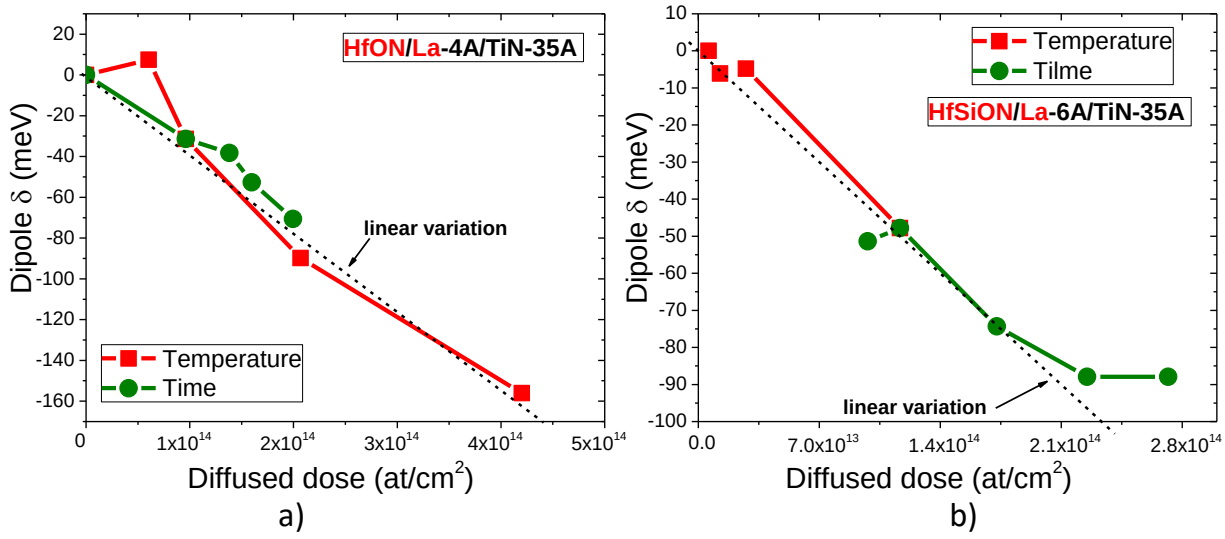


Figure 5.19: Variation of dipole at the SiO_2/high-k interface versus the diffused La dose on a) HfON and b) HfSiON

5.4 Conclusion

In this chapter, the impact of the drive-in-anneal (DIA) conditions, of the sacrificial gate stack, on the diffusion of La and Al additives and the effective workfucntion (WF_{eff}) are investigated. For both the additives, the diffused dose changes significantly with temperature and follows the Arrhenius law of exponential behavior in the DIA temperature. Moreover, the diffusion of these additives, in steps before the DIA step, is significant and a model is proposed to correct the dose. With this correction, an increase in the temperature activation energy and time power law exponent is observed. Studied as a fucntion of time, the variation of dose does not necessarily follow the ideal power law, and in the case of Al the dose even saturates for higher DIA times. This might be due to the ultrathin and bilayer gate stack, chemical reaction of dopants the gate oxides or due to formation of AlN clusters in them.

For both La and Al, the dose diffused in HfON is higher than in HfSiON, probably due to a lower volumetric material density and higher number of defects in HfON. Moreover, in both HfON and HfSiON, diffused dose of Al is higher compared to La, which might be due by the smaller size of Al atoms compared to La. The diffused dose increases as the activation energy is decreased, which can explain the differences in diffused dose between the different additives and high-k's.

The effective workfunction shift with the DIA conditions has been correlated with the diffused additive doses. For La addition, on both HfON and HfSiON, the workfunction changes linearly with the diffused dose, quite similar to studies done earlier, irrespective of the method used to introduce the dose (by varing DIA time or temperature). In the case of Al addition, the dose follows a linear behaviour for small doses but then deviates and saturate for higher doses. Finally, these results show that Al or La dipole on HfON is stronger than on HfSiON, and La dipole is stronger than that of Al on both the high-k's.

6. X-ray Photoelectron Spectroscopy under bias

This chapter presents the development and validation of XPS under bias technique to analyze HKMG stack band energies, with the main aim to localize different dipoles. Section 6.1 presents the need for this technique and a comparison of its benefits with capacitance measurements will be done. The state of the art and the major issues related to the conventional the XPS and XPS under bias techniques will be presented.

Section 6.2.1 presents the specific test structures, used to perform XPS under bias measurements, and their process flow. In section 6.2.2, the biasing issues occuring due to different parasitic serial resistances will be discussed, and section 6.2.3 will present the biasing solutions developed by a specific methodology that combines electrical measurements, XPS under bias measurements and modeling on different devices.

The various experimental issues, related to the XPS equipment and the samples, and their solutions will be discussed. This includes the issue of device location inside the XPS equipment in section 6.3.1, and the estimation of the exact size and the position error of the X-ray beam used to perform XPS measurements in section 6.3.2. Section 6.3.3 deals with the impact of the X-ray beam on the devices, during XPS measurements. Lastly, in section 6.3.4, the experimental methods employed to reduce the parasitic bias drops in the Si substrate will be shown.

In section 6.4, XPS under bias technique will be validated. This involves the development of a procedure to fit the XPS signals from the different layers of the HKMG gate stack, in section 6.4.1. Next, in section 6.4.2, the binding energy values obtained from fitting the XPS signals will be compared to the values obtained with electrical modelling.

Lastly, in section 6.5, XPS under bias technique will be used to localize and quantify dipoles related to the addition of La or Al, and dipoles occuring due to TiN gate thickness modulation. In section 6.5.2, the methodology to use XPS measurements at zero bias for estimating the values at the flatband condition will be presented.

6.1 State of the art and the need for XPS under bias

The development of CMOS devices needs effective tools to compare various technologies with specific architectures. Their characterization relies on the measurement of various parameters, such as the effective workfunction (WF_{eff}) or the flat band voltage (V_{FB}), the effective oxide thickness (EOT), and the various charges and dipoles present in the gate stack. In section 2.1.1, we have seen that EOT, WF_{eff} and silicon surface potential and its charge, can be extracted from capacitance voltage C(V) measurements. WF_{eff} is the parameter that depends on the difference between metal and semiconductor Fermi levels. It is affected by, and contains information about, the fixed charges present in the gate oxide, dipoles at various HKMG interfaces and interface states (D_{it}) (see Eq. 1.13). Despite the fact that WF_{eff} contains combined information about these parameters, calculating their individual values and localizing them inside the gate stack is complicated.

Although in Figure 2.5, it was shown that the amount of fixed charges and dipoles can be obtained by C(V) measurements using a bevel interlayer, however, it was possible only with the

prior knowledge that the dipoles are present only at the high-k/SiO_2 interface and fixed charges only at the SiO_2/Si interface. Moreover, this analysis requires measurements on a large number of devices. In addition, obtaining D_{it} information from C(V) measurements at very low frequency can get complicated, due to measurement difficulties faced at low frequencies.

The information of charges, dipoles and interface states inside the HKMG stack can be obtained through the knowledge of relative band energies of the gate stack layers, which can not be obtained from C(V) measurements. Band energy diagram become even more complicated in the case of a HKMG stack, as it contains many ultra-thin layers and their interfaces which can have different dipoles and charges. Moreover, process development involves changing gate stack materials, process conditions and introducing additives that can change the relative band energy position in the gate stack, leading to a change in the leakage current, WF_{eff} and the threshold voltage (V_T). For this purpose, the introduction of Lanthanum (La) and Aluminum (Al) additives [17] [47], forming dipoles at the SiO_2/high-k interface, has been proposed. They can also lead to interface states (D_{it}) at the Si/SiO_2 interface. Lastly, metal gate thickness variation can also result in WF_{eff} modulation [37][164] [165] [166]. Thus, in order to build a full understanding of the CMOS devices and for its further development, a comprehensive knowledge of band-energies of different layers in the gate stack is required.

In contrast, X-ray Photoelectron Spectroscopy (XPS) (see section 2.2.1) is a technique that is able to measure the core levels (CL) and valence bands (VB) binding energies (BE) of different materials within its analysis depth of around 10 nm. XPS, when applied on a HKMG stack, can give information about the relative core levels and hence on the band energy levels of the different layers composing the gate stack. Moreover, as stated in section 2.2.1, CL BE's depend on the chemical state of a material and its electrostatic potential [131], and hence XPS can also be used to identify bias drops inside the gate stack. Its high elemental and chemical specificity makes it extremely useful for analyzing device gate stack. Hereafter, we will review the studies that has been conducted and different experimental difficulties associated with XPS analysis.

Fulton et al.[167] analyzed a Zr/ZrO_2/SiO_2/Si HKMG stack, with combined UPS (ultraviolet photoelectron spectroscopy) and XPS measurements, after deposition of every gate stack layer (starting from the Si substrate) in a stepwise manner. XPS was used to mesure the bias drops in the gate stack layers, and UPS to measure the valence band offsets and vaccum level position. From these values, the authors calculated the values of dipoles formed at each interface of the gate stack, due to a charge transfer across the interface. However, as the measurements are done after deposition of each layer, it does not take into account the effect of this deposition on the previously deposited stack. Indeed, deposition of each layer can introduce charges and dipole at the interfaces or inside the different layers of the already present stack. Thus, this analysis of interface dipoles and valence band offsets does not represent the band energies of the final HKMG stack. Moreover, as the device was not biased during the XPS measurements, the calculated values can be influenced by the charging effects explained below.

Charging effects: These effects are one of the principle intrinsic problems associated with the conventional XPS, in determining the energy positions of the gate stack layers. Generally gate oxides contain a finite number of charges. Moreover, electrons emitted during an XPS measurement makes the oxide positively charged. Both these effects result in the band bending inside the oxide and in the Si substrate, and hence causes a shift in the measured CL's. This might

be the reason for variations in CL positions measured by many researchers. In a previous study, this problem was solved by Chikata et al. [168] by biasing the gate of the devices to ground, together with the Si substrate. This solution allowed to take care of the charging effects and hence controlling the bias at the metal gate and its interface with the gate oxide, and so provided accurate BE values for the oxide. The effect of charging on the CL's and the control of the CL shift by grounding the metal gate, is shown in Figure 6.1 [168]. In this study, the authors made separate measurements on different high-k layers and SiO_2, and calculated their conduction band and valence bands energies to form the band energies of the entire HKMG stack. One drawback of this work arises due to the fact that these measurements were conducted on individual layers of the gate stack and not on a complete gate stack, and thus the results does not represent the real device that will include interface effects and band bending.

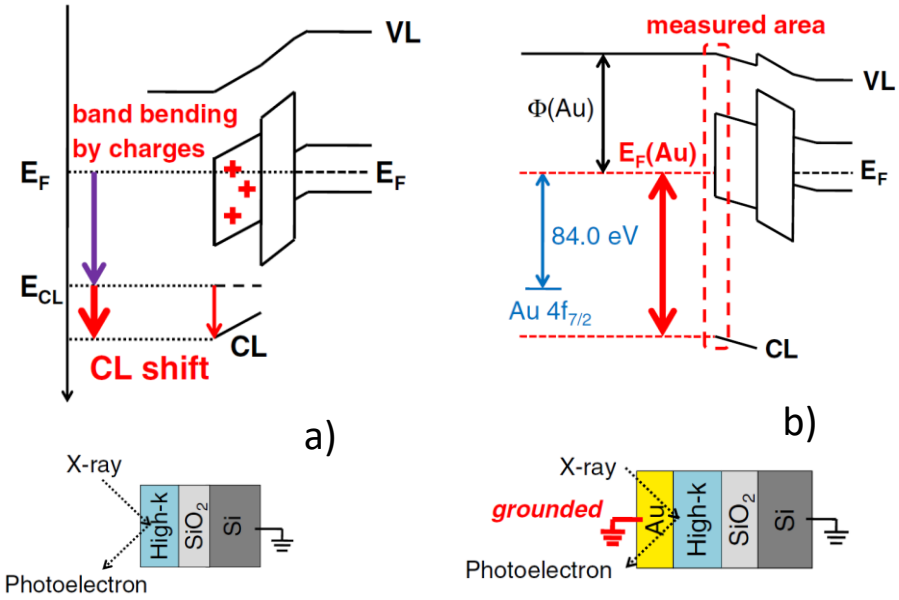


Figure 6.1: Comparison of energy bands for gate stacks a) without a metal gate and b) with a grounded metal gate [168]

While the charging effects can be solved by grounding the metal gate, a non-zero bias must be applied in order to obtain key information from XPS under bias results, of the HKMG stack. The interest of applying a non-zero bias comes from the following points:

- The best condition to compare energy bands of different gate stacks is the flatband condition (shown in Figure 1.3b). Indeed at the flatband condition, the BE signal from a specific gate stack layer will be the same for its entire thickness, due to the bands being flat, and not an average over the thickness. This would make the task of comparing the band energies of different gate stacks, accurate and much less complicated.

- The study of BE shifts with the applied bias, can enable us to obtain the profile of interface states. In a previous study, Yamashita et al. made measurements on devices with HKMG stacks to determine the density of interface states at the Si/SiO_2 interface, by studying the BE shift of XPS spectra under bias [169].

- In order to show that the BE's obtained from XPS under bias technique can be used to represent the MOS gate stack energy bands, CL BE shifts of different gate stack layers with bias must be shown to correlate with the expected bias drops in these gate stacks, obtained from electrical measurements. To achieve this, a range of bias has to be applied to the device, as will be done in detail in section 6.4.3.

However, knowing the effective bias across the MOS device can get very complicated. This complication arises due to a difference between the applied bias and the effective bias occuring over the MOS gate stack, due to parasitic bias drops. This would lead to an unkown biasing across the device and hence might lead to a wrong analysis of the band energy results obtained from XPS measurements. An example of this kind of improper analysis was obtained, we think, in the work of Yamashita et al. [170], due to parasitic bias drops, leading to conficting results as shown in Figure 6.2. In this work, the authors have analysed the HKMG stack through the XPS CL BE shifts with bias. The goal was to compare gate stacks with two different metal gates. The authors concluded that a potential drop occurred at the Pt/HfO_2 interface, and this voltage drop varies as a function of the applied voltage. Such an interpretation is difficult to understand because it implies some charge variation at the high-k/metal gate interface. Usually the state of charge at one interface is defined by their relative position to the adjacent Fermi level, and there are no reason that it would depend more on the Fermi level position of a remote electrode as suggested in Figure 6.2.

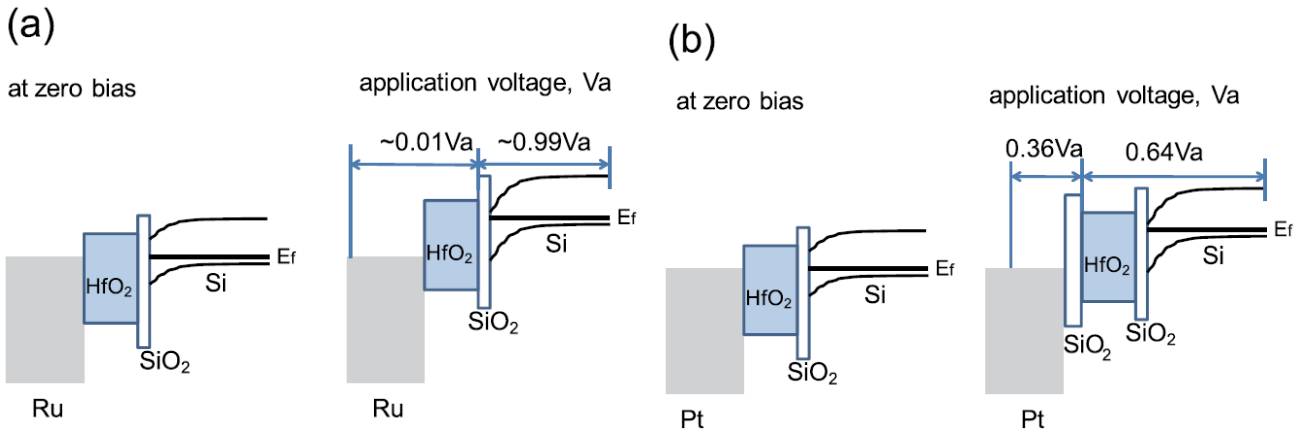


Figure 6.2: Band diagram for the a) Ru/HfO_2/SiO_2/Si stack and (b) Pt/HfO_2/SiO_2/Si stack, at zero bias and a higher bias [170]

The issue of estimating the effective biasing across the MOS device is the main concern for us, and will be further discussed, solutions will be presented in the next section.

6.2 Test structures and their biasing issues

6.2.1 Test structures and their process flow

A basic requirement to perform XPS under bias measurements is the realization of specific test structures, that will allow to bias the device with simultaneous XPS measurements. For this, MOS capacitors were fabricated with a process flow quite similar to the one introduced in section 1.3.2. We discuss it in Figure 6.3, it consists of a 500 nm Silicon Trench Isolation (STI) SiO_2 oxide deposited on a N-type or P-type Si substrate. Then, cavities are patterned by photolithography and etching, defining the active area for the MOS capacitances. Then, HKMG stack is deposited, consisting of 1.2 nm SiO_2 interlayer (IL), HfSiON or HfON as the high-k and the sacrificial gate stack deposition by the approach of Figure 1.18, in order to introduce La or Al additives. The sacrificial gate is removed and finally a 5 nm TiN and poly-Si films are deposited, followed by source drain (S-D) anneal. Gate patterning is done by photolithography to isolate individual capacitors. As the depth of XPS beam analysis is less than 10 nm, the poly-Si layer is removed in order to analyze all the gate stack. As this poly-Si have endured the S-D anneal, it could not be removed by the chemical solutions that are normally implemented. Thus, a new wet etching solution was developed, that could remove the poly-Si layer after the S-D anneal.

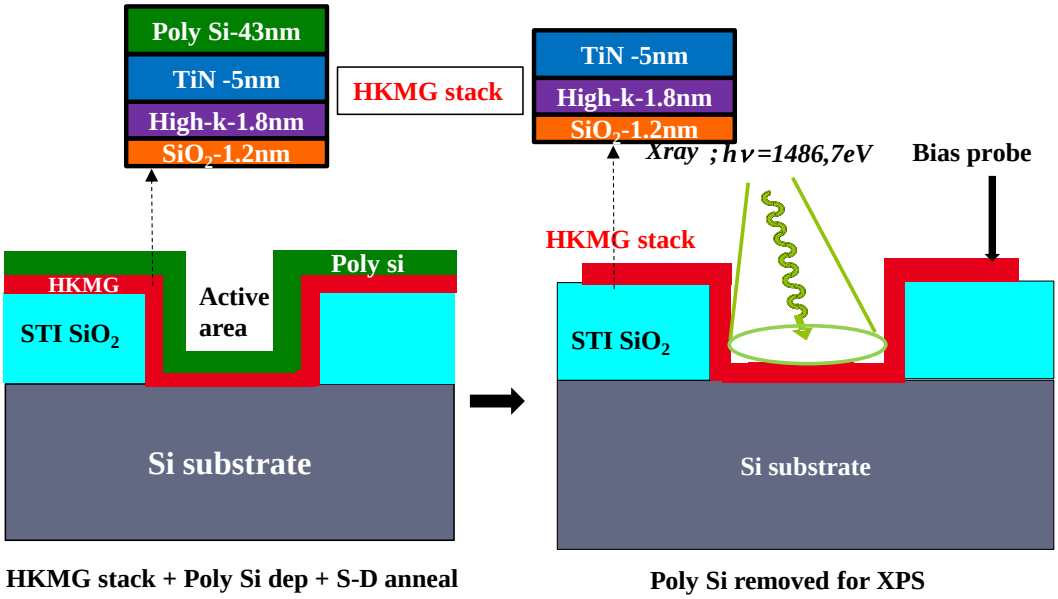


Figure 6.3: Process flow of MOS capacitors used for XPS under bias studies

Inside the cavity (called the active area), the HKMG stack is in contact with the Si substrate and hence forms a MOS capacitor, which is termed as the device under test (DUT) or device. The advantage of this structure is that while the XPS measurements are performed inside the cavity, the device can be biased outside the cavity on the thick SiO_2 part. The thick SiO_2 insulating layer makes it possible to electrically isolate the cavities from each other and to electrically polarize the structures reliably and reproducibly, without damaging the MOS device located at the bottom of the cavities.

6.2.2 Biasing issues in the test structures

Despite having several advantages for XPS under bias, these structures still face a few challenges. A first difficulty comes from the need to illuminate, with the X-ray beam, only the cavity area (or DUT). Indeed, if some part of the X-ray beam falls on the thick SiO_2 area, XPS measurements would be altered by the response of the signal from the high-k and SiO_2 layers which are not part of the MOS device. This would lead to a distortion of the XPS spectrum and even to a shift in BE of the peaks. The second difficulty with these structures consists of a precise contol of the voltage applied to each of the different layers comprising the gate stack. In other words, when a voltage Vg is applied to a structure of sufficiently large size to allow XPS characterization, the distribution of voltage in the different layers is not known. Figure 6.4 schematically illustrates the biasing of the device, where voltage Vg is applied to the gate on the thick oxide part and the backside Si substrate is grounded. It shows that the device bias (V_{DUT}) occuring across what we consider as the DUT, comprising of the HKMG stack and a part of the Si layer at the SiO_2/Si interface and having a thickness of the order of 1 µm (depletion region), is different from the applied bias Vg.

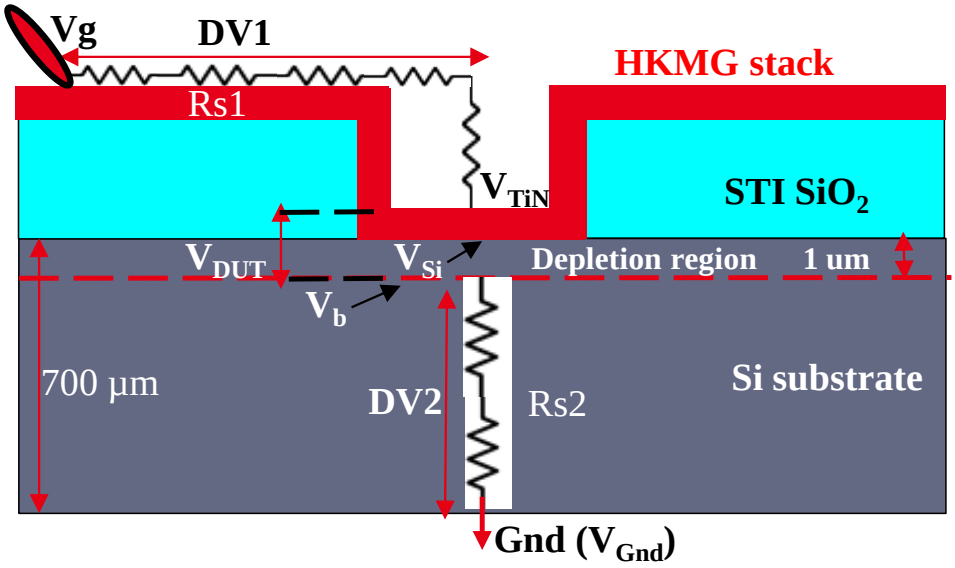


Figure 6.4: Description of serial resistances and bias drops present in the test structure used for XPS under bias

The difference between Vg and V_{DUT} (between gate bias V_{TiN} and substrate bias V_b) must be accounted for, to know the exact device biasing, that would allow to obtain accurate BE for different layers. This difference DV (Vg - V_{DUT}) or parasitic bias drop is a combination of bias drops DV1 and DV2, occuring due to series resistances of the metal gate (Rs1) and the bulk Si substrate (Rs2) respectively. The basic equations for this can be written as:

$$DV1 = Vg - V_{TiN} \; ; \; DV2 = V_b - V_{Gnd} \; ; \; V_{DUT} = V_{TiN} - V_b \qquad 6.1$$

XPS under bias requires the use of relatively bigger devices (bigger cavity), which has the advantage of a better focusing of the XPS beam inside it, and thus avoiding an error signal from outside the cavity. Moreover, for a bigger cavity, a larger beam size could be used, which would

increase the XPS signal and thus significantly reducing the measurement time. Despite having these advantages for XPS measurements, bigger devices suffer from a large substrate serial resistance (Rs2) effect, that makes the parasitic voltage drop DV2 even more significant in these large devices. The impact of Rs2 on the bias drop DV2 increases as the device size increases. An evidence of the increase of DV2 with the device size can be seen in Figure 6.5, where the leakage current density J(V) decreases as the size increases due to an increase of DV2. In turn, this effect decreases the V_{DUT} (V_{DUT} = Vg - DV1 - DV2).

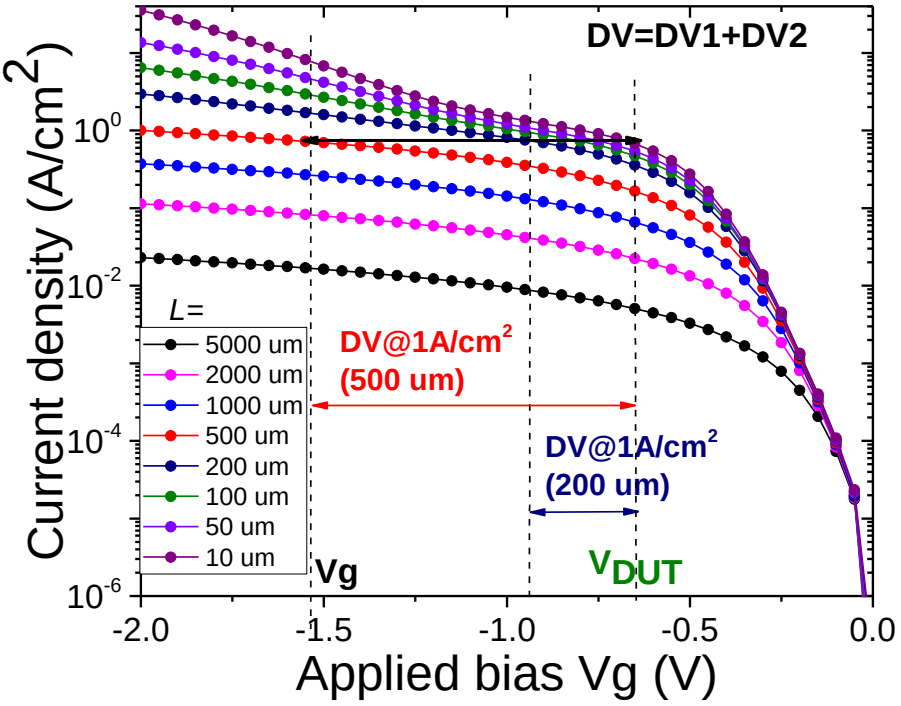


Figure 6.5: Leakage current density versus applied bias for different square capacitances of edge size L, a same current density corresponds to a same device bias V_{DUT}

An ideal condition for band energies analysis of a MOS device is the flat band condition requiring a device biasing at V_{FB}, shown in Figure 1.3b. C(V) characteristics can be used to identify V_{FB}, as well as bias drops in the different gate stack layers. This includes the dielectric bias drop V_{OX} and silicon substrate bias drop V_{Si}, that we can estimate by using the Poisson-Schrödinger model introduced in section 2.1.1. However on bigger devices, like the leakage current density, a large substrate serial resistance effect also impacts the C(V) measurements [171]. Figure 6.6 shows the capacitance degradation as the device size increases. Thus, exact biasing of the device (control of V_{DUT}) and analysis of biasing inside the gate stack from C(V) become great issues.

Parasitic effects DV1 and DV2 on smaller devices (less than about 10 µm) are negligible due to a smaller leakage current, inducing a smaller bias drop in the serial resistances. This would mean that the applied bias Vg is equal to V_{DUT} and negligible degradation on C-V characteristics will take place (as seen on Figure 6.5 and Figure 6.6). From an electrical point of view small devices seem almost perfect, but they can not be used for XPS measurements as the XPS beam size is much bigger than their size.

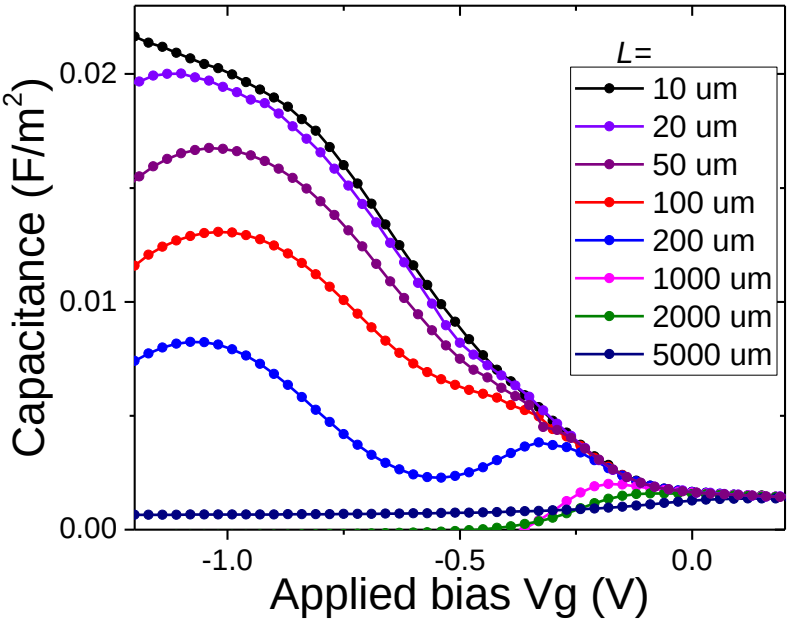


Figure 6.6: C-V measurements showing capacitance degradation with device size (L is the edge size) due to serial resistance

6.2.3 Biasing solutions and electrical modelling

In order to overcome these challenges, inherent to both bigger and smaller devices, a set of test structures with different cavity (device) sizes has been utilized in this work, with a combination of electrical modelling on smaller devices and XPS measurements on bigger ones [PK-3] [PK-5]. Figure 6.7 shows the top view of the test structures (maskset) used in this work, where test structures with different size of square cavities are represented, ranging from 1*1 μm^2 to 5*5 mm^2.

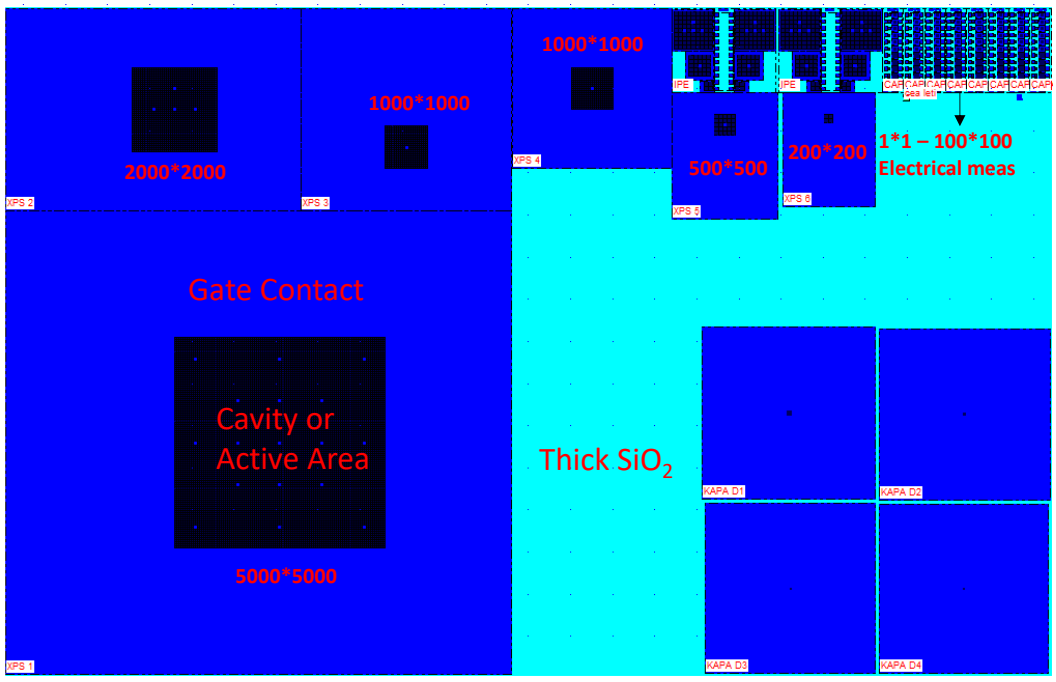


Figure 6.7: Top view of the test structures with various device sizes, used for XPS and electrical measurements

Figure 6.7 shows the top view of the test structures that were introduced in Figure 6.3 (in cross section view). The black part represents the cavity (device) and the blue part is the gate contact outside the cavity on thick STI SiO_2, on which the bias will be applied. The cyan color indicates the thick STI SiO_2 without any metal contact, that is isolating the different test structures. There are mainly two types of devices that are relevant for our studies:

- First are the devices for XPS measurements (or XPS device), with big square cavities ranging from 200 um to 5 mm in edge length. The area of the gate contact is large enough to allow a good contact during device biasing in an XPS equipment. Indeed the probe size used in an XPS equipment is quite large, in the mm range.

- Second are the devices for electrical characterization (C(V) and I(V)), with relatively smaller square cavities ranging from 1 μm to 100 μm in edge length. The gate contact area for these devices is large enough to make a contact during measurements on an electrical bench.

Next, using a combination of bigger and smaller devices, we will present a strategy for XPS under bias measurements. This involves the use of the smaller devices, for electrical measurements to determine V_{FB} and device biasing (V_{DUT}, DV1, DV2), and the bigger devices for XPS under bias measurements. The detailed procedure is described below:

- **V_{FB} calculation**: First, on smaller devices of area 10*10 μm^2, leakage current density measurements $J_S(V)$ and C(V) measurements are conducted on an electrical characterization bench. Then, from the C(V) characteristics, flatband voltage is extracted by Poisson Schrödinger quantum simulations. As the bias drops DV1 and DV2 are negligible on this smaller device, $J_S(V)$ will be the ideal leakage current density characteristic of the HKMG stack.

- **DV calculation**: On bigger devices (or XPS devices), and on the same electrical bench, leakage current density measurements $J_B(V)$ are conducted. As the gate stack is the same on big or small devices, the bias across them (or V_{DUT}) will be the same on both the devices at the same leakage current density. Thus V_{DUT} on bigger devices can be identified by comparing J_B with J_S, like has been done in Figure 6.5. In order to be at a certain current density, a bias Vg^B must be applied to the bigger device, which is higher than the Vg^S bias leading to the same current density on the smaller device. Then, by comparison at this same current density, Vg^S would be equal to the device bias V_{DUT}^B for the bigger device, i.e. $V_{DUT}^B = Vg^S$. Hence the total parasitic bias drop for the bigger device can be calculated as $DV^B = Vg^B - V_{DUT}^B$. In this way, for a bigger device, the device bias V_{DUT} and the total parasitic drop DV (= DV1+DV2) can be calculated for any applied bias Vg and leakage current.

- **DV1 and DV2 calculations**: During DV calculation on the bigger device, DV1 can also be calculated individually by an additional measurement. Figure 6.8 shows the schematic of such measurement on a 200*200 μm^2 cavity test structure, in cross sectional and top view. A bias Vg is applied on the metal contact over the thick SiO_2 and the bias is measured at many points very close to the active area (in black) boundary (in white), but still on the

thick SiO_2 part. Then an average is taken over these points to calculate the bias at the active area edge or V_{TiN}. From the measurement of V_{TiN}, DV1 and DV2 can be calculated by Eq. 6.2 below:

$$DV1 = Vg - V_{TiN} \; ; \; DV2 = DV - DV1 \qquad 6.2$$

In this way, as shown in Figure 6.9, we can calculate for the bigger devices the V_{DUT}, V_{Ti}, DV, DV1 and DV2 for each value of leakage current,.

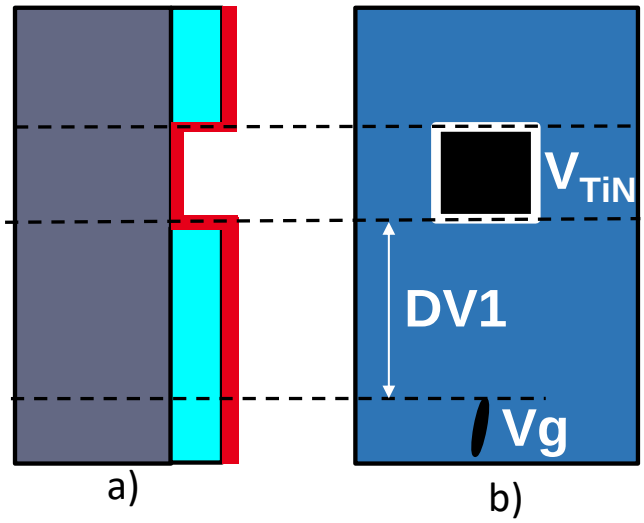


Figure 6.8: A XPS test structure in a) cross-section view and b) top view, and the method to calculate DV1

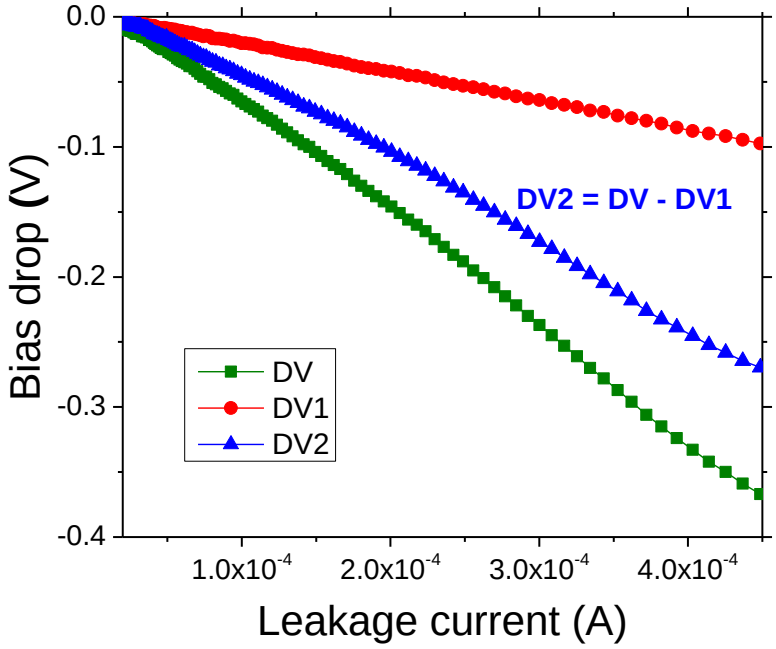


Figure 6.9: Bias drop evaluations in the gate (DV1) and the Si substrate (DV2)

The presence of substrate bias drop DV2 in a bigger device (see Figure 6.4 and Figure 6.9) will shift the band energies of the entire gate stack by an amount DV2. Figure 6.10 compares the band energy diagrams, at the flatband condition, for a 10 µm device (ideal device) and a XPS device having a larger size and therefore a larger substrate bias drop DV2. As a reminder, the 10 µm device have a negligible substrate bias drop DV2 and therefore the Fermi level at the back side EF_{BS}, which is grounded to the XPS spectrometer, is also the Fermi level at the silicon surface. On the contrary, in case of a bigger device, the pseudo Fermi level of holes follows the bias drop DV2. For the gate stack under test or DUT, the CL is still referenced to the Fermi level at the silicon surface EF_{Si}, leading to the E_{CL} energy level. But for the spectrometer, the same CL will be measured with reference to the EF_{BS}, leading to the E_{raw} measured. In order to understand the band diagrams from an XPS point of view, we must keep in mind that the binding energies are increasing as we go to deeper core levels (contrary to what is done for classical analysis in semiconductor physics). Thus, EF_{Si} - EF_{BS} is equal to the bias drop in the substrate, leading to the relation between E_{CL} and E_{raw} given in Eq. 6.3.

$$E_{corrected} = E_{CL} = E_{raw} - DV2 \qquad 6.3$$

More generally, the above equation will account for the effect of any bias drop in the gate stack, between real core level that we want to measure (E_{CL}) and the measured CL (E_{raw}).

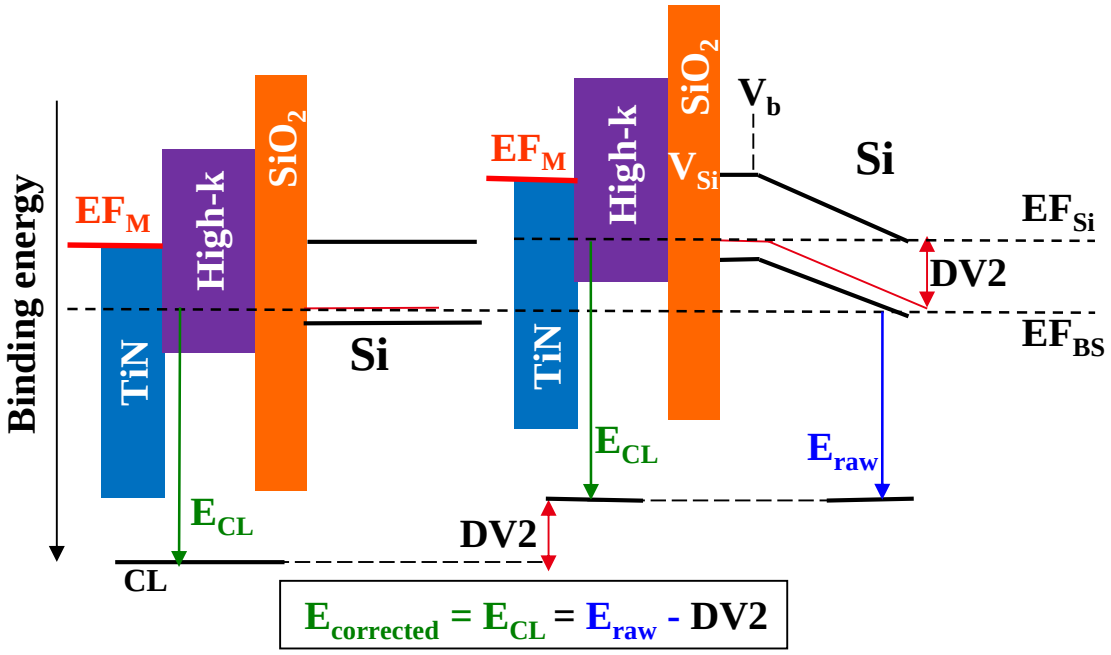


Figure 6.10: Effect of substrate bias drop DV2 on energy bands at flatband condition during XPS under bias measurements

- **XPS measurements**: XPS is performed inside the cavity of a XPS device by using a Versaprobe II spectrometer from ULVAC-PHI equipped with a 4-contacts sample holder, as depicted in Figure 6.11. A copper film is placed on the sample holder and the sample which contains the test structure, is mounted on it (Figure 6.11b). Then, the test structure is biased by contacting two probe pins over the metal contact, one to apply the bias and the other to measure and check the bias. Two other pins are placed on the copper film, one to apply the ground potential and the other to measure and check this potential.

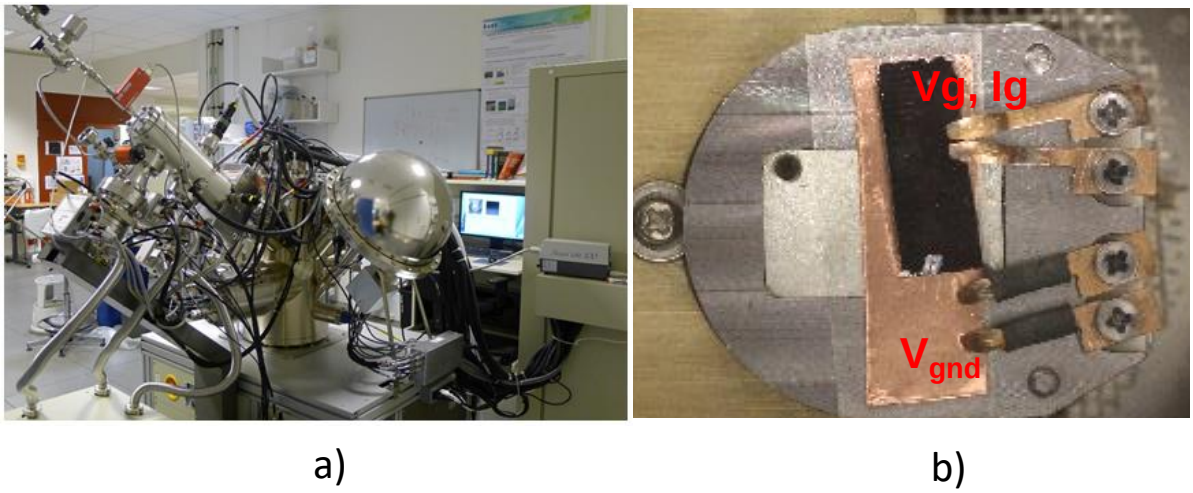


Figure 6.11: a) Versaprobe II spectrometer used to perform XPS under bias measurements and b) test structure connection on a 4 probe sample holder, used inside the XPS equipment

As a reminder, the relationship between the V_{DUT} and leakage current for a bigger device has been calculated using the characteristic J(V) on a smaller device. During a XPS under bias measurement, control of the biasing can be performed by two ways: 1) Applying a constant leakage current value correponding to a particular V_{DUT} across the gate stack or, 2) Applying a constant Vg, and the V_{DUT} value is calculated by measuring the leakage current during the XPS measurement.

6.3 Issues related to the XPS equipment

In the previous section, we have solved the biasing issues and have developed a strategy to bias the devices at any desired band energy condition of the gate stack. Before this technique can be applied to make XPS under bias measurements on the devices, several checks must be made and issues related to the XPS equipment should be solved. This will make this techique more robust and reliable, these issues are discussed below:

6.3.1 Device location

Inside the XPS equipment, the sample can be observed with the help of a secondary electron imaging. In normal cases, this image is used to localize the relevant area (device) on the sample, where XPS analysis would be done. In our case, the relevant area is the cavity containing the MOS

gate stack. Secondary electron imaging contrast corresponds to the chemical composition of the layer being analyzed, and this forms the basis to differentiate between various areas on the sample. In our test structures, the TiN layer on top is over the cavity, as well as outside (Figure 6.3). Thus, the cavity cannot be differentiated from the TiN contact area, this issue can be seen from the secondary electron image in Figure 6.12. Therefore, this becomes a big experimental hindrance in carrying out precise XPS measurements, when the device of interest cannot be located.

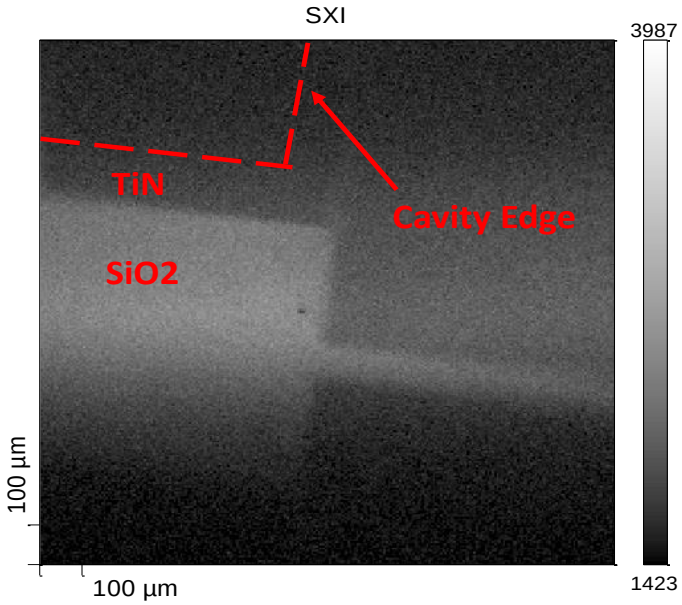


Figure 6.12: Secondary electron imaging of the test structures, the cavity area cannot be seen, and only the border between thick SiO_2 and TiN contact can be seen

We have used two different ways to solve the problem of device location:

- First, as we see in Figure 6.12, the border between thick STI SiO_2 and the TiN metal contact can be clearly seen. The distance of the cavity area from the TiN contact/thick SiO_2 border is known from the maskset or test structure dimension of Figure 6.7. Thus, by locating the TiN/SiO_2 border from the secondary electron image, the location of the cavity can be determined and XPS analysis can be conducted. The precision of this method depends on the resolution of the TiN/SiO_2 border, meaning on the border thickness in the image of Figure 6.12. This thickness was calculated on many samples and the maximum observed value was < 20µm, which is quite good considering the size of the cavity.

- A complimentary method can be implemented to localize the cavity area by doing XPS anaysis of the suspected area and analysing its Si signal. Si substrate signal (Si^0) will only be observed from inside the cavity, and on the TiN contact part it will be masked by the thick SiO_2 layer. Thus, by analyzing the Si substrate signal, the cavity can be located in some way. Figure 6.13 shows the total Si signal for XPS measurements done inside and outside the cavity, and the Si substrate signal at a BE of around 99 eV disappears outisde the cavity.

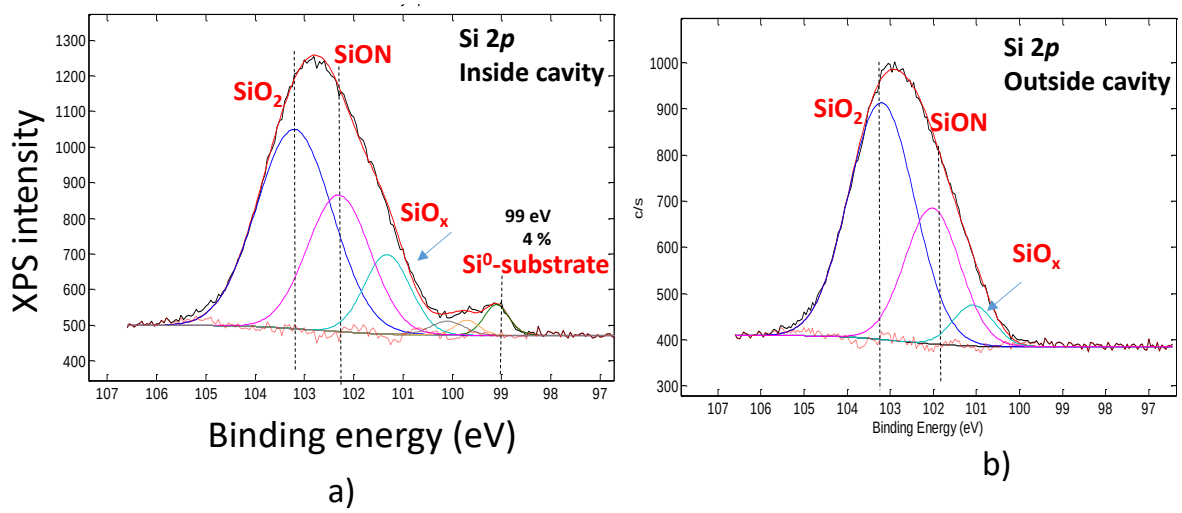


Figure 6.13: Si-2p XPS spectra of the test structures a) inside the cavity and b) outside the cavity

As the Si substrate signal is very low, the second method is not quantifiable, meaning that the area of the Si substrate peak can not be used to know if the XPS beam was partially outside the cavity. Therefore, in this work, we have used only the first method. Apart from device location issues and its errors, there can be an error in the size and position of the X-ray beam in space and/or time, which is discussed in the next section.

6.3.2 XPS beam position error and size estimation

In order to perform XPS measurements with minimum error, the XPS beam must be completely contained inside the cavity. If a part of the X-ray beam overlaps outside the cavity, SiO_2 and Hf signal from that area can influence the main signal from the cavity. Thus, the precision in the position and size of the X-ray beam spot must be checked. The X-ray spot can be shifted in distance, and its position can also change with time as well. Certain measurements were conducted to check this point, such as: XPS analysis of the same area at two different times, and XPS analysis in a known area on the sample. These measurements showed that the XPS beam is quite precise in terms of position and it does not shift with time.

The X-ray beam has a gaussian shape, with its intensity being maximum at the center and decreasing with the distance from the center, as shown in Figure 6.14a. Normally, in the equipment manuals, the data on beam size (theoretical) does not represent the total size of the beam but rather a value related to its full width at half maximum. In order to check the total or real size of the X-ray beam, linescan measurements were conduced on an area containing both TiN and thick SiO_2 parts, as shown in Figure 6.14b. Starting from the SiO_2 part and moving towards the TiN part, XPS measurements were conducted at many points with a theoretical beam size of 50 and 100 µm, and Ti signal was analysed. This Ti signal is plotted in Figure 6.15, with the distance from the starting point in SiO_2. The distance between the Ti signal maximum and minimum will be equal to the total size of the X-ray beam. From this analysis, its is clear that the real size of the beam is almost twice of its expected size. The real beam diameter is about 100 µm and 200 µm, for beams having theoretical diameter of 50 µm and 100 µm respectively.

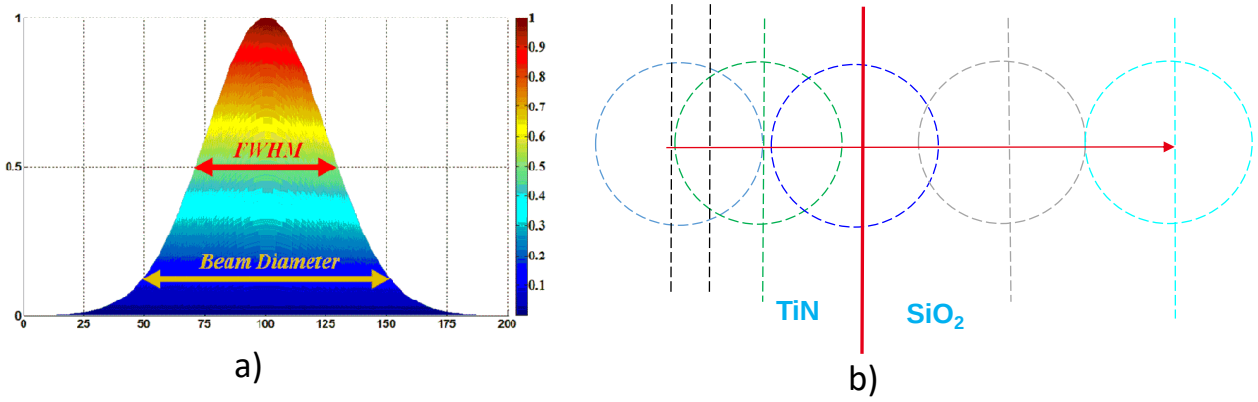


Figure 6.14: a) Gaussian profile of an X-ray beam [172] and b) method used to obtain its total size

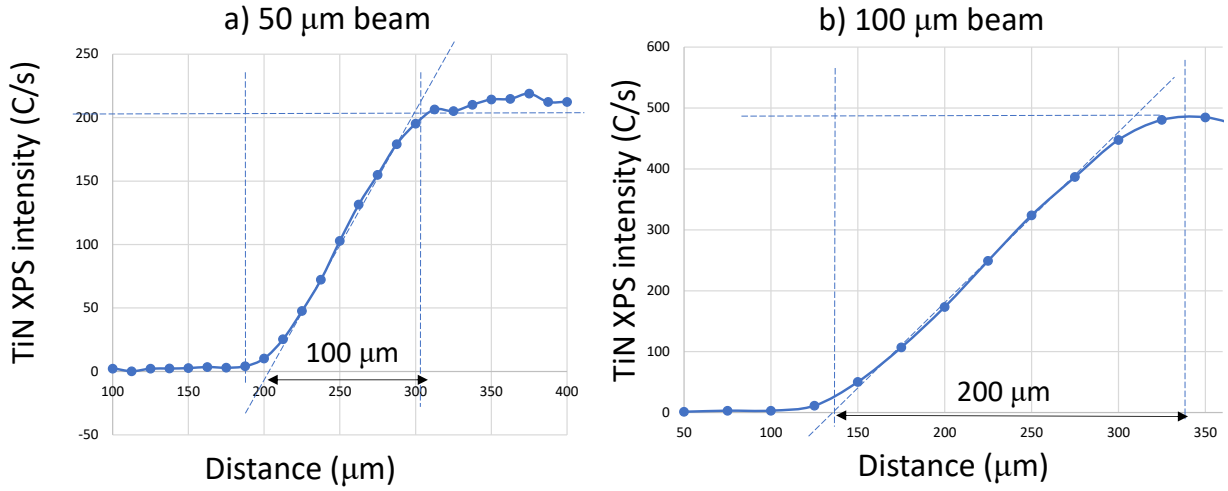


Figure 6.15: Line scan results showing the calculation of the total X-ray beam diameters, with theoretical diameters of a) 50 um and b) 100 um

6.3.3 Impact of X-ray beam on the gate stack

The X-ray beam used during XPS measurements interacts with the layers of the HKMG stack. Thus, it could change their properties and thereby cause an increase in the device leakage current. An increase in the leakage current of an XPS device would mean that its gate stack is not the same as that of a smaller device used for electrical measurements. Thus, the method presented earlier to calculate V_{DUT} of an XPS device, by comparing its leakage current density with that of a smaller device, will need to be modified. In order to check this point, leakage current measurements I(V) were done with the sample mounted inside the XPS equipment, with or without the use of the XPS analysis beam. Figure 6.16 compares these two I(V) curves, and it is clear that the gate leakage current characteristic is independent of the use of the X-ray beam. This demonstrates the validity of the use of measured leakage current during XPS measurements to identify the V_{DUT}.

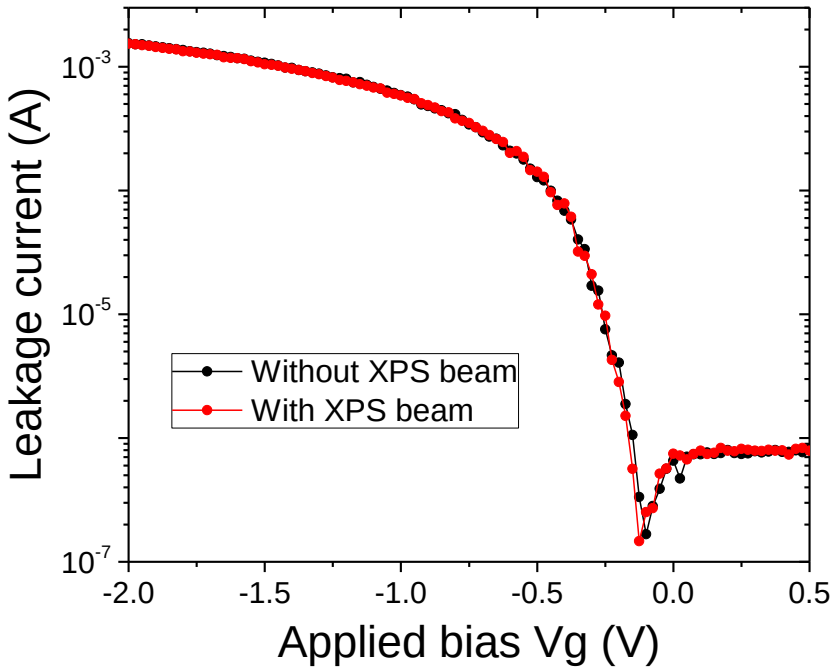


Figure 6.16: Leakage current versus applied bias for device connected in the XPS equipment, with or without illumination with an X-ray beam

6.3.4 Reduction of substrate resistance

It has been stated earlier, that parasitic bias drop DV2, caused by the resistance in the Si substrate, gives rise to biasing issues and band energy shifts, and must be as low as possible. The DV values presented in Figure 6.9 are actually for samples on which a specific treatment was conducted on their back side. Intially on the raw samples, the value of parasitic substrate serial resistance (Rs2) was very high (>5000 Ω), instead of a usual value of < 500 Ω, which resulted in such a large value of DV2 that it would be impossible to perform reliable XPS under bias measurements. For normal electrical tests, the impact of this backside resistance is not a problem, because much smaller device are used and thus have a smaller leakage current. However, for XPS devices (> 200 μm) this became a major problem. The exact reason for this large back side resistance is still unknown, but it could result from a physical modification of the substrate backside induced during a step in the device process flow.

Thus, in order to reduce the backside resistance, significant experimental efforts were made which resulted in resistance reduction by more than a decade. In order to achieve this, first the samples were etched by hand using a diamond stylus to remove any films possible with it. In order to further remove other films, not possible by diamond etch, the samples were etched by an Ar plasma sputtering. Then, a stack of highly conductive metal layer of Ni/Au was deposited on the back side, allowing a good electrical contact of the sample with the ground connection. Figure 6.17 shows the total bias drop DV2 versus the device leakage current, for a raw sample and after the process described above (termed as gold), for 200 μm and 500 μm devices. A significant decrease in DV is obtained, from about 0.9 V to about 0.3 V.

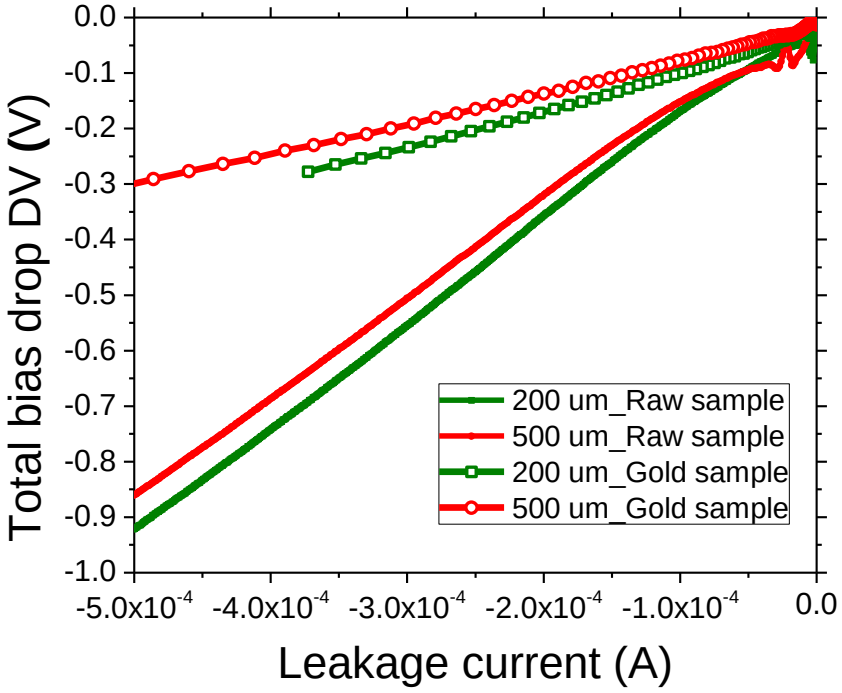


Figure 6.17: Total parasitic bias drop before and after Ar etching and Au/Ni stack deposition on the sample back side, for 200 μm and 500 μm devices

6.4 XPS under bias technique validation

In the previous sections we have been able to solve the electrical biasing issues, by defining a methodology to calculate effective bias across the device and bias drops in the metal gate and Si substrate. Further, we have addressed the issues related to the XPS equipment in terms of X-ray beam size estimation, device location and decrease of substrate resistance. In this section, XPS measurements and electrical modelling will be used simultaneously to validate XPS under bias technique and to improve its use [PK-3] [PK-5].

Analysis of the MOS device by electrical measurements, such as C(V) and I(V), to calculate Si surface potential (in section 2.1.1) and the potential drop in the metal gate (DV1 and V_{Ti} in section 6.2.3) have been well established. It is the first time that the XPS under bias technique would be used in this manner, using this new measurement strategy presented in the above sections. To validate the accuracy of this technique, we must prove that the different CL BE shifts in the various gate stack layers correspond to the bias shifts expected from MOS gate stack simulations.

6.4.1 XPS spectra fitting methodology

For these studies, XPS under bias measurements were performed on a 200*200 μm^2 device. C(V) measurements, performed on a 10*10 μm^2 device, allowed to obtain the flatband voltage. Moreover, combination of I(V) measurements performed on both these devices provided the relationship between the leakage current value and V_{DUT}, V_{Ti}, DV1 and DV2 for the 200 μm device, with the method presented in section 6.2. A X-ray beam with a theoretical diameter of 100 μm and real diameter of about 200 μm (Figure 6.15b) was used for XPS measurements. The device size and the X-ray beam size were chosen by taking into account the serial resistance effect and the power of the beam. In other words, a 200 μm device size was chosen to minimize the substrate serial resistance effect (DV2 value) and that the X-ray beam can be contained completely inside the cavity.

The device gate stack used here is as shown in Figure 6.3, with the HKMG stack consisting of the SiO_2 interlayer, HfSiON as high-k and 5 nm TiN, deposited on a P-type Si substrate inside the cavity. This device does not contain any La or Al additives. The sample, containing the 200 μm device was mounted on the 4-probe sample holder, as shown in Figure 6.11, and placed inside the XPS equipment. The device was biased at different voltages between 0 and -2V and XPS acquisition were conducted for about 6 hours, for each bias. XPS signal of Ti-2p, Hf-4f and Si-2p were acquired from inside the cavity. These three signals originate from different gate stack layers and can correspond to various chemical compounds of Ti, Hf or Si elements. Each of these compounds will have several peaks in binding energy depending on the elements oxidation state, bonding atoms involved and doublets due to spin orbit coupling, as described in section 2.2.1. Thus, each signal is fitted with the convolution of its compounds, and their doublets. From these three signals, BE of one of the compounds was selected to represent each gate stack layer.

Figure 6.18, Figure 6.19 and Figure 6.20 shows the acquired Ti-2p, Hf-4f and Si-2p signal respectively, obtained on a device biased at 0 V. In Figure 6.18, Ti is present only in the metal gate, and so the entire signal corresponds to it. This Ti signal is well fitted by using a Shirley background, as well as three components: TiN, TiON and TiO_2 and their respective doublets, for all our analysis. The binding energy difference between the compounds and their doublets, as well as area ratio between them was taken from literature and were fixed (given in Annex 1). The TiN peak has the largest area amongst the three peaks, which shows that it is present in the maximum amount inside the metal gate. Therefore in our studies, BE of the TiN-2p peak will be used as the BE reference for the metal gate.

In Figure 6.19, Hf-4f XPS signal originates only from the high-k layer. Similar to Ti signal, the Hf signal is fitted using a Shirley background and three specific components: HfOSi, HfO_2 and HfN and their respective doublets. As explained before, difference in binding energy from one oxidation state to another and between their doublets, were kept constant (Annex 1). Therefore, any of the peaks can be used as the BE reference for the high-k. For our studies, we decided to choose the HfO_2-4f peak BE to represent the high-k.

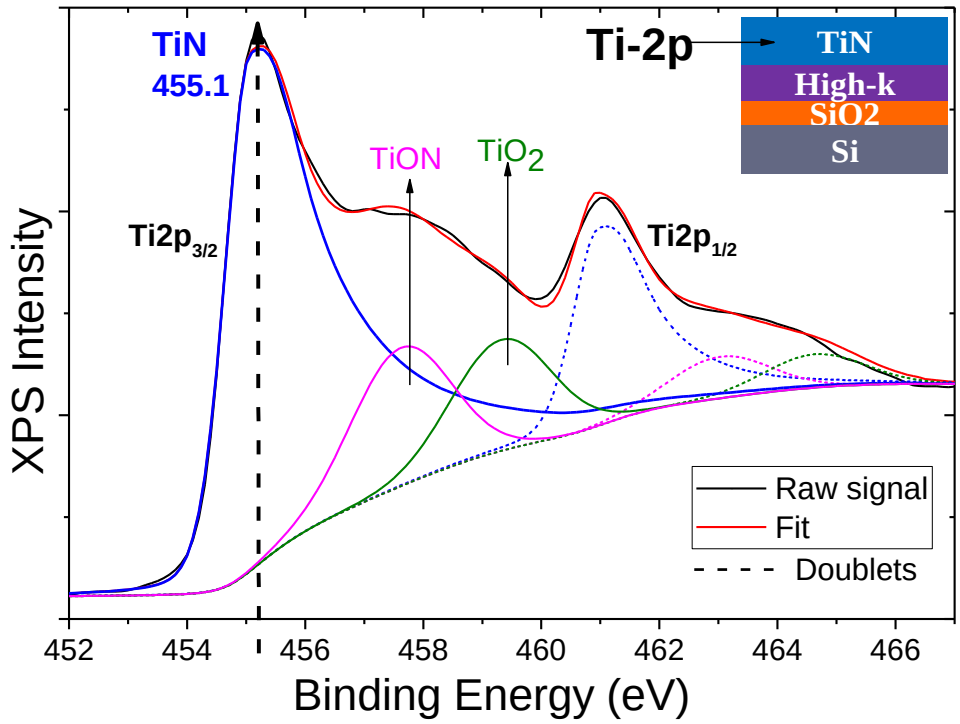


Figure 6.18: XPS spectra of Ti-2p showing fitting with different components, BE of the TiN-2p peak is used as reference for the metal gate

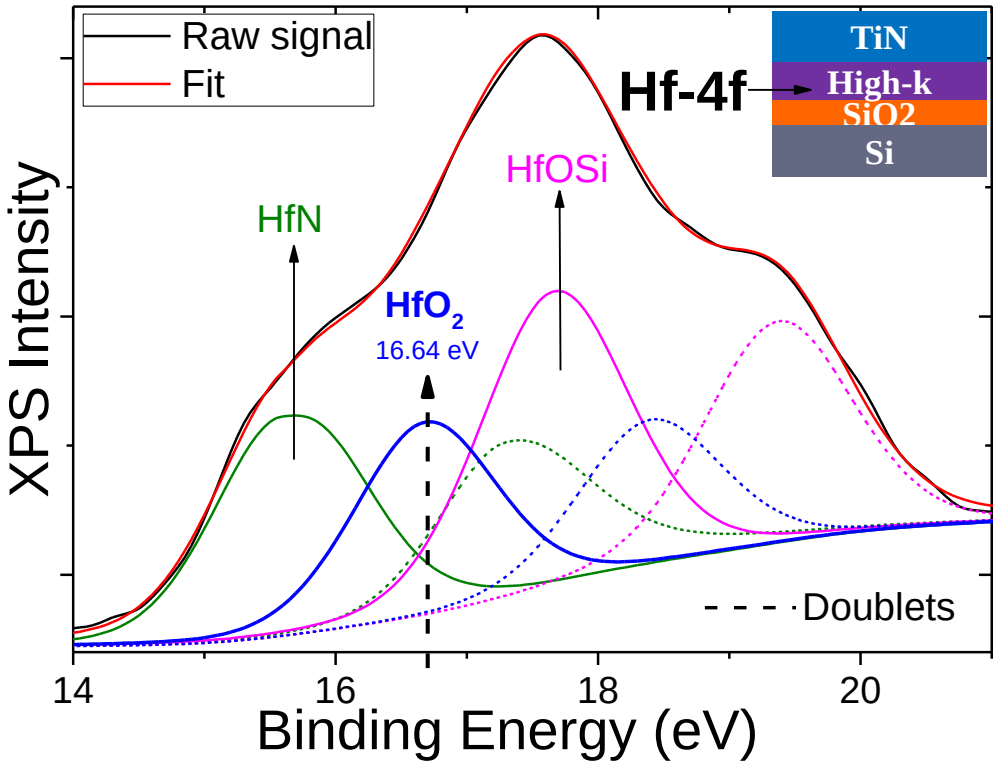


Figure 6.19: XPS spectra of Hf-4f showing fitting with different components, BE of the HfO_2-4f peak is used as reference for the high-k

In Figure 6.20 for the Si-2p signal, contrary to the Ti-2p and Hf-4f signals, the XPS signal originates from three different layers in the device: the silicon surface at the SiO_2/Si interface (Si^0-2p), the interlayer (SiO_2-2p) and the high-k (SiOHf-2p). The total Si-2p signal is fitted with six components: Si^0, Si^{+1}, Si^{+2}, Si^{+3}, SiOHf+SiON and SiO_2 (or Si^{+4}), and the BE increases with the oxidation state of Si in the chemical compounds. The Si^0-2p peak at a BE of 99.37 eV is taken as the reference BE for the Si surface and the SiO_2-2p peak was taken as the reference for the interlayer. As the various Si-2p components are from different layers of the HKMG stack, their band energies will shift in a different manner as the bias is varied.

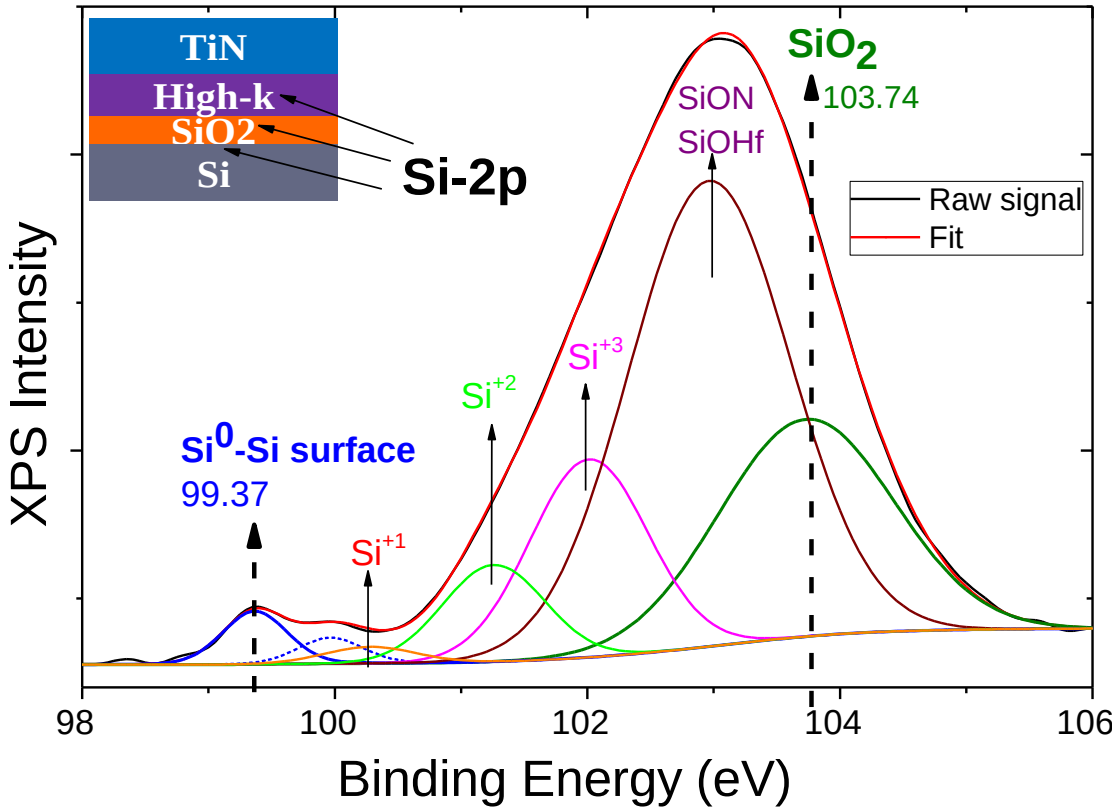


Figure 6.20: XPS spectra of Si-2p showing fitting with different components, BE of the Si^0-2p and SiO_2-2p peak are used as reference for Si substrate and SiO_2 IL respectively

6.4.2 Shift of binding energies with bias

XPS spectras for Ti, Hf and Si were acquired for each applied bias between 0 and strong accumulation (0 to -2 V). These signals were fitted with the method described in Figure 6.18, Figure 6.19 and Figure 6.20, and the BE were extracted for TiN (TiN-2p), high-k (HfO_2-4f), SiO_2 (SiO_2-2p) and Si surface (Si^0-2p). As the applied bias is varied, the band energies of the metal gate is expected to change and so the CL BE measured by XPS. Indeed, the Ti signal from the metal gate shifts with bias, as seen in Figure 6.21. Similar shifts with bias were observed for Hf signal from the high-k and the Si signal, shown in Figure 6.23. The observation of such shifts with bias, as observed in previous work [131], is the proof that our system described in Figure 6.11 for in-situ XPS with bias is working properly.

A first thing to check will be that the biasing of the device does not affect the general shape of our various spectra, and so keeping the same spectra fitting procedure described in the section above. In order to do this, Ti, Hf and Si signals were fitted at each bias and a factor D is calculated for each signal as:

$$D\,(Vg) = BE(0) - BE(Vg) \qquad 6.4$$

where BE(Vg) and BE(0) are the extracted BE's at a specific applied voltage Vg and at Vg = 0 respectively. D is calculated for BE's corresponding to different gate stack layers, i.e. TiN (TiN-2p), high-k (HfO_2-4f), Si surface (Si^0-2p). Next the Ti, Hf and Si spectras at each bias are shifted by the value of D (Vg) calculated for TiN-2p, HfO_2-4f and Si^0-2p respectively. Figure 6.22, Figure 6.23 and Figure 6.24 shows the summary of the shifted Ti, Hf and Si spectra. In these Figures, we have plotted the XPS spectras with BE + D, for which D changes with bias. Shifted Ti and Hf signals coincide with the signal at Vg = 0, which confirms that the same signal fitting, as done in Figure 6.18 and Figure 6.19, applies at every bias.

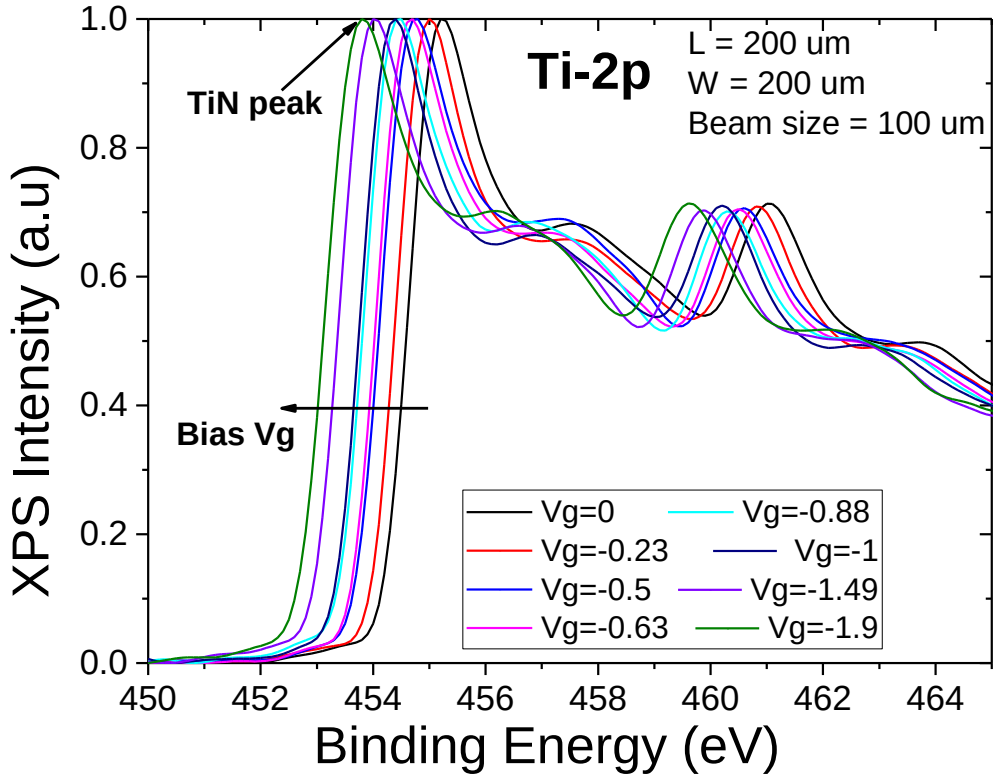


Figure 6.21: Shift of Ti-2p XPS spectra with applied bias

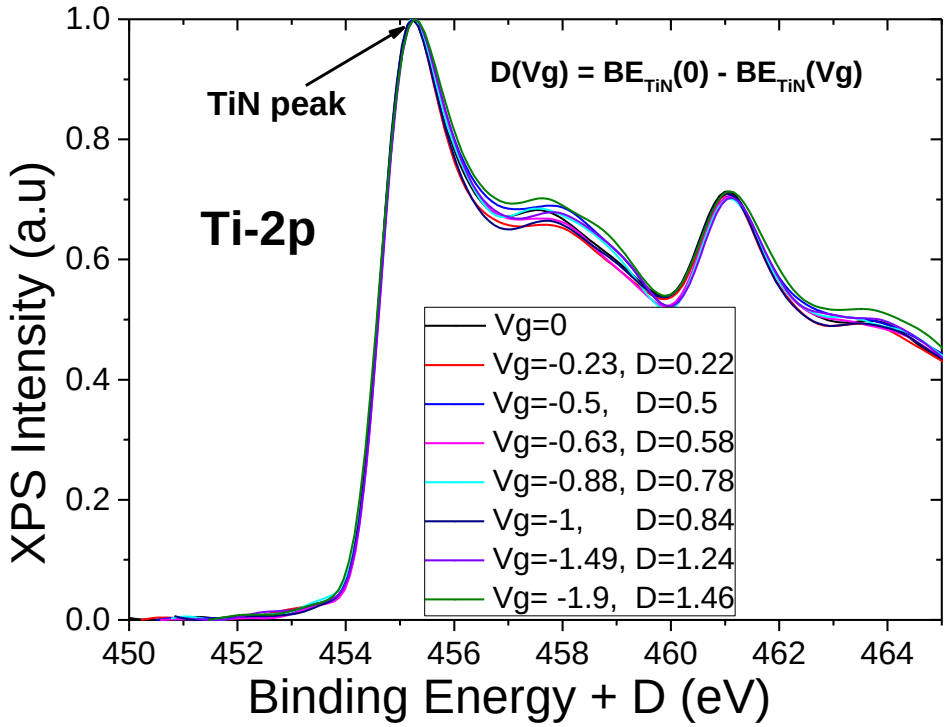


Figure 6.22: Ti-2p spectra shifted by D, showing that TiN peaks at different bias merge with that at Vg=0 and thus the validity of the fitting technique at every bias

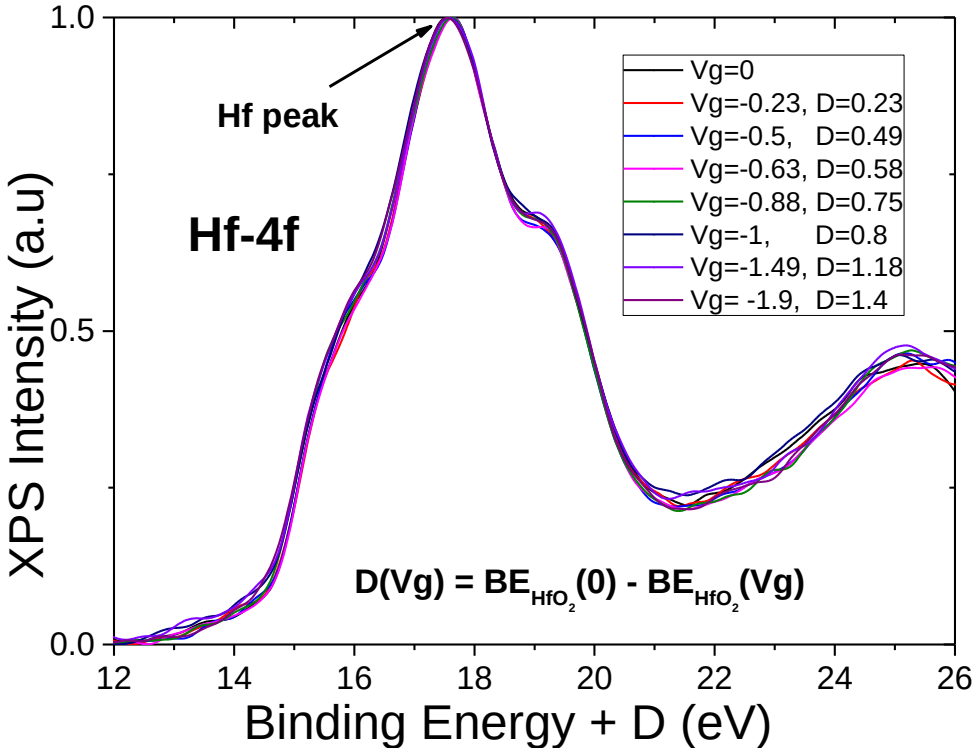


Figure 6.23: Hf-4f spectra shifted by D, showing that Hf peaks at different bias merge with that at Vg=0

In the case of Si signal in Figure 6.24, where D is calculated from the Si^0-2p signal, we notice that the Si^0 peaks concide but the SiO_2 peaks do not. This confirms that the Si spectra from the HKMG stacks corresponds to different layers in the gate stack. As the band bending with bias is different for SiO_2 and Si, the SiO_2 peaks in Figure 6.24 do not concide when Si^0 peak does. Note that in the other case if D is calculated from the SiO_2-2p peaks, the shifted SiO_2 peaks will concide but not the shifted Si^0 peaks.

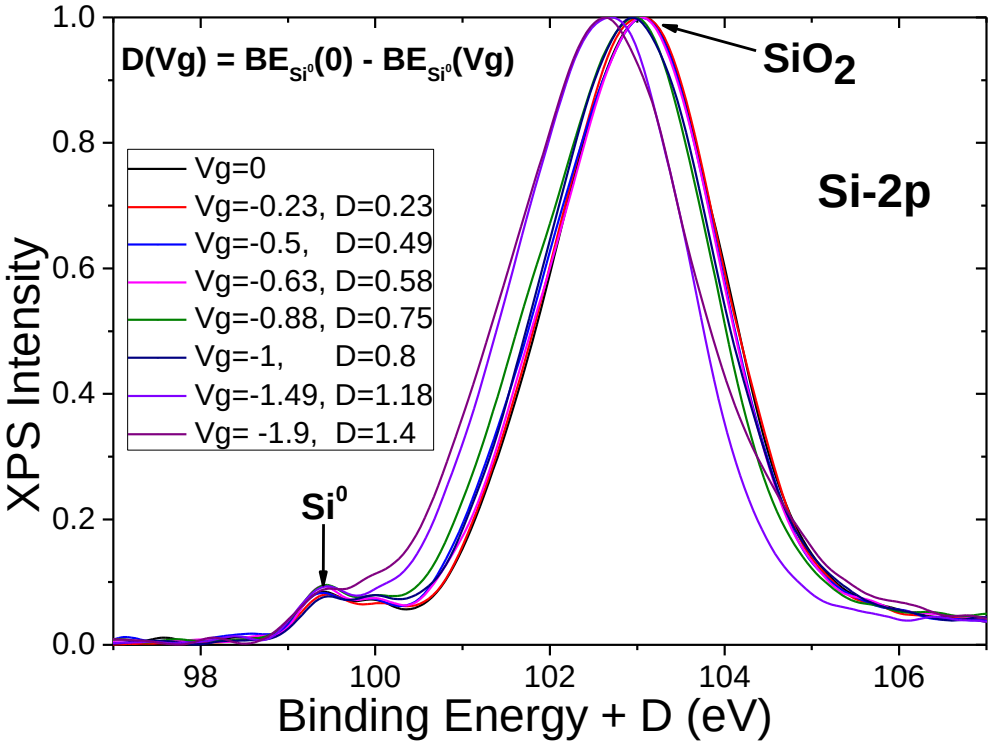


Figure 6.24: Si spectra shifted by D, showing that Si^0 peaks at different bias merge with that at V=0 but not the SiO_2 peaks

The shift of BE with bias (BE(Vg) - BE(0)) for each gate stack layer (TiN-2p, HfO_2-4f, SiO_2-2p and Si^0-2p) is plotted in Figure 6.25. The dotted black line respresents the expected ideal values for the metal gate, i.e. in case there is no bias drop in metal gate (DV1= 0 in Figure 6.4). BE of all the gate stack layers shift with bias, and the shift of TiN BE is lower than the expected shift of the gate (dotted line), revealing that parasitic bias drops in TiN gate (DV1) are not zero. HfO_2 BE shift follows the TiN BE shifts very closely at all the applied bias, which means that the bias drop or band bending in the high-k layer is not significant. This result can be explained by the high permitivity value of the high-k layer and so from Eq. 1.4, the bias drop V_{OX} inside the high-k layer would be very low. Si^0 and SiO_2 BE shifts follows the TiN BE shift for low applied voltages, but deviates significantly at higher voltages. The difference between TiN and Si^0 corresponds to bias drop inside the dielectrics, which will be mostly in the SiO_2 interlayer. Moreover, the signal related to SiO_2 is intermediate beween Si^0 and TiN as expected, detailed explaination will be given in section 6.5.2.

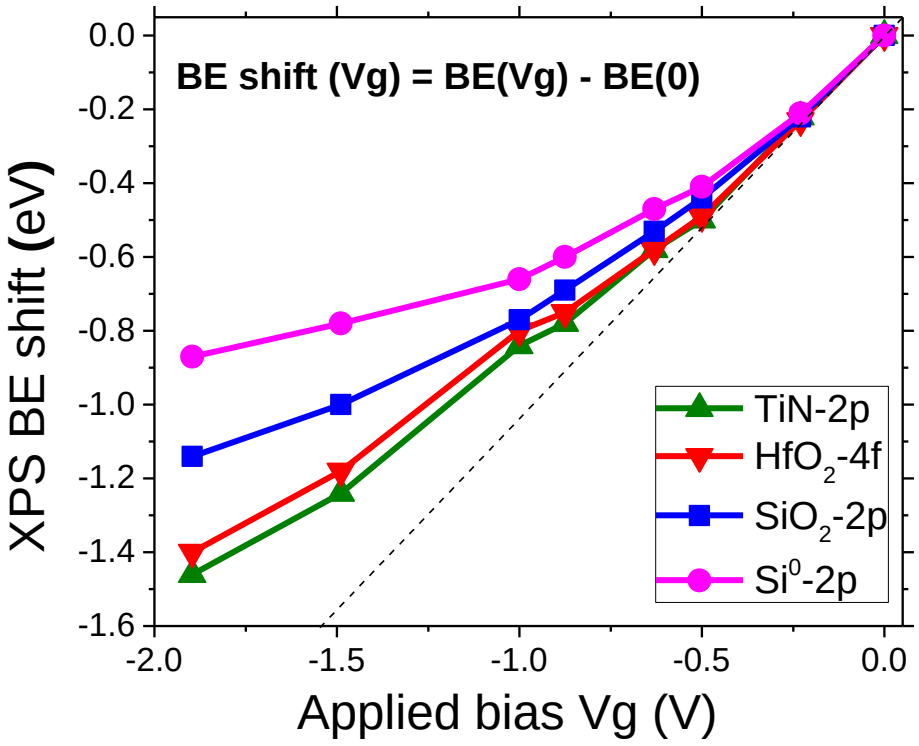


Figure 6.25: Variation of binding energy (BE) of the gate stack layers, extracted from their respective signal fitting, with applied bias

6.4.3 Technique validation

As stated at the begining of this section, with a combination of measurements on the XPS device of 200 µm and a 10 µm device, we can evaluate for the 200 µm device the TiN potential V_{TiN}, the bias across the device V_{DUT} and the substrate bias drop DV2 at any applied gate bias (Vg). In order to validate XPS under bias technique, we must check that the shift of XPS binding energy of the different gate stack layers obtained in Figure 6.25 follows what is expected from their electrical modelling. This can be done by the analysis of two potentials: the TiN potential shift and the Si surface potential shift.

TiN potential shift: V_{TiN} was calculated for every applied bias Vg by the following equation:

$$V_{TiN}(Vg) = V_{DUT}(Vg) + DV2(Vg) \qquad 6.5$$

V_{TiN} and DV2 are calculated from the respective I(V) measurements of 200 µm and 10 µm devices, and the TiN-2p BE shift obtained from XPS signal fitting (of Figure 6.25) are compared at different device bias V_{DUT} in Figure 6.26. We see that an accurate correlation is seen between the two measurements, demonstrating the validity of XPS TiN BE extraction and that, it can be used to estimate the potential of the metal gate V_{Ti}.

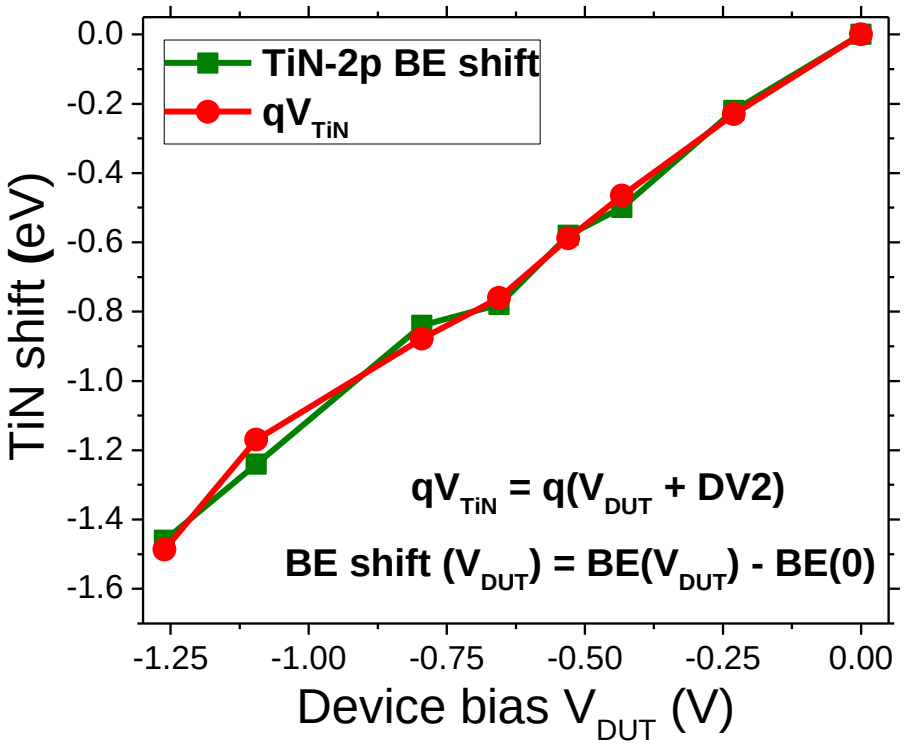


Figure 6.26: Comparison of TiN-2p BE shift, extracted from Ti-2p XPS signal fitting, with TiN bias V_{TiN} (Calculated from leakage current modelling)

Si surface potential shift: As explained in Figure 6.10, substrate bias drop DV2 will also result in a shift of gate stack band energies. Thus, Si surface potential for a 200 µm device (V_{Si}), with respect to the backside substrate, will be obtained from the evaluation of the Si surface potential on a 10 µm device ($\Psi_{Si\,(10*10)}$) at a particular V_{DUT} plus the substrate bias drop DV2, leading to the following equation:

$$V_{Si}\,(V_{DUT}) = \Psi_{Si\,(10*10)}\,(V_{DUT}) + DV2\,(V_{DUT}) \qquad 6.6$$

Thus, apart from utilizing I(V) measurements to calculate DV2, the analysis of V_{Si} requires the use of C(V) to calculate $\Psi_{Si\,(10*10)}$. It can be calculated by fitting the C(V) curves obtained for the 10 µm device, using a self consistent Poisson-Schrodinger simulation presented in section 2.1.1, from Eqs. 2.5 and 2.6. Now, we will compare the shift of V_{Si} (ΔV_{Si}) with the shift in Si^0-2p binding energy obtained in Figure 6.25. ΔV_{Si} is the Si surface potential value at a particular V_{DUT} with respect to the its value at 0 bias, given as:

$$\Delta V_{Si}\,(V_{DUT}, 0) = V_{Si}\,(V_{DUT}) - V_{Si}\,(0) \qquad 6.7$$

Figure 6.27 compares Si^0-2p BE shift obtained from XPS measurements with the Si surface bias shift ΔV_{Si}, for the 200 µm device at different device bias V_{DUT}. We see that Si^0-2p BE shift matches quite accurately with the ΔV_{Si} at low V_{DUT}, but then deviates as the V_{DUT} increases. This might mean that there is either a problem in the estimation of DV2 or in the extraction of $\Psi_{Si\,(10*10)}$ from C(V) fitting. A good correlation of TiN shifts in Figure 6.26 shows that the calculated value

of DV2 is correct at every bias. Hence, we will investigate the extraction of $\Psi_{Si\ (10*10)}$ from experimental C(V) curves fitting, as shown in Figure 6.28.

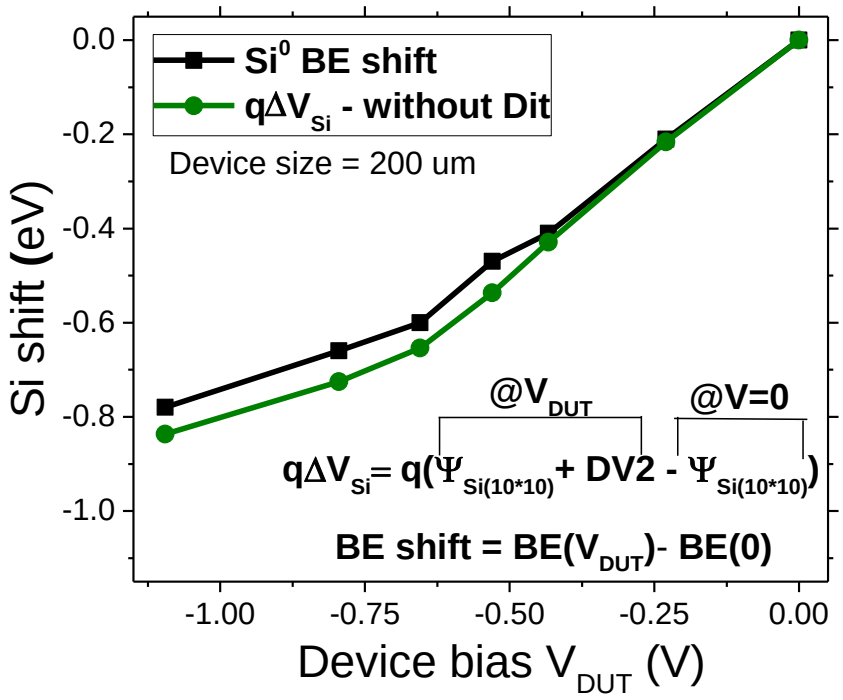


Figure 6.27: Comparison of Si^0-2p BE shift, calculated from Si-2p XPS signal, with Si surface bias (V_{Si}) shift (calculated from CV modelling of a 10*10 device without Dit effect)

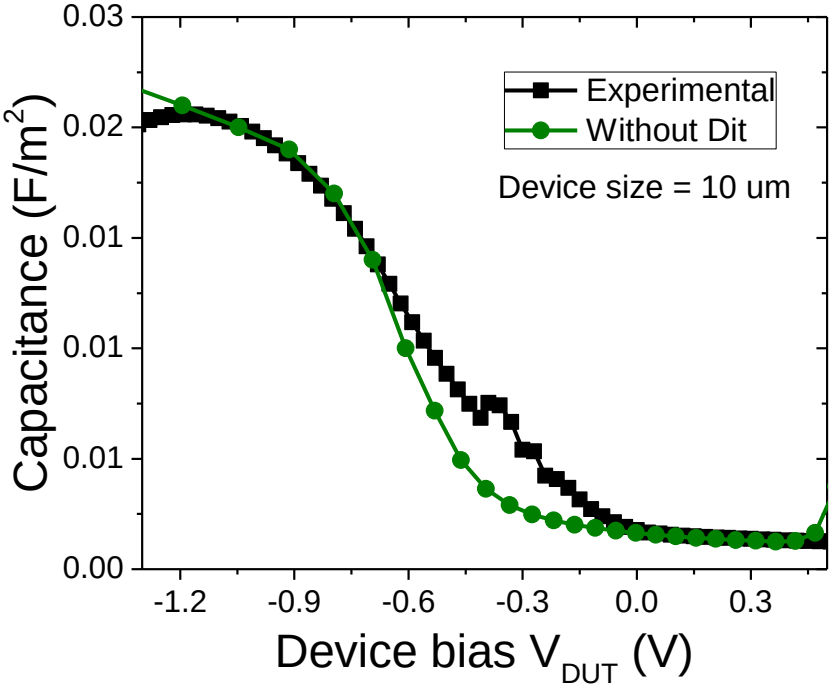


Figure 6.28: Fitting of C-V curves by Poisson-Schrodinger model without D_{it} effect

Figure 6.28 shows that the fitting curve, obtained from Poisson-Schrödinger simulations, does not fit the experimental C(V) curve at low voltages. This type of behavior observed on the experimental curve is usually due to the presence of interface states D_{it}. Depending on the amount of D_{it} present and the bias applied, D_{it} will have two effects on the gate stack:

1) First, depending on the small AC bias applied across the device in addition to the static bias, the D_{it} gets filled or unfilled, depending on the frequency of the AC signal. Thus, it can give rise to an additional component C_{it} in the total capacitance of Eq. 2.1, as follows:

$$\frac{1}{C_{exp}} = \frac{1}{C_{OX}} + \frac{1}{C_{Si}+C_{it}} \qquad 6.8$$

The added capacitance C_{it} affects the measured capacitance C_{exp}, and so changes the shape of the experimental C(V) curves (in black) in Figure 6.28.

2) Secondly, due to filling of the D_{it} with the static bias, a charge Q_{it} at the Si/SiO_2 interface is necessarily generated, and will affect the substrate charge and its potential.

Eq. 2.1 and 2.3 has been used to get the C(V) fit in Figure 6.28, by Poisson-Schrodinger simulations (as done in section 2.2.1) and gives the relationship between the applied bias across the gate stack V_{DUT} and the total charge. For simplicity, Eq. 2.3 is used here again as Eq. 6.9.

$$V_{DUT} = V_{FB} - Q_{Si}/C_{OX} + \Psi_{Si} \qquad 6.9$$

As explained in point 2) above also, Eq .6.9 must be modified, leading then to the following equation 6.10.

$$V_{DUT} = V_{FB} - (Q_{Si}^* + Qit)/C_{OX} + \Psi_{Si}^* \qquad 6.10$$

In Eq. 6.10, the bias drop in Si is then called Ψ_{Si}^* because the relationship between Ψ_{Si} and V_{DUT} will be modified, it is the same for Q_{Si}^*. But we must notice that the relation $\Psi_{Si}^*(Q_{Si}^*)$ is still the same relation as Ψ_{Si} (Q_{Si}), which is typical of the Si substrate. Moreover, in Eq. 6.10, Q_{it} is a function of Ψ_{Si}^*, and by definition at V_{FB} ($\Psi_{Si}^* = 0$) $Q_{Si}^* = 0$.We must also notice that, in Eq. 6.8 we assume that the AC signal frequency is significantly low to allow the response of all the D_{it}. Then, C_{it} is given by Eq. 6.11:

$$C_{it} = -dQ_{it}/d\Psi_{Si^*} \qquad 6.11$$

and Qit is given by Eq. 6.12

$$Q_{it} = \int_{VG1}^{VG2} q \cdot D_{it} \cdot d\Psi_{Si^*} \qquad 6.12$$

In Figure 6.29a, we propose a D_{it} profile given here as $Q_{it}(V_{DUT})$ that fits well on the experimental C(V) curve in Figure 6.29b.

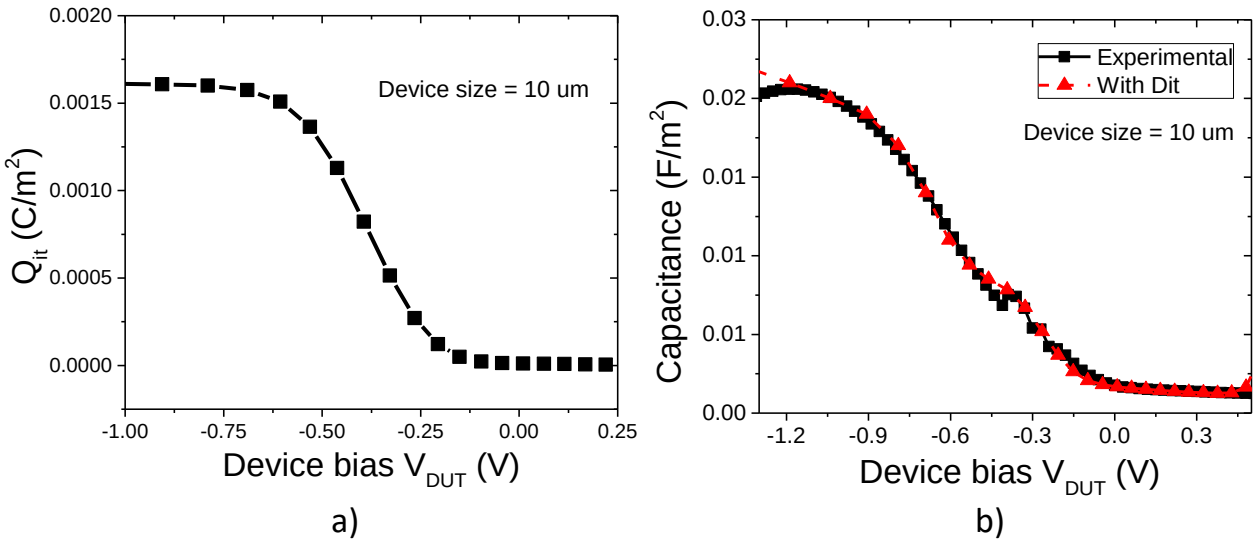


Figure 6.29: a) D_{it} charge (Q_{it}) found, that fit experimental C(V) and b) fitting of C-V curves by Poisson-Schrödinger model including the Qit found in (a)

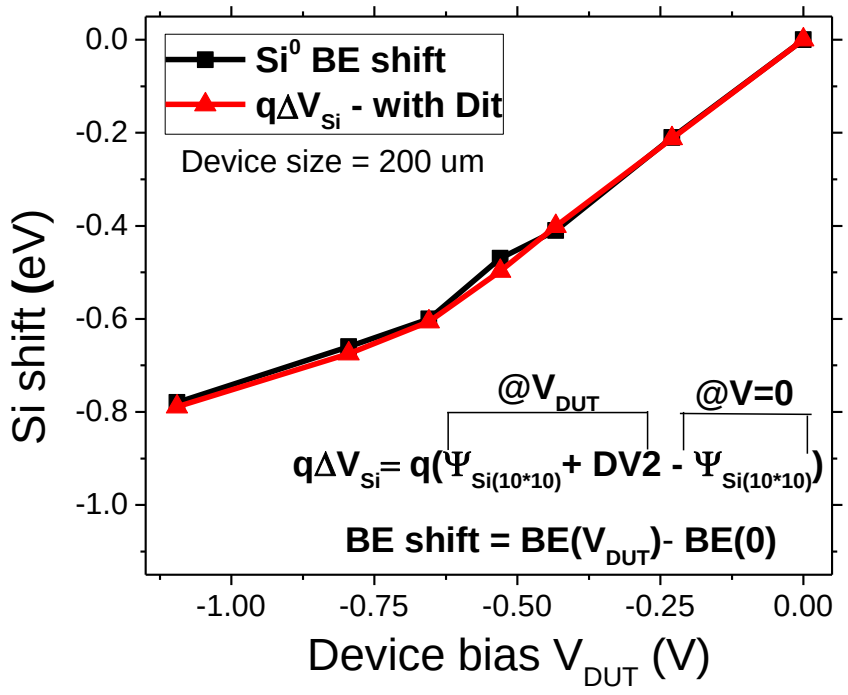


Figure 6.30: Comparison of Si^0-2p BE shift, calculated from Si-2p XPS signal, with Si surface bias (V_{Si}) shift (calculated from CV modelling of a 10*10 device with Dit effect)

It is also possible to extract the D_{it} profile from the comparison between simulation and XPS measurements. We assume that the over estimation observed for ΔV_{Si} in Figure 6.27 is due to the fact that we do not take into account the interface states Dit.

Thus, we determine for each V_{DUT}, Ψ_{Si}* that would fit with XPS BE shift of Figure 6.27, leading to Eq .6.13.

$$\Psi_{Si}^* = [BE_{Si}^0(V_{DUT}) - BE_{Si}^0(V_{FB})] - [DV2(V_{DUT}) - DV2(V_{FB})] \quad 6.13$$

where $BE_{Si}^0(V_{DUT})$ and $BE_{Si}^0(V_{FB})$ are the Si^0-2p BE's, and $DV2(V_{DUT})$ and $DV2(V_{FB})$ are the substrate bias drops for the 200 µm device at a certain V_{DUT} and at V_{FB} respectively. After having calculated Ψ_{Si}*, the Ψ_{Si} (Q_{Si}) Poisson-Schrodinger database allows to calculate the substrate charge Q_{Si}* associated with Ψ_{Si}*. Two simulations, $\Psi_{Si}^*(V_{DUT})$ and $\Psi_{Si}(V_{DUT})$, that account to a same C(V) shape in large accumulation (Figure 6.28) assume in fact a same V_{FB} value in Eq. 6.9 and 6.10. Thus, Eq. 6.9 and 6.10 can be equated at the same V_{DUT} and gives rise to Eq. 6.14 below.

$$\Psi_{Si}^* - (Q_{Si}^* + Q_{it})/C_{OX} = \Psi_{Si} - Q_{Si}/C_{OX} \quad 6.14$$

Thus, from Eq. 6.14, Q_{it} is given by:

$$Qit = (\Psi_{Si}^* - \Psi_{Si}) \cdot C_{OX} - (Q_{Si}^* - (Q_{Si}) \quad 6.15$$

Figure 6.31 presents the value of Q_{it}, calculated from the above equation, versus the V_{DUT} (red curve). A comparison with the Q_{it} found in Figure 6.29a, by fitting the low frequency curve, is also made, and we see that both saturation value of D_{it} are close. Q_{it} profile on the two curves are not exactly the same, and can be due to the fact that the C(V) curve used to obtain Q_{it} (curve in black) is not done exactly at the lowest frequency required to obtain the entire D_{it}. Moreover, the profile obtained from Si^0 BE shift (curve in red) has only a limited number of measurements or data points, and the level of Q_{it} is also quite low compared to the accuracy and capability of XPS measurements, which leads to a noisy behavior in the Q_{it} data. Nevertheless, if Q_{it} values are high enough, it could be precisely measured by the method presented above.

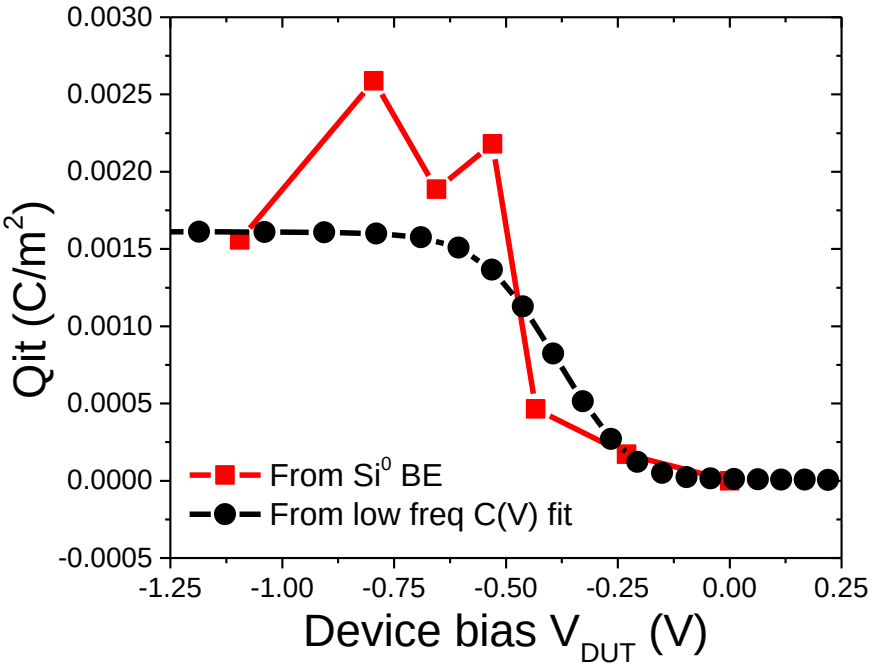


Figure 6.31: D_{it} charge (Q_{it}) obtained from the difference between ΔV_{Si} without D_{it} and Si BE shift

6.5 Dipole localization by XPS under bias

In the previous sections, we have presented XPS and electrical measurements on a set of big and small devices, to solve the biasing issues related to the test structures used for XPS measurements. Then, a method to fit the different XPS signals, from the MOS device with the HKMG stack, was presented in order to obtain BE for the different layers. The accuracy of the biasing calculations and BE extraction is justified or validated, by comparison between XPS BE shifts and the bias drops calculated from our electrical modelling based on C(V) and I(V) measurements. Now, after having validated an XPS under bias methodology, we can apply it to localize various dipoles in the gate stack [PK-3].

6.5.1 Dipoles related to La and Al addition

In section 1.2.4, La or Al additives were introduced to tune the flatband voltage of the MOS devices. These additives form dipoles inside the HKMG stack, causing MOS flatband voltage shift. In this section, these dipoles will be localized and quantified inside the HKMG stack, by the XPS under bias technique we have developed. The test structures and their process flow, for MOS devices used in this study, were introduced in section 6.2.1 and Figure 6.3. The HKMG stack consists of SiO_2 interlayer, HfON or HfSiON as high-k and TiN as metal gate. La or Al addtives were introduced by the sacrificial gate process of Figure 1.18, with drive in anneal (DIA) temperature of 900°C and DIA time of 10 s or 30 s. The sacrificial gate stack, DIA conditions, diffused additive dose and the resulting flatband voltage modulation are shown in Table 6.1.

Sacrifical gate stack	DIA conditions	Diffused dose (at/cm^2)	Flatband voltage V_{FB}(V)	Δ V_{FB}(V)
HfON/La-0/TiN	900°C, 10s	0	-0.553	0
HfON/La-2Å/TiN	900°C, 10s	$2.36 \cdot 10^{14}$	-0.688	-0.135
HfON/La-4Å/TiN	900°C, 10s	$4.5 \cdot 10^{14}$	-0.778	-0.225
HfON/TiN10Å/Al2.2Å/TiN	900°C, 30s		-0.453	0.1
HfSiON/La-0/TiN	900°C, 10s	0	-0.53	0
HfSiON/La-6Å/TiN	900°C, 30s	$1.2 \cdot 10^{14}$	-0.65	-0.12
HfSiON/TiN10Å/Al2.2Å/TiN	900°C, 10s	$1.7 \cdot 10^{14}$	-0.488	0.042

Table 6.1: Sacrificial gate stack and its DIA conditions, used to introduce La and Al inside the gate stack, and the resulting diffused dose and flatband voltage of the MOS devices

For each of the splits reported in Table 6.1, I(V) measurements were perfomed on a 10*10 µm^2 and a 200*200 µm^2 device. Following the methodology introduced in section 6.2.3, I(V) measurements on both sizes were used to calculate for the 200 µm device, the relationship between the leakage current and device bias V_{DUT} and substrate bias drop DV2. Identifying such relationship allows the calculation of the leakage current value (I_{FB}) corresponding to a V_{DUT} = V_{FB}. XPS

measurements were performed on the same 200 μm device, with a total XPS beam size of around 200 μm. During the XPS measurement the device was biased to two different conditions: 1) at 0 V and 2) at the flatband condition. Ti-2p, Hf-4f and Si-2p signals were acquired during XPS measurements, at both 0 and flatband condition, and the BE for TiN-2p, HfO_2-4f, SiO_2-2p and Si^0-2p were extracted.

For devices with different La or Al dose, we will present the results on BE at V_{FB}, by comparing them with the reference devices without any La or Al additives (splits 1 and 5 in Table 6.1). Thus, we plot the shift in BE (ΔBE) versus the shift in V_{FB} (ΔV_{FB}). As stated earlier, the best condition to compare energy bands of different gate stacks is the flatband condition, because the BE signal from a specific layer inside the gate stack will be the same for its entire thickness. Thus, we would present the results at the flatband condition, to compare the BE of devices with different La or Al doses.

These results are presented in Figure 6.32 for splits 1-4 in Table 6.1, at their respective flatband condition. From this figure, no clear trend is observed between ΔBE and ΔV_{FB}, for BE of any of the gate stack layers. Indeed, modulation of dipole at an interface in the HKMG stack would have resulted in a linear relationship for some layers and constant for others. More importantly, Si^0 BE at flatband, corresponding to the Si substrate, is expected to be the same irresepective of the additive dose (or gate stack). This means that ΔBE for Si^0 should have been equal to 0, but this is not the case in Figure 6.32. The value of substrate bias drop DV2 at V_{FB} is also shown for each split, and it is not constant for the different splits due to a different V_{FB} and leakage current for each. The unexpected behaviour observed in Figure 6.32 is actually due to the parasitic bias drop in Si substrate DV2. As explained in Figure 6.10, DV2 will shift the BE of all the layers of the stack, and as its value varies among the four samples, the BE shift caused by it will be different.

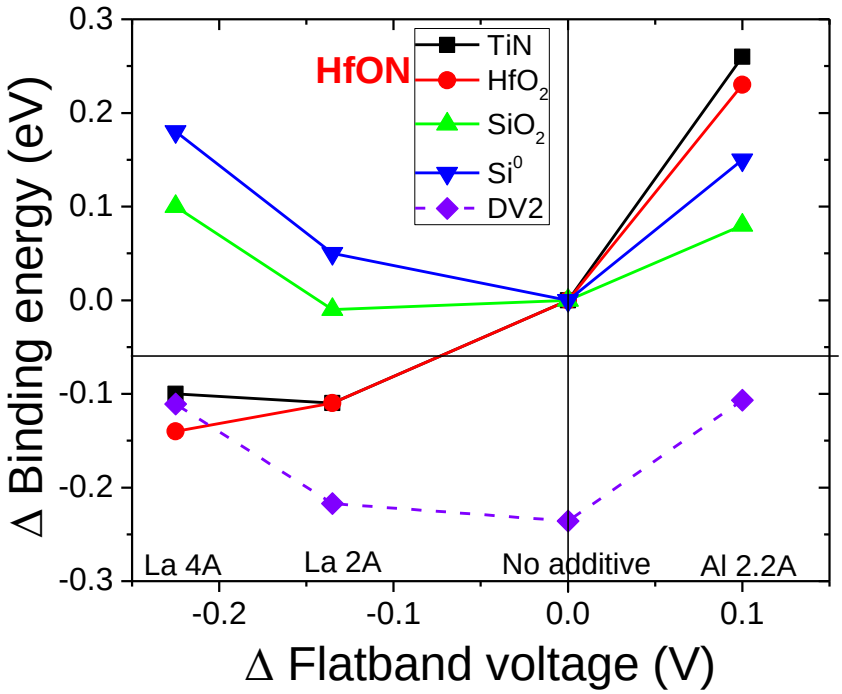


Figure 6.32: Raw binding energy shift at V_{FB}, as well as DV2, versus flatband voltage shift

We assume that the lack of a regular trend in Figure 6.32 is due to DV2 variation among devices. It shows the importance of the methodology developed in this work to accurately estimated the value of DV2. In order to compare gate stacks with different values of DV2, their experimental BE values at flatband (raw values) must be corrected by their respective DV2 value at flatband (Figure 6.10), as given below:

$$BE_{corrected}(V_{FB}) = BE_{raw}(V_{FB}) - DV2(V_{FB}) \qquad 6.16$$

Figure 6.33 shows the corrected BE shift versus the flatband voltage shift. We see that Si^0 and SiO_2 BE's remains quite constant for all the splits, meaning that no dipole modulation is taking place at this interface due to La or Al addition. On the other hand, TiN and HfO_2 BE's shift by the same amount and moreover follows a linear trend: $\Delta BE = q\Delta V_{FB}$. TiN and HfO_2 BE's shifting by the same amount and that of Si and SiO_2 remaining constant, means that there is a dipole modulation only at the high-k/SiO_2 interface and not at the others. Secondly, TiN and HfO_2 BE's following the trend $\Delta BE = q\Delta V_{FB}$ means that the value of the dipole modulation by La or Al at this interface can by calculated by the shift in these BE's. As the Si^0 values should have been constant, its slight variation gives an indication of the accuracy of this technique, and we see that an excellent accuracy (< 50 meV) is obtained for our results.

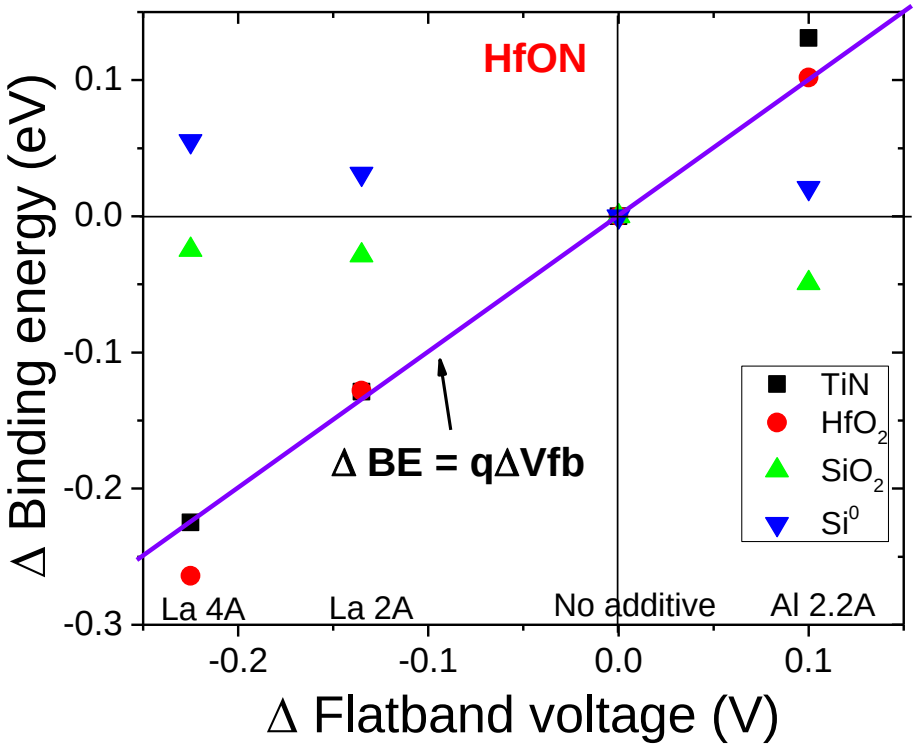


Figure 6.33: Binding energy shift at flatband, corrected by substrate bias DV2, versus the flatband voltage shift for samples with La and Al additives on HfON

Figure 6.34 shows similar results of ΔBE versus ΔV_{FB} with different La or Al dose, but on HfSiON high-k. For La addition, similar to results obtained on HfON in Figure 6.33, Si and SiO_2 BE's remains quite constant, and TiN and HfO_2 BE's shift follows the trend $\Delta BE = q\Delta V_{FB}$. In case

of Al addition, the trend is not clear as the V_{FB} shift caused by Al is only 42 meV, which is within the range of the accuracy of this XPS under bias technique.

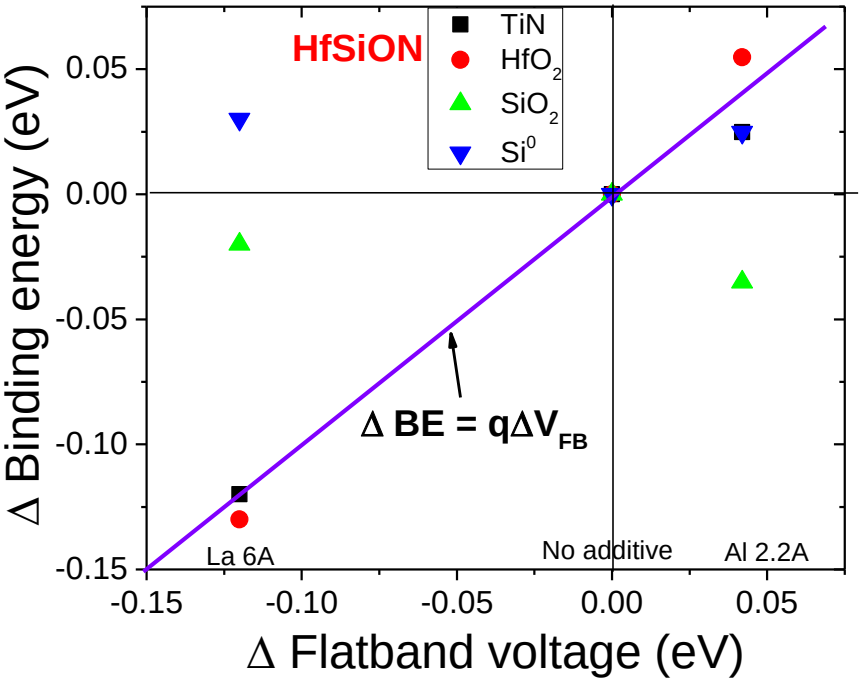


Figure 6.34: Binding energy shift at flatband, corrected by substrate bias DV2, versus the flatband voltage shift for samples with La and Al additives on HfSiON

Thus, we prove here for the first time that Al and La additives induce no dipole modification at high-k/metal gate interface or at the SiO_2/Si interface, but only at the SiO_2/high-k interface that is equal to $q \cdot \Delta V_{FB}$ as shown by the band diagram in Figure 6.35.

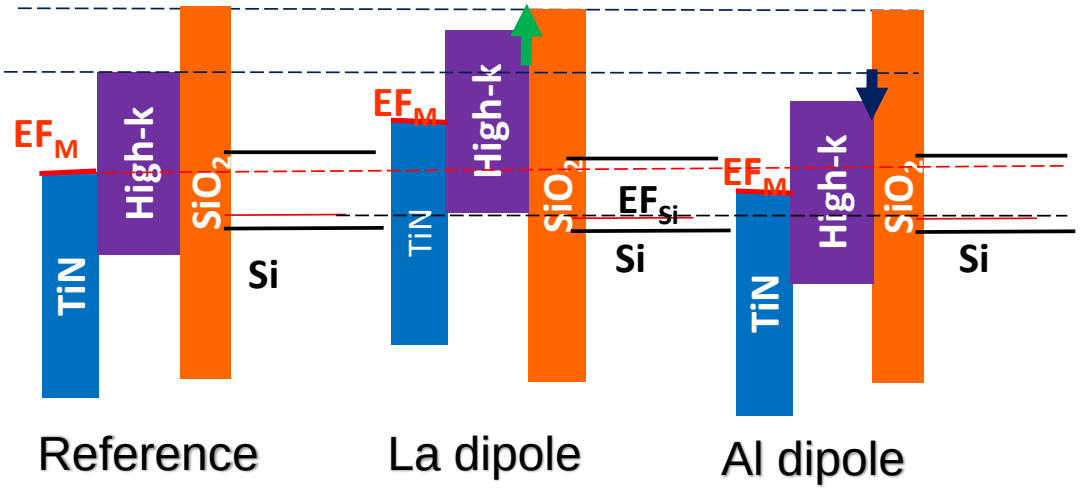


Figure 6.35: HKMG stack band diagram at flatband, corrected by substrate bias DV2, showing the relative energy bands shifts. Relative energies change at the SiO_2/ high-k interface due to La/Al dipole modulation that is responsible for flatband change

6.5.2 Methodology to reduce experimental difficulties

In the previous section, we have presented the application of XPS under bias measurements, done at the flatband condition, to localize and quantify La or Al dipoles inside the HKMG stack. Indeed in terms of HKMG band energy analysis, measurements at the flatband condition are the most accurate, but experimentally it faces two major challenges:

1) As the leakage current at V_{FB} is non-zero, electrical modelling is required on small and big test structures, to be able to calculate device bias V_{DUT} and to evaluate substrate bias drop DV2.

2) In order to perform measurements at V_{FB}, the XPS tool must be equipped with an additional system to apply a given bias to the test structure. This is generally not the case, and most XPS tools are compatible to apply only a ground potential to the device.

On the contrary, XPS under bias measurements are easier to perform at zero bias, as there is no leakage current and so no need to calculate V_{DUT} and DV2 by electrical modelling. Moreover, there is no need of specific biasing connection inside the XPS tool and so normal XPS tools can be utilized for these measurements. Even after these benefits at zero bias, we still need the binding energy values at the flatband condition. Thus, here we define a strategy to extrapolate from results at 0 bias, the expected BE values at V_{FB}. Figure 6.36 reports ΔBE versus ΔV_{FB}, for different La or Al doses, for measurements done at zero bias.

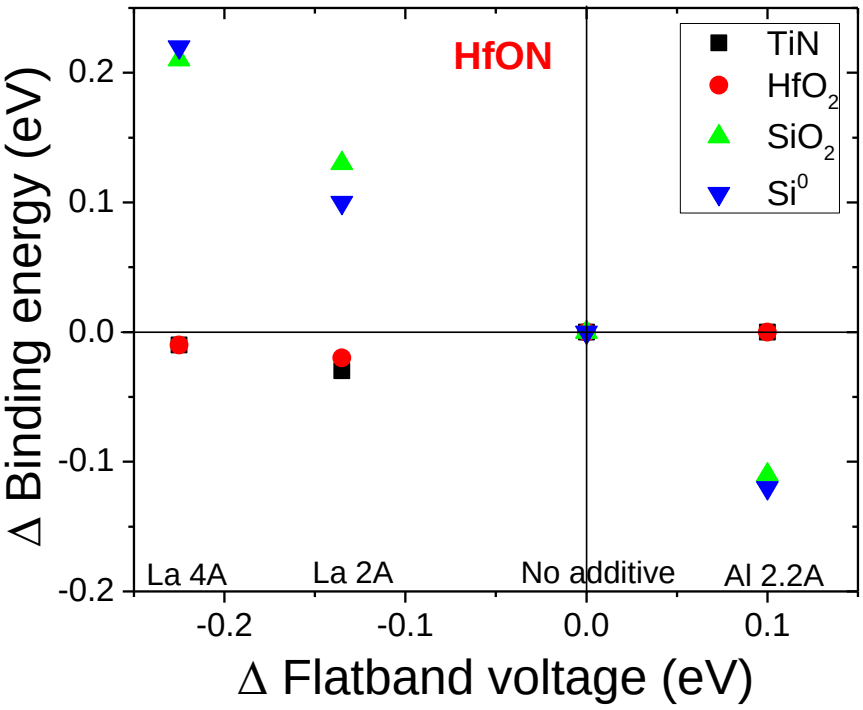


Figure 6.36: Binding energy shift, at zero bias, versus flatband voltage shift

In order to extrapolate to the flatband condition, the BE values at zero bias are shifted by δ, where δ is calculated differently for each gate stack layer. By doing this, we will be extrapolating the values for an ideal device with a zero leakage current, meaning that it does not have any bias drops in the gate (DV1) or in the substrate (DV2). Therefore, from Eq. 6.1 and Figure 6.4, at any bias, the gate potential V_{TiN} will be equal to the applied bias Vg. Thus, it will be equal to V_{FB} at the flatband condition.

$$V_{FB} = V_{TiN} = Vg \tag{6.17}$$

From Figure 6.26, the difference between TiN BE at V_{FB} and at zero bias matches well with the V_{TiN}, and hence it can be used to estimate the TiN BE at the flatband condition by Eq. 6.18.

$$BE_{TiN}(V_{FB}) = BE_{TiN}(0) + V_{TiN} = BE_{TiN}(0) + V_{FB} \tag{6.18}$$

In Figure 6.25, there is very close correlation between the HfO_2 BE shift and the TiN BE shift, at all the bias conditions. It means that bias drop inside HfO_2, as well as its variation with gate bias, is not significant. Thus, the same shift applied for TiN must be applicable to the HfO_2 BE levels, as given in Eq. 6.19.

$$BE_{HfO2}(V_{FB}) = BE_{HfO2}(0) + V_{FB} \tag{6.19}$$

By putting DV2 = 0 in Eq. 6.6, Eq. 6.7 for the silicon surface potential shift at flatband from zero bias, can be written as:

$$\Delta\Psi_{Si}(V_{FB}, 0) = \Psi_{Si}(V_{FB}) - \Psi_{Si}(0) \tag{6.20}$$

BE for Si^0 is expected to follow the shift of biasing in silicon, leading to:

$$BE_{Si0}(V_{FB}) = BE_{Si0}(0) + \Psi_{Si}(V_{FB}) - \Psi_{Si}(0) \tag{6.21}$$

As the silicon surface potential at flatband should be zero, Eq. 6.14 becomes

$$BE_{Si0}(V_{FB}) = BE_{Si0}(0) - \Psi_{Si}(0) \tag{6.22}$$

From zero bias to flatband, the band energy of SiO_2 at the interface with Si will move by $-\Psi_{Si}(0)$. Since the BE value measured at zero bias is an average over the entire SiO_2 layer thickness, we must also consider the bias drop on the whole SiO_2 layer (V_{OX}). The measurement for SiO_2 is an average on the two extreme shifts: $-\Psi_{Si}(0)$ at Si interface and $-[\Psi_{Si}(0) + V_{OX}(0)]$ at the top of the interlayer, i.e. $-(\Psi_{Si}(0) + V_{OX}/2)$ .Thus SiO_2 BE at flatband becomes:

$$BE_{SiO2}(V_{FB}) = BE_{SiO2}(0) - \Psi_{Si}(0) - V_{OX}/2 \tag{6.23}$$

If we neglect bias drop in HfO_2, as done above, V_{OX} is then equal to $-(\Psi_{Si}(0) + V_{FB})$, thus Eq. 6.23 becomes:

$$BE_{SiO2}(V_{FB}) = BE_{SiO2}(0) + \{V_{FB} - \Psi_{Si}(0)\}/2 \tag{6.24}$$

Figure 6.37 reports the projected BE values extrapolated at flatband, using zero bias and the calculations described above.

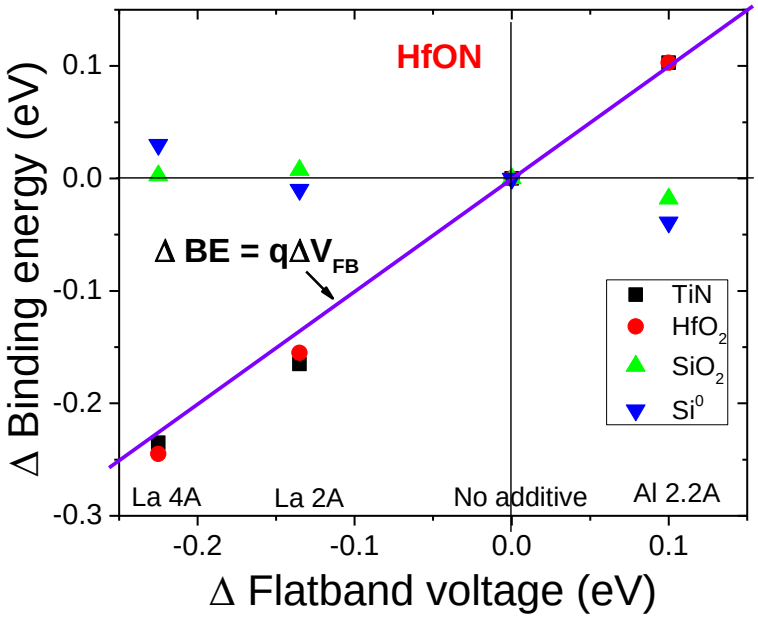


Figure 6.37: Extrapolated binding energy shift at flatband, from data at zero bias, versus flatband voltage shift for samples with La and Al additives

The results are very similar to the one obtained experimentally at the flatband condition in Figure 6.33. The differences x between Figure 6.33 and Figure 6.37, shown in Figure 6.38, are very small, with a mean value $\bar{x}$ = -1meV and a standard deviation of σ = 25 meV. These results demonstrates a very close correlation between the two methods, and a powerful way to obtain the BE values at the flatband condition from the XPS measurements done only at zero bias.

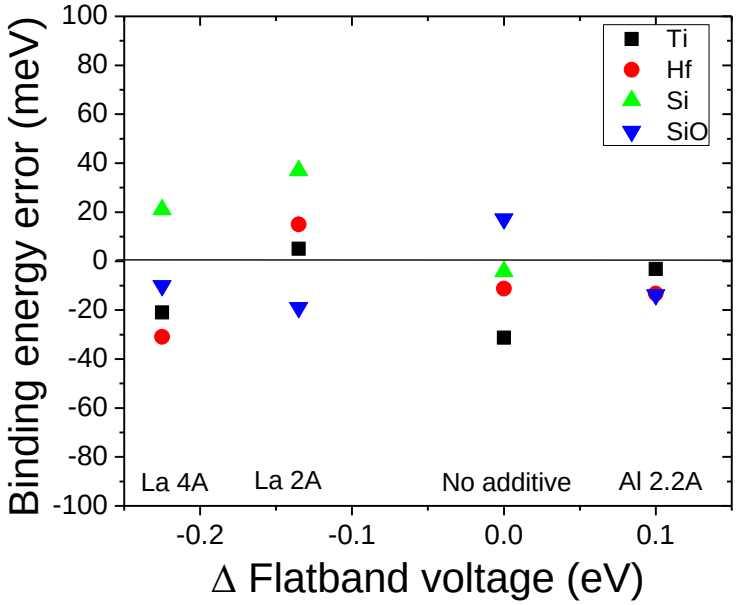


Figure 6.38: Binding energy discrepancies at V_{FB} condition between experimental results at V_{FB} and projected values from measurements at Vg = 0

6.2.3 Dipoles related to TiN thickness modulation

The effective workfunction or flatband voltage of MOS devices with HKMG stack usually increases as TiN thickness is increased. This effect is observed on both HfON and HfSiON high-k [37][164][165][166]. In a study conducted by Kai et al., the authors explained such behaviour by the variation of crystal behaviour of TiN films with its thickness, and then its further saturation [164]. In another study [165], it was shown that the effect of TiN thickness on V_{FB} depends on the capping layer. Electrodes that contains less amount of silicon relative to poly-Si, such as tungsten (W), have negligible effect and it was related to the migration of oxygen from gate oxides and oxidation of the TiN and poly-Si layer, which changed the WF_{eff}. An increase in TiN thickness reduces the oxidation of poly-Si layer on top and hence increaes its own oxidation, but TiN oxidation will be the same with W irrespective of its thickness due to low reactivity of W with oxygen [165]. However the main mechanism behind this flatband voltage modulation remains ambiguous, and the reported values in literature are quite different from one study to another.

In order to solve this ambiguity in the effect of TiN on V_{FB}, XPS under bias measurements were performed on devices with HKMG stack consisting of a bevel SiO_2 interlayer, HfSiON as high-k and TiN as metal gate with a range of TiN thickness (3.5, 5, 6.5 and 10 nm). The measurements were only done at 0 bias, and the BE values at flatband were calculated from the methodology presented in the previous section. Extrapolated results at V_{FB}, of ΔBE versus ΔV_{FB}, are presented in Figure 6.39. As with La/Al additives results, the Si & SiO_2 BE's remain quite constant and TiN & HfO_2 BE's shift by the same amount and follows the line $\Delta BE = q \cdot \Delta V_{FB}$. We prove here for the first time that it is the modulation of dipole at the SiO_2/High-k interface that is responsible for V_{FB} shift with TiN thickness, as depicted in the band diagrame of Figure 6.40.

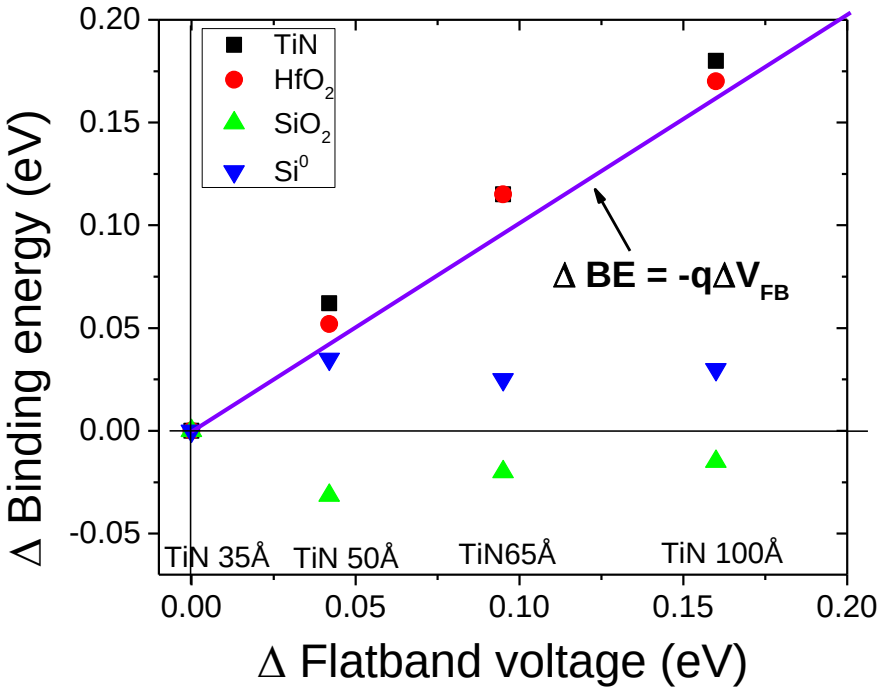


Figure 6.39: Extrapolated binding energy shift at flatband, from data at 0 bias, versus flatband voltage shift for samples with different TiN thickness

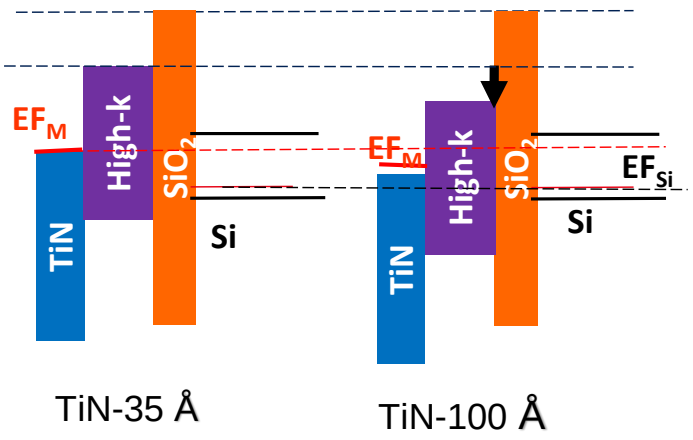


Figure 6.40: HKMG stack band diagram at flatband condition for TiN-35 Å and TiN-100 Å gate thicknesses, showing dipole modulation at the SiO_2/ high-k interface

6.6 Conclusion

In this chapter, a robust and accurate XPS under bias technique has been developed, able to determine the different band energies shift inside the HKMG stack layers of MOS devices [PK-3] [PK-5]. New test structures has been presented that enable to perform XPS under bias measurements. The biasing issues related to these test structures, has been identified and modelled. Moreover, an electrical model and appropriate methodology are introduced for the first time, in order to determine the parasitic bias drops and correct biasing for the flatband condition and at other device biases.

Experimental issues related to the XPS equipment such as device location and XPS beam size, that can influence XPS under bias measurements, has been identified and solutions has been found to deal with them. Moreover, it has been shown that the XPS beam does not have any electrical impact on the MOS device during XPS under bias measurements. The issue concerning a large parasitic substrate resistance, that presented a hindrance in performing accurate XPS measurements, have been significantly reduced.

XPS under bias technique has been validated by first defining a robust fitting methodology for various XPS signals, and then comparing XPS binding energy shifts with what is expected from our electrical model. Moreover, a strategy has been presented, by combining XPS binding energies and capacitance measurements, to calculate the interface states present at the SiO_2/Si interface.

Finally, this technique has been applied to analyze the origin of effective work function (WF_{eff}) shift related to La and Al additives or to TiN thickness modulation. It has been shown for the first time, by relative band energy measurements, that in all these cases the dipole modulation at the SiO_2/high-k interface is responsible for these observed WF_{eff} changes. Moreover, it has been shown that band energies at flatband condition can be obtained from measurements at zero bias condition. This technique has the potential to be used to characterize devices beyond CMOS, for other technologies.

Conclusions and perspectives

This Ph.D. thesis has focused on the impact of high-k metal gate (HKMG) stack processes on the reliability parameters of the FD-SOI MOSFET devices, their effective workfunction (WF_{eff}) and the diffusion of additives in them. These parameters have been measured by combined electrical and physicochemical characterization techniques, and also involves the development of an innovative physical characterization technique (XPS under bias). The main objectives of this thesis were to investigate : 1) the effects of La and Al additives on the BTI (bias temperature instability) and TDDB (time dependent oxide breakdown) reliability, 2) the impact of DIA conditions on the diffusion of these additives and device WF_{eff}, 3) the impact of TiN metal gate deposition conditions on its microstructure in the context of reducing threshold voltage (V_T) variability, mechanical properties and device WF_{eff}, and 4) the development of XPS (X-ray Photoelectron Spectroscopy) under bias technique to determine the band energies of the MOSFET HKMG stack.

In the work we have focused on building an understanding of the working of the MOSFET device used in this thesis and its performance parameters, i.e. the effective work function (WF_{eff}) and the impact of charges and dipoles in the gate stack on these parameters. We have presented a simplified process flow, used to fabricate the different MOS devices analyzed in this work. The importance of TiN microstructure on MOSFET V_T variability has been shown. We have discussed on the current understanding and the mechanisms associated to the reliability phenomena studied in this work. Various techniques to characterize electrical, physical and mechanical properties of the gate stack have been presented and reviewed, such as the X-Ray Fluorescence (XRF), X-Ray Diffraction (XRD) and X-ray Photoelectron Spectroscopy (XPS).

In the third chapter, we have investigated the reliability effects (NBTI, PBTI and TDDB) of La and Al addition into the HKMG stack, used for threshold voltage adjustment in FDSOI gate dielectrics [PK-1]. For the BTI studies, the V_T shift follows, as expected, a power law in time and increases with the stress voltage amplitude. In literature, the latest mechanisms for BTI propose its dependence on the oxide field (E_{OX}), instead of the gate voltage (Vg). Thus, we have primarily assumed an E_{OX} dependence for BTI degradation and have analyzed our results accordingly. We have also made a distinction between device degradation at the same E_{OX} and the same Vg, by comparing device degradation and lifetime results at both conditions. Moreover, in literature NBTI and PBTI mechanism are expected to be influenced by the defects inside the interlayer and the high-k respectively. Therefore, we have explained our results at the same E_{OX} by the creation or passivation of these defects. Comparison at the same E_{OX} shows that La causes significant enhancement of intrinsic NBTI and only a slight reduction of PBTI device lifetime. On the contrary, Al causes significant enhancement of PBTI and only slight reduction in case of NBTI device lifetime. We have explained the NBTI results by a possible defect creation by La and negligible impact of Al in the interlayer SiO_2. For PBTI, we linked the results to the likely passivation of oxygen vacancies by La and defect creation by Al in HfON high-k. On the other hand, comparison at the same gate voltage shows no effect of La on NBTI and no clear effect on PBTI, contrary to the results at the same E_{OX}. We explained these different results by our assumption of BTI

dependence on E_{OX}: at the same Vg, BTI effect caused due to increasing La dose is influenced by the decreasing E_{OX} due to dipole formation by La, which leads to an almost direct compensation on NBTI but on PBTI there seems to be a more complicated relationship. Effect of Al on NBTI and PBTI is quite similar if we analyze it at the same Vg or at the same oxide field due to the relatively smaller V_T change by Al addition. We have also studied the effect of La and Al addition on TDDB reliability or device breakdown time (T_{BD}). In all cases, as expected, the breakdown times fit the Weibull distribution over a large range of time and the average time to breakdown (T_{BD}) follows a power law in voltage. Al addition in PMOS devices increases the device T_{BD}. We have correlated this increase in T_{BD} to the decrease in the leakage current with Al addition, which in turn seems responsible for lower defect creation in the gate, similar to mechanisms reported in literature. On the other hand, we found a non-monotonous effect of La addition on T_{BD} of PMOS devices, T_{BD} decreases with initial La addition but then increases for a higher La dose. Like Al addition, we correlated this effect by the variation of leakage current with La dose. In the case of La addition inside NMOS devices, the T_{BD} first decreases and then saturates due to La addition, but unlike NMOS devices this behavior could not be explained by the leakage current variation. Similar to some studies in literature on NMOS, we have been able to explain these result by the SILC measurements, which have shown that defect generation is higher with initial La addition but it saturates for further La addition. It is important to note that classically La is used in NMOS, and there seems to be no drastic degradation for PBTI but an impact only on TDDB. Moreover, it will be a problem if La is used in PMOS in future technologies due to degradation of both NBTI and TDDB. Similarly, Al is classically used in PMOS and we only observed a slight degradation on NBTI and moreover a large positive impact on TDDB, so it will not be a big issue for PMOS. However, it will be a problem if Al is used in future NMOS devices. Lastly, T_{BD} mechanism seems to depend on the use of pedestal TiN and in this case, it does not have a monotonous behaviour with La dose.

The above reliability studies were conducted on HfON high-k. In a future work, it would be interesting to study the impact of these additives on HfSiON, to have additional results and to improve our understanding of the degradation behavior. Moreover, due to limited number of devices, TDDB studies with Al were conducted only on NMOS, investigations could be done also on PMOS devices.

In the fourth chapter, we have investigated the effect of TiN metal gate RF-PVD deposition conditions (chamber pressure, RF power and substrate temperature) on the grain orientation, grain size, mechanical stress, resistivity, effective workfunction (WF_{eff}) and wafer-level non-uniformity [PK-2]. First, we have measured the variation of TiN deposition rate with its deposition parameters. For CMOS process ultra-thin TiN metal gate films are required, and so information on the deposition rate becomes very important for a good thickness control of these TiN films. According to literature, a small grain size of metal gate or a unique orientation might have an impact on reducing threshold voltage variability. Thus, we have evidenced that the grain size and orientation are significantly modulated by varying RF power and chamber pressure during TiN RF-PVD process, and grain size as low as 5.2 nm has been obtained. Moreover, a unique TiN crystal orientation have been obtained for the first time, for few optimized process conditions [PK-2][PK-6]. We have validated this unique grain orientation process for industrial applications, by showing

that the sheet resistance values and wafer non-uniformity in grain size and thickness are within or even lower than the process of reference. We have also investigated the modulation of MOS WF_{eff} due to a change in TiN gate process. We found that RF power and pressure changes the WF_{eff}, but the modulation in metal workfunction (WF_M) value could not be extracted from WF_{eff} due to a bigger change in the dipole at SiO_2/high-k interface. However, we showed that WF_{eff} variations were well correlated to the Ti/N ratio change, suggesting an effect related to this dipole. Lastly, we have shown that ASTAR technique is capable to map and image all crystal orientations present in the TiN metal layer with an excellent accuracy, which was not possible with XRD. The results on the grain size and relative amount of crystal orientations are quite similar to as obtained from XRD measurements. In a future work, TiN film with a unique orientation could be measured by ASTAR technique, to get information on all its grain orientations and so to measure the true crystallinity.

In the fifth chapter, we have investigated the impact of the drive-in-anneal (DIA) conditions, of the sacrificial gate stack, on the diffusion of La and Al additives and their impact on the WF_{eff}. We observed that for both the additives, the diffused dose changes significantly with temperature and follows the Arrhenius law of exponential behavior in the DIA temperature, and a power law in DIA time. Moreover, the diffusion of these additives, in steps other than the DIA, is significant and is taken into account to correct the dose. Thus, we have defined a methadology to correct the experimental dose and with this correction, we observed an increase in the temperature activation energy. The variation of dose with time does not necessarily follow the ideal power law, and in the case of Al the dose even saturates for higher DIA times. This non-ideal behaviour with time is in accordance with some results in literature, we have attributed this behaviour due to the ultrathin and bilayer gate stack, chemical reaction of dopants the gate oxides or due to formation of AlN clusters in them. The variation with time power law exponent. For both La and Al, the diffused dose in HfON is higher than in HfSiON, probably due to a lower mass density and higher number of defects in HfON. Moreover, in both HfON and HfSiON, diffused dose of Al is higher compared to La, which might be due by the smaller size of Al atoms compared to La. Dopant diffusion increases as the activation energy of diffusion decreases. We have correlated this increase in diffused dose with the activation energy, which can explain the differences in diffused dose between the different additives and high-k materials. Then, the effective workfunction shift with the DIA conditions has been correlated with the diffused additive doses. For La addition, on both HfON and HfSiON, the workfunction changes linearly with the diffused dose, quite similar to studies done earlier, irrespective of the method used to introduce the dose (by varing DIA time or temperature). In the case of Al addition, the dose follows a linear behaviour for small doses but then deviates and saturate for higher doses. Finally, these results show that for the same diffused dose, the electrical impact (WF_{eff} shift) of Al or La dipole at the SiO_2/high-k interface on HfON is stronger than on HfSiON. Moreover, we also show that this impact of La dipole is stronger than that of Al on both the high-k materials.

We have presented many results on the diffusion of La and Al on different high-k materials and also correlated it to the WF_{eff}, but we were limited by the measurement of only the total dose inside the HKMG stack but not its profile. However, a complete diffusion study can be conducted in order

to better understand the diffusion of these additives. This can involve SIMS measurements to evaluate additive dose profile inside the gate stack and also taking into account the diffusion of additives without the annealing step, by which the diffusion equations can be properly formulated.

In the final chapter, we have developed a robust and accurate XPS under bias technique, able to determine the band energies relative positions of the HKMG stack layers of MOS devices [PK-3][PK-5]. Dipoles and interface states inside the HKMG stack can modulate device electrical parameters, specially at the SiO_2/high-k interface arising due to La or Al (chapter 4) or TiN process change (chapter 5). Thus, obtaining information of these dipoles and that of interface states at the SiO_2/Si interface becomes very important, which is possible from band energy analysis of the gate stack. In previous studies, researchers have tried to obtain information on band energies through XPS, but faced many experimental challenges such as proper biasing of the devices. In this work, we have presented new test structures that enable to perform correct XPS under bias measurements. We have identified and modelled the biasing issues related to these test structures, to calculate the effective bias across the device. Addtionally, we have identified the experimental issues related to the XPS equipment such as a method to making sure that the XPS beam irradiates only the device under test. This was done by defining a strategy to correctly localize the device and knowing the correct XPS beam size. Moreover, we have been able to significantly reduce the issue concerning a large parasitic substrate resistance, that presented a hindrance in performing accurate XPS measurements. Solving these experimental problem then allowed us to validate the XPS under bias technique by first defining a robust fitting methodology for various XPS signals, and then comparing XPS binding energy shifts with what is expected from our electrical model. This correlation, done for the first time between two previosuly unrelated measurements, thus validated this technique. Then, we have applied this technique to measure gate stack electrical properties by first defining a strategy to calculate the interface states present at the SiO_2/Si interface. Finally, this technique has been applied to analyze the origin of effective work function (WF_{eff}) shift related to La and Al additives or to TiN thickness modulation. It has been shown for the first time, by relative band energy measurements, that in all these cases the dipole modulation at the SiO_2/High-k interface is responsible for these observed WF_{eff} changes. Moreover, it has been shown that band energies at flatband condition can be obtained from measurements at zero bias condition. This technique has the potential to characterize devices beyond CMOS, and for other technologies.

In this work, we have conducted the most important part regarding XPS under bias technique, that was solving the biasing and experimental issues related to this technique and its development and validation. We have also localized and quantified dipoles formed due to additives and TiN thickness modulation. For future work, this technique can be used to localize the dipoles related to SiGe channel, modulation of TiN intrinsic workfunction by its process conditions, and the origin of the roll-off effect.

[1] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices. John Wiley & Sons, 3rd edition*. 2007.

[2] R. Jha, J. Gurganos, Y. H. Kim, R. Choi, J. Lee, and V. Misra, "A capacitance-based methodology for work function extraction of metals on high-/spl kappa/;," *IEEE Electron Device Lett.*, vol. 25, no. 6, pp. 420–423, 2004.

[3] B. E. Deal, "The current understanding of charges in the thermally oxidized silicon structure," *J. Electrochem. Soc.*, vol. 121, p. 198C– 205C, 1974.

[4] B. E. Deal, "Standardized terminology for oxide charges associated with thermally oxidized silicon," *IEEE Trans. Electron Devices*, vol. 27, no. 3, pp. 606–608, 1980.

[5] M. Charbonnier *et al.*, "Measurement of Dipoles/Roll-Off /Work Functions by Coupling CV and IPE and Study of Their Dependence on Fabrication Process," *Electron Devices, IEEE Trans.*, vol. 57, pp. 1809–1819, Sep. 2010.

[6] C. Suarez-Segovia *et al.*, "Effective work function engineering by sacrificial lanthanum diffusion on HfON-based 14 nm NFET devices," in *2015 45th European Solid State Device Research Conference (ESSDERC)*, 2015, pp. 246–249.

[7] J. K. Schaeffer, L. R. C. Fonseca, S. B. Samavedam, Y. Liang, P. J. Tobin, and B. E. White, "Contributions to the effective work function of platinum on hafnium dioxide," *Appl. Phys. Lett.*, vol. 85, no. 10, pp. 1826–1828, Sep. 2004.

[8] J. K. Schaeffer *et al.*, "Application of group electronegativity concepts to the effective work functions of metal gate electrodes on high-κ gate oxides," *Microelectron. Eng.*, vol. 84, no. 9, pp. 2196–2200, 2007.

[9] H. N. Alshareef *et al.*, "Composition dependence of the work function of Ta1−xAlxNy metal gates," *Appl. Phys. Lett.*, vol. 88, no. 7, p. 72108, Feb. 2006.

[10] T. C. Chen *et al.*, "Band-Edge High-Performance High-k/Metal Gate n-MOSFETs Using Cap Layers Containing Group IIA and IIIB Elements with Gate-First Processing for 45 nm and Beyond," in *2006 Symposium on VLSI Technology, 2006. Digest of Technical Papers.*, 2006, pp. 178–179.

[11] K. Kakushima *et al.*, "Origin of flat band voltage shift in HfO2 gate dielectric with La2O3 insertion," *Solid State Electron.*, vol. 52, no. 9, pp. 1280–1284, 2008.

[12] Tseng, H. H., Kirsch, P., Park, C. S., Bersuker, G., et al., "The progress and challenges of threshold voltage control of high-k/metal-gated devices for advanced technologies," *Microelectron. Eng*, vol. 86, pp. 1722–1727, 2009.

[13] K. Shiraishi *et al.*, "Physics in Fermi level pinning at the polySi/Hf-based high-k oxide interface," in *Digest of Technical Papers. 2004 Symposium on VLSI Technology, 2004.*, 2004, pp. 108–109.

[14] E. Cartier *et al.*, "Role of oxygen vacancies in V_{FB}/V_t stability of pFET metals on HfO_2," in *Digest of Technical Papers. 2005 Symposium on VLSI Technology, 2005.*, 2005, pp. 230–231.

[15] K. Kita and A. Toriumi, "Origin of electric dipoles formed at high-k/SiO2 interface," *Appl. Phys. Lett.*, vol. 94, no. 13, p. 132902, Mar. 2009.

[16] G. Bersuker *et al.*, "Origin of the Flatband-Voltage Roll-Off Phenomenon in Metal/High- k Gate Stacks," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2047–2056, 2010.

[17] C. Leroux *et al.*, *Investigating doping effects on high-κ metal gate stack for effective work function engineering*, vol. 88. 2013.

[18] D. A. Buchanan, "Scaling the Gate Dielectric: Materials, Integration, and Reliability," *IBM J.*

Res. Dev., vol. 43, no. 3, pp. 245–264, May 1999.
[19] S. V Hattangady, H. Niimi, and G. Lucovsky, "Controlled nitrogen incorporation at the gate oxide surface," *Appl. Phys. Lett.*, vol. 66, no. 25, pp. 3495–3497, Jun. 1995.
[20] D. J. DiMaria and J. H. Stathis, "Trapping and trap creation studies on nitrided and reoxidized-nitrided silicon dioxide films on silicon," *J. Appl. Phys.*, vol. 70, pp. 1500–1509, 1991.
[21] X. Meng *et al.*, "Atomic Layer Deposition of Silicon Nitride Thin Films: A Review of Recent Progress, Challenges, and Outlooks," *Materials (Basel).*, vol. 9, no. 12, p. 1007, Dec. 2016.
[22] E. Gusev, *Impact of high-k properties on MOSFET electrical characteristics*. Springer Netherlands, 2006.
[23] S. Guha and V. Narayanan, "High-κ/Metal Gate Science and Technology," *Annu. Rev. Mater. Res.*, vol. 39, no. 1, pp. 181–202, Jul. 2009.
[24] J. S. Jeon, C. Wong, J. Gray, H. S. Kim, and and B. Ogle, "Preparation of sub 20a thick ultra-thin stack gate dielectrics by in-situ RTCVD processes," *Electrochem. Soc. Proc.*, pp. 183–188, 2001.
[25] J. Robertson, "Band offsets of wide-band-gap oxides and implications for future electronic devices," *J. Vac. Sci. Technol. B Microelectron. Nanom. Struct. Process. Meas. Phenom.*, vol. 18, no. 3, pp. 1785–1791, May 2000.
[26] M. S. M. Koike, T. Ino, Y. Kamimuta, M. Koyama, Y. Kamata and and Y. T. Y. Mitani, A. Nishiyama, "Effect of Hf-N bond on properties of thermally stable amorphous HfSiON and applicability of this material to sub-5o nm technology node LSIs," in *IEDM, Technical Digest*, 2003, p. 4.7.1–4.7.4.
[27] M. R. Visokay, J. J. Chambers, A. L. P. Rotondaro, A. Shanware, and L. Colombo, "Application of HfSiON as a gate dielectric material," *Appl. Phys. Lett.*, vol. 80, no. 17, pp. 3183–3185, Apr. 2002.
[28] M. Bohr, R. Chau, T. Ghani, and K. Mistry, "The high-k solution," *IEEE Spectr.*, pp. 30–35, 2007.
[29] H. Shimada, Y. Hirano, T. Ushiki, K. Ino, and Ohmi.T, "Tantalum-gate thin-film SOI nMOS and pMOS for low-power applications," *IEEE Trans. Electron Devices*, vol. 44, pp. 1903–1907, 1997.
[30] Y.-C. Yeo *et al.*, *Dual-metal gate CMOS technology with ultrathin silicon nitride gate dielectric*, vol. 22. 2001.
[31] L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu, "Gate Length Scaling and Threshold Voltage Control of Double-Gate MOSFETs," *IEDM, Tech. Dig.*, pp. 719 – 722, 2000.
[32] O. Weber *et al.*, "Work-function engineering in gate first technology for multi-VT dual-gate FDSOI CMOS on UTBOX," *2010 Int. Electron Devices Meet.*, p. 3.4.1-3.4.4, 2010.
[33] E. P. Gusev, V. Narayanan, and M. M. Frank, "Advanced High-&k Dielectric Stacks with polySi and Metal Gates: Recent Progress and Current Challenges," *IBM J. Res. Dev.*, vol. 50, no. 4/5, pp. 387–410, Jul. 2006.
[34] E. K.-E. *et al.*, "Diffusion Barrier Deposition on a Copper Surface by Atomic Layer Deposition," *Chem. Vap. Depos.*, vol. 8, no. 4, pp. 149–153, Jul. 2002.
[35] M. A. L. and A. J. Lichtenberg, *Principles of plasma discharges and materials processing*. ohn Wiley & Sons, 2009.
[36] "Applied Endura2 RF PVD chamber manual," 2009.

[37] S. Baudot, “Elaboration et caracterisation des grilles metalliques pour les technologies CMOS 32/28 nm a base de dielectrique haute permittivite,” Universite de Grenoble, 2012.
[38] C.-H. Jan *et al.*, *A 32nm SoC platform technology with 2nd generation high-k/metal gate transistors optimized for ultra low power, high performance, and high density product applications*. 2010.
[39] O. Weber *et al.*, *14nm FDSOI technology for high speed and energy efficient applications*. 2014.
[40] C. S. Segovia, “Electrical and physicochemical characterization of metal gate processes for work function modulation and reduction of local VTH variability in 14FDSOI technologies,” 2016.
[41] N. Planes *et al.*, “28nm FDSOI technology platform for high-speed low-voltage digital applications,” in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 133–134.
[42] S. Guha, E. Cartier, M. A. Gribelyuk, N. A. Bojarczuk, and M. C. Copel, “Atomic beam deposition of lanthanum- and yttrium-based oxide thin films for gate dielectrics,” *Appl. Phys. Lett.*, vol. 77, no. 17, pp. 2710–2712, Oct. 2000.
[43] D. J. Lichtenwalner *et al.*, “Lanthanum silicate gate dielectric stacks with subnanometer equivalent oxide thickness utilizing an interfacial silica consumption reaction,” *J. Appl. Phys.*, vol. 98, no. 2, p. 24314, Jul. 2005.
[44] S. Kamiyama *et al.*, *Band edge gate first HfSiON/metal gate n-MOSFETs using ALD-La(2)O(3) cap layers scalable to EOT=0.68 nm for hp 32 nm bulk devices with high performance and reliability*. 2008.
[45] K. Y. and K. K. and A. Toriumi, “Study of La-Induced Flat Band Voltage Shift in Metal/HfLaO x /SiO 2 /Si Capacitors,” *Jpn. J. Appl. Phys.*, vol. 46, no. 11R, p. 7251, 2007.
[46] T. Ando *et al.*, “Understanding mobility mechanisms in extremely scaled HfO2 (EOT 0.42 nm) using remote interfacial layer scavenging technique and Vt-tuning dipoles with gate-first process,” in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.
[47] C. Suarez-Segovia, C. Leroux, F. Domengie, and G. Ghibaudo, “Quantitative Analysis of La and Al Additives Role on Dipole Magnitude Inducing Vt Shift in High-K/Metal Gate Stack,” *IEEE Electron Device Lett.*, vol. 38, no. 3, pp. 379–382, 2017.
[48] H. K. Kim, S. Y. Lee, I. H. Yu, T. J. Park, R. Choi, and C. S. Hwang, “Gate Engineering in TiN/La/TiN and TiLaN Metal Layers on Atomic-Layer-Deposited HfO2/Si,” *IEEE Electron Device Lett.*, vol. 33, no. 7, pp. 955–957, 2012.
[49] E. P. Gusev, M. Copel, E. Cartier, I. J. R. Baumvol, C. Krug, and M. A. Gribelyuk, “High-resolution depth profiling in ultrathin Al2O3 films on Si,” *Appl. Phys. Lett.*, vol. 76, no. 2, pp. 176–178, Jan. 2000.
[50] K. Choi, H.-C. Wen, G. Bersuker, R. Harris, and B. H. Lee, “Mechanism of flatband voltage roll-off studied with Al2O3 film deposited on terraced oxide,” *Appl. Phys. Lett.*, vol. 93, no. 13, p. 133506, Sep. 2008.
[51] Y. Kamimuta *et al.*, “Comprehensive Study of V<inf>FB</inf> Shift in High-k CMOS - Dipole Formation, Fermi-level Pinning and Oxygen Vacancy Effect,” in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 341–344.
[52] M. Charbonnier *et al.*, “Investigation of mechanisms shifting metal effective workfunction towards P+ for various Al incorporation scenarii,” in *Proceedings of 2010 International Symposium on VLSI Technology, System and Application*, 2010, pp. 84–85.

[53] K. Iwamoto *et al.*, “Re-examination of Flat-Band Voltage Shift for High-k MOS Devices,” in *2007 IEEE Symposium on VLSI Technology*, 2007, pp. 70–71.
[54] S. Baudot *et al.*, “Understanding reversal effects of metallic aluminum introduced in HfSiON/TiN PMOSFETs,” *Microelectron. Eng.*, vol. 88, no. 7, pp. 1305–1308, 2011.
[55] C. Suarez-Segovia *et al.*, “Effective work function modulation by sacrificial gate aluminum diffusion on HfON-based 14nm NMOS devices,” *Microelectron. Eng.*, vol. 147, pp. 113–116, 2015.
[56] R. Smoluchowski, “Anisotropy of the Electronic Work Function of Metals,” *Phys. Rev.*, vol. 60, no. 9, pp. 661–674, Nov. 1941.
[57] N. D. Lang and W. Kohn, “Theory of Metal Surfaces: Charge Density and Surface Energy,” *Phys. Rev. B*, vol. 1, no. 12, pp. 4555–4568, Jun. 1970.
[58] H. F. Dadgour, K. Endo, V. K. De, and K. Banerjee, “Grain-Orientation Induced Work Function Variation in Nanoscale Metal-Gate Transistors-Part I: Modeling, Analysis, and Experimental Validation,” *IEEE Trans. Electron Devices*, vol. 57, no. 10, pp. 2504–2514, 2010.
[59] N. D. Lang, *Theory of Metal Surfaces: Charge Density and Surface Energy*, vol. 1. 1970.
[60] T. Matsukawa *et al.*, “Suppressing Vt and Gm variability of FinFETs using amorphous metal gates for 14 nm and beyond,” in *2012 International Electron Devices Meeting*, 2012, p. 8.2.1-8.2.4.
[61] K. Ohmori *et al.*, “Impact of additional factors in threshold voltage variability of metal/high-k gate stacks and its reduction by controlling crystalline structure and grain size in the metal gates,” in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1–4.
[62] H. Dadgour, K. Endo, V. De, and K. Banerjee, “Modeling and analysis of grain-orientation effects in emerging metal-gate devices and implications for SRAM reliability,” in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1–4.
[63] S. Markov, A. S. M. Zain, B. Cheng, and A. Asenov, “Statistical variability in scaled generations of n-channel UTB-FD-SOI MOSFETs under the influence of RDF, LER, OTF and MGG,” in *2012 IEEE International SOI Conference (SOI)*, 2012, pp. 1–2.
[64] A. Asenov, “Simulation of Statistical Variability in Nano MOSFETs,” in *2007 IEEE Symposium on VLSI Technology*, 2007, pp. 86–87.
[65] F. Boeuf, M. Sellier, A. Farcy, and T. Skotnicki, “An Evaluation of the CMOS Technology Roadmap From the Point of View of Variability, Interconnects, and Power Dissipation,” *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1433–1440, 2008.
[66] A. Vandooren *et al.*, “Mixed-signal performance of sub-100nm fully-depleted SOI devices with metal gate, high K (HfO/sub 2/) dielectric and elevated source/drain extensions,” in *IEEE International Electron Devices Meeting 2003*, 2003, p. 11.5.1-11.5.3.
[67] C. Fenouillet-Beranger *et al.*, “Fully-depleted SOI technology using high-k and single-metal gate for 32 nm node LSTP applications featuring 0.179 μm^2 6T-SRAM bitcell,” in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 267–270.
[68] Y. Morita *et al.*, “Smallest Vt variability achieved by intrinsic silicon on thin BOX (SOTB) CMOS with single metal gate,” in *2008 Symposium on VLSI Technology*, 2008, pp. 166–167.
[69] A. Yagishita *et al.*, “Improvement of threshold voltage deviation in damascene metal gate transistors,” *IEEE Trans. Electron Devices*, vol. 48, no. 8, pp. 1604–1611, 2001.
[70] V. Huard, M. Denais, and C. Parthasarathy, “NBTI degradation: From physical mechanisms

to modelling," *Microelectron. Reliab.*, vol. 46, no. 1, pp. 1–23, Jan. 2006.
[71] Y. Miura and Y. Matukura, "Investigation of silicon–silicon dioxide interface using MOS structure," *Jpn J Appl Phys*, vol. 5, p. 180, 1966.
[72] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, *Characteristics of the Surface State Charge QHH of Thermally Oxidized Silicon*, vol. 114. 1967.
[73] D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," *J. Appl. Phys.*, vol. 94, no. 1, pp. 1–18, Jun. 2003.
[74] V. Reddy *et al.*, "Impact of negative bias temperature instability on digital circuit reliability," in *2002 IEEE International Reliability Physics Symposium. Proceedings. 40th Annual (Cat. No.02CH37320)*, 2002, pp. 248–254.
[75] S. Chakravarthi, A. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, "A comprehensive framework for predictive modeling of negative bias temperature instability," in *2004 IEEE International Reliability Physics Symposium. Proceedings*, 2004, pp. 273–282.
[76] D. K. Schroder, "Negative bias temperature instability: What do we understand?," *Microelectron. Reliab.*, vol. 47, no. 6, pp. 841–852, 2007.
[77] K. Saminadayar and J. C. Pfister, "Evolution of surface-states density of Si/wet thermal SiO2 interface during bias-temperature treatment," *Solid. State. Electron.*, vol. 20, no. 11, pp. 891–896, 1977.
[78] T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, "A two-stage model for negative bias temperature instability," in *2009 IEEE International Reliability Physics Symposium*, 2009, pp. 33–44.
[79] B. Kaczer *et al.*, "Recent trends in bias temperature instability," *J. Vac. Sci. Technol. B*, vol. 29, no. 1, p. 01AB01, Jan. 2011.
[80] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectron. Reliab.*, vol. 45, no. 1, pp. 71–81, 2005.
[81] K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," *J. Appl. Phys.*, vol. 48, no. 5, pp. 2004–2014, May 1977.
[82] S. Ogawa and N. Shiono, "Generalized diffusion-reaction model for the low-field charge-buildup instability at the Si-SiO2 interface," *Phys. Rev. B*, vol. 51, no. 7, pp. 4218–4230, Feb. 1995.
[83] B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, "Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification," in *2005 IEEE International Reliability Physics Symposium, 2005. Proceedings. 43rd Annual.*, 2005, pp. 381–387.
[84] Lee *et al.*, "Negative Bias Temperature Instability Dependence on Dielectric Thickness and Nitrogen Concentration in Ultra-scaled HfSiON Dielectric/TiN Gate Stacks," *Jpn. J. Appl. Phys.*, vol. 45, no. 4S, p. 2945, 2006.
[85] V. Huard *et al.*, "New characterization and modeling approach for NBTI degradation from transistor to product level," in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 797–800.
[86] T. Grasser, B. Kaczer, and W. Goes, "An energy-level perspective of bias temperature instability," in *2008 IEEE International Reliability Physics Symposium*, 2008, pp. 28–38.

[87] B. Kaczer *et al.*, “Ubiquitous relaxation in BTI stressing—New evaluation and insights,” in *2008 IEEE International Reliability Physics Symposium*, 2008, pp. 20–27.
[88] H. Reisinger, T. Grasser, W. Gustin, and C. Schlünder, “The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress,” in *2010 IEEE International Reliability Physics Symposium*, 2010, pp. 7–15.
[89] D. M. Fleetwood *et al.*, “Unified model of hole trapping, 1/f noise, and thermally stimulated current in MOS devices,” *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 2674–2683, 2002.
[90] W. Goes *et al.*, “Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence,” *Microelectron. Reliab.*, vol. 87, pp. 286–320, 2018.
[91] X. Garros, J. Mitard, C. Leroux, G. Reimbold, and F. Boulanger, “In Depth Analysis of VT Instabilities in HFO2 Technologies by Charge Pumping Measurements and Electrical Modeling,” in *2007 IEEE International Reliability Physics Symposium Proceedings. 45th Annual*, 2007, pp. 61–66.
[92] G. Bersuker *et al.*, *Mechanism of Electron Trapping and Characteristics of Traps in HfO2 Gate Stacks*, vol. 7. 2007.
[93] G. Reimbold, J. Mitard, X. Garros, C. Leroux, G. Ghibaudo, and F. Martin, “Initial and PBTI-induced traps and charges in Hf-based oxides/TiN stacks,” *Microelectron. Reliab.*, vol. 47, no. 4, pp. 489–496, 2007.
[94] J. Mitard *et al.*, “Large-Scale Time Characterization and Analysis of PBTI In HFO2/Metal Gate Stacks,” in *2006 IEEE International Reliability Physics Symposium Proceedings*, 2006, pp. 174–178.
[95] X. Garros *et al.*, “PBTI mechanisms in La containing Hf-based oxides assessed by very Fast IV measurements,” in *2010 International Electron Devices Meeting*, 2010, p. 4.6.1-4.6.4.
[96] K. Onishi *et al.*, *Bias-temperature instabilities of polysilicon gate HfO2 MOSFETs*, vol. 50. 2003.
[97] G. Ribes *et al.*, “Review on high-k dielectrics reliability issues,” *IEEE Trans. Device Mater. Reliab.*, vol. 5, no. 1, pp. 5–19, 2005.
[98] C. Shen *et al.*, “Negative U traps in HfO_2 gate dielectrics and frequency dependence of dynamic BTI in MOSFETs,” in *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.*, 2004, pp. 733–736.
[99] S. Bruyère, “Etude des mécanismes de dégradation et de défaillance des oxydes ultra minces, application à la fiabilité des technologies CMOS sub-0.25um,” Thèse de l’Institut National Polytechnique de Grenoble, 2000.
[100] C. Leroux, D. Blachier, O. Briere, and G. Reimbold, “Light emission microscopy for thin oxide reliability analysis,” *Microelectron. Eng.*, vol. 36, no. 1, pp. 297–300, 1997.
[101] E. Wu, J. Stathis, and L.-K. Han, *Ultra-thin oxide reliability for ULSI application*, vol. 15. 2000.
[102] J. Suñé, I. Placencia, N. Barniol, E. Farrés, F. Martín, and X. Aymerich, “On the breakdown statistics of very thin SiO2 films,” *Thin Solid Films*, vol. 185, no. 2, pp. 347–362, 1990.
[103] R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, “A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides,” in *Proceedings of International Electron Devices Meeting*, 1995, pp. 863–866.
[104] R. Degraeve *et al.*, “New insights in the relation between electron trap generation and the statistical properties of oxide breakdown,” *IEEE Trans. Electron Devices*, vol. 45, no. 4, pp.

904–911, 1998.
[105] J. H. Stathis, “Percolation models for gate oxide breakdown,” *J. Appl. Phys.*, vol. 86, no. 10, pp. 5757–5766, Nov. 1999.
[106] J. Sune, E. Y. Wu, D. Jimenez, R. P. Vollertsen, and E. Miranda, “Understanding soft and hard breakdown statistics, prevalence ratios and energy dissipation during breakdown runaway,” in *International Electron Devices Meeting. Technical Digest (Cat. No.01CH37224)*, 2001, p. 6.1.1-6.1.4.
[107] M. A. Alam *et al.*, “An anode hole injection percolation model for oxide breakdown-the ‘doom’s day’ scenario revisited,” in *International Electron Devices Meeting 1999. Technical Digest (Cat. No.99CH36318)*, 1999, pp. 715–718.
[108] E. Y. Wu and J. Suñé, “Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability,” *Microelectron. Reliab.*, vol. 45, no. 12, pp. 1809–1834, 2005.
[109] P. E. Nicollian, A. T. Krishnan, C. A. Chancellor, and R. B. Khamankar, “The Traps that cause Breakdown in Deeply Scaled SiON Dielectrics,” in *2006 International Electron Devices Meeting*, 2006, pp. 1–4.
[110] J. H. Stathis, “Physical and predictive models of ultrathin oxide reliability in CMOS devices and circuits,” *IEEE Trans. Device Mater. Reliab.*, vol. 1, no. 1, pp. 43–59, 2001.
[111] P. E. Nicollian, A. T. Krishnan, C. Bowen, S. Chakravarthi, C. A. Chancellor, and R. B. Khamankar, “The roles of hydrogen and holes in trap generation and breakdown in ultra-thin SiON dielectrics,” in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, 2005, pp. 392–395.
[112] J. W. McPherson, “Extended Mie-Grüneisen molecular model for time dependent dielectric breakdown in silica detailing the critical roles of O–Si□O3 tetragonal bonding, stretched bonds, hole capture, and hydrogen release,” *J. Appl. Phys.*, vol. 99, no. 8, p. 83501, Apr. 2006.
[113] K. Schuegraf and C. hu, *Effects of Temperature and Defects on Breakdown Lifetime of Thin SiO2 at Very Low Voltages*, vol. 41. 1994.
[114] D. J. DiMaria and E. Cartier, “Mechanism for stress-induced leakage currents in thin silicon dioxide films,” *J. Appl. Phys.*, vol. 78, no. 6, pp. 3883–3894, Sep. 1995.
[115] M. Rafik, “Caractérisation et Modélisation de la Fiabilité des Transistors Avancés à Diélectriques de Haute Permittivité et à Grille Métallique,” 2008.
[116] P. Avouris *et al.*, “Breaking individual chemical bonds via STM-induced excitations,” *Surf. Sci.*, vol. 363, no. 1, pp. 368–377, 1996.
[117] K. Stokbro, B. Hu, C. Thirstrup, and X. C. Xie, *First principles theory of inelastic currents in a scanning tunneling microscope*, vol. 58. 1998.
[118] E. Cartier and A. Kerber, “Stress-induced leakage current and defect generation in nFETs with HfO2/TiN gate stacks during positive-bias temperature stress,” in *2009 IEEE International Reliability Physics Symposium*, 2009, pp. 486–492.
[119] T. Fenfen *et al.*, *TDDB characteristic and breakdown mechanism of ultra-thin SiO2/HfO2 bilayer gate dielectrics*, vol. 35. 2014.
[120] C. Leroux, F. Allain, A. Toffoli, G. Ghibaudo, and G. Reimbold, “Automatic statistical full quantum analysis of C-V and I-V characteristics for advanced MOS gate stacks,” *Microelectron. Eng.*, vol. 84, no. 9, pp. 2408–2411, 2007.
[121] J. Maserjian, G. Petersson, and C. Svensson, “Saturation capacitance of thin oxide MOS

structures and the effective surface density of states of silicon," *Solid. State. Electron.*, vol. 17, no. 4, pp. 335–339, 1974.
[122] M. Charbonnier, C. Leroux, F. Allain, A. Toffoli, G. Ghibaudo, and G. Reimbold, "Automatic full quantum analysis of CV measurements for bulk and SOI devices," *Microelectron. Eng.*, vol. 88, no. 12, pp. 3404–3406, 2011.
[123] C. Leroux, "Contribution à la caractérisation electrique et à l'analyse des dispositifs utilisés en microélectronique. Habilitation à diriger des recherches, Institut National Polytechnique de Grenoble," 2010.
[124] C. Leroux, G. Ghibaudo, and G. Reimbold, "Accurate determination of flat band voltage in advanced MOS structure," *Microelectron. Reliab.*, vol. 47, no. 4, pp. 660–664, 2007.
[125] F. M. Smits, "Measurement of sheet resistivities with the four-point probe," *Bell Syst. Tech. J.*, vol. 37, no. 3, pp. 711–718, 1958.
[126] C. M. Friend, "Miscellanea. Physics at Surfaces. By A. Zangwill. Cambridge University Press, Cambridge 1988. xiii, 454 pp., hardcover £ 40.00.—ISBN 0-521-32147-6," *Angew. Chemie*, vol. 100, no. 10, pp. 1463–1464, Sep. 2018.
[127] M. P. Seah, *ISO 15472:2001. Surface Chemical Analysis—X-ray Photoelectron Spectrometers—Calibration of Energy Scales. Surf. Interface Anal. 2001; 31: 721*, vol. 31. 2001.
[128] P. O. Gartland, S. Berge, and B. J. Slagsvold, "Photoelectric Work Function of a Copper Single Crystal for the (100), (110), (111), and (112) Faces," *Phys. Rev. Lett.*, vol. 28, no. 12, pp. 738–739, Mar. 1972.
[129] P. A. Tipler and R. Llewellyn, *Modern Physics*, no. p. 3. W. H. Freeman, 1999.
[130] "PHI VersaProbe II Manual, Physical Electronics."
[131] S. Suzer, *XPS investigation of a Si-diode in operation*, vol. 4. 2012.
[132] "http://www.fischer-technology.com/en/united states/knowledge/methods/material-testing/x-ray-fluorescence2/." .
[133] P. Mandracci, "Master nanotech lecture on XRD." Politecnico di torino, 2013.
[134] W.Wong-Ng, H. McMurdie, B. Paretzkin, C. Hubbard, and A. Dragoo, "The JCPDS database (1998), data set number: 38-1420. Powder Diffraction 2, 200," 1987.
[135] F. Panciera, S. Baudot, K. Hoummada, M. Gregoire, M. Juhel, and D. Mangelinck, "Three-dimensional distribution of Al in high-k metal gate: Impact on transistor voltage threshold," *Appl. Phys. Lett.*, vol. 100, no. 20, p. 201909, May 2012.
[136] B. Mohamad, C. Leroux, D. Rideau, M. Haond, G. Reimbold, and G. Ghibaudo, "Robust EOT and effective work function extraction for 14 nm node FDSOI technology," in *2016 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, 2016, pp. 135–138.
[137] D. Liu and J. Robertson, "Passivation of oxygen vacancy states and suppression of Fermi pinning in HfO2 by La addition," *Appl. Phys. Lett.*, vol. 94, no. 4, p. 42904, Jan. 2009.
[138] M. Copel, S. Guha, N. Bojarczuk, E. Cartier, V. Narayanan, and V. Paruchuri, "Interaction of La2O3 capping layers with HfO2 gate dielectrics," *Appl. Phys. Lett.*, vol. 95, no. 21, p. 212903, Nov. 2009.
[139] J. W. McPherson, "Time dependent dielectric breakdown physics – Models revisited," *Microelectron. Reliab.*, vol. 52, no. 9, pp. 1753–1760, 2012.
[140] E. Cartier, B. P. Linder, V. Narayanan, and V. K. Paruchuri, "Fundamental understanding and

optimization of PBTI in nFETs with SiO2/HfO2 gate stack," in *2006 International Electron Devices Meeting*, 2006, pp. 1–4.
[141] R. O'Connor *et al.*, "SILC defect generation spectroscopy in HfSiON using constant voltage stress and substrate hot electron injection," in *2008 IEEE International Reliability Physics Symposium*, 2008, pp. 324–329.
[142] A. Chaoumead, Y. Sung, and D.-J. Kwak, *The Effects of RF Sputtering Power and Gas Pressure on Structural and Electrical Properties of ITiO Thin Film*, vol. 2012. 2012.
[143] F. M. D'Heurle and J. M. E. Harper, "Note on the origin of intrinsic stresses in films deposited via evaporation and sputtering," *Thin Solid Films*, vol. 171, no. 1, pp. 81–92, 1989.
[144] X. Pang, L. Zhang, H. Yang, K. Gao, and A. A. Volinsky, "Residual Stress and Surface Energy of Sputtered TiN Films," *J. Mater. Eng. Perform.*, vol. 24, no. 3, pp. 1185–1191, 2015.
[145] T. Matsukawa *et al.*, "Comprehensive analysis of variability sources of FinFET characteristics," in *2009 Symposium on VLSI Technology*, 2006, pp. 118–119.
[146] A. R. Brown, N. M. Idris, J. R. Watling, and A. Asenov, "Impact of Metal Gate Granularity on Threshold Voltage Variability: A Full-Scale Three-Dimensional Statistical Simulation Study," *IEEE Electron Device Lett.*, vol. 31, no. 11, pp. 1199–1201, 2010.
[147] H. Cheng and Y. Li, "Random work function variation induced threshold voltage fluctuation in 16-nm bulk FinFET devices with high-k-metal-gate material," in *2010 14th International Workshop on Computational Electronics*, 2010, pp. 1–4.
[148] F. Fillot, S. Maitrejean, F. Pierre, and B. Chenevier, "Work function tuning of TixSiyNz electrodes using partial saturation of chemisorbing surface during pulsing chemical vapor deposition," *Electrochem. Solid State Lett.*, vol. 12, pp. H272–H274, 2009.
[149] C. Adelmann *et al.*, "Thermally stable high effective work function TaCN thin films for metal gate electrode applications," *J. Appl. Physics,* vol. 105, p. 053516:1–8, 2009.
[150] J. A. Thornton, *Influence of Apparatus Geometry and Deposition Condition the Structure and Topography of Thick Sputtered Coatings*, vol. 11. 1974.
[151] L. Dong and D. J. Srolovitz, "Mechanism of texture development in ion-beam-assisted deposition," *Appl. Phys. Lett.*, vol. 75, no. 4, pp. 584–586, Jul. 1999.
[152] D. Dobrev, *Ion-beam-induced texture formation in vacuum-condensed thin metal films*, vol. 92. 1982.
[153] Y. G. Shen, "Effect of deposition conditions on mechanical stresses and microstructure of sputter-deposited molybdenum and reactively sputter-deposited molybdenum nitride films," *Mater. Sci. Eng. A*, vol. 359, no. 1, pp. 158–167, 2003.
[154] G. N. Van Wyk and H. J. Smith, "Crystalline reorientation due to ion bombardment," *Nucl. Instruments Methods*, vol. 170, no. 1, pp. 433–439, 1980.
[155] S. Hu, J. Zheng, L. Feng, W. Li, and D. Yang, "The Influence of Preparing Conditions on the Properties of CdS Polycrystalline Thin Films Prepared by Magnetron Sputtering," in *2010 Asia-Pacific Power and Energy Engineering Conference*, 2010, pp. 1–4.
[156] S.-H. Lee, R. Choi, and C. Choi, "Effects of composition and thickness of TiN metal gate on the equivalent oxide thickness and flat-band voltage in metal oxide semiconductor devices," *Microelectron. Eng.*, vol. 109, pp. 160–162, 2013.
[157] "NanoMEGAS website, www.nanomegas.com." .
[158] E. F. Rauch and L. Dupuy, "Rapid spot diffraction patterns identification through template

matching," *Arch. Met. Mater*, vol. 50, no. 1, pp. 87–99, 2005.
[159] A. VALERY, "Caractérisation de microtextures par la technique ACOM-TEM dans le cadre du développement des technologies avancées en microélectronique," University grenoble alps, 2017.
[160] "Master nanotech lecture-Diffusion." Politecnico di torino, 2013.
[161] "Introduction To Materials Science FOR ENGINEERS, Ch. 5." University of Tennessee, Dept. of Materials Science and Engineering, 2015.
[162] Z. Essa *et al.*, "Evaluation and modeling of lanthanum diffusion in TiN/La2O3/HfSiON/SiO2/Si high-k stacks," *Appl. Phys. Lett.*, vol. 101, no. 18, p. 182901, Oct. 2012.
[163] E. Martinez *et al.*, "Lanthanum diffusion in the TiN/LaOx/HfSiO/SiO2/Si stack," *Microelectron. Eng.*, vol. 88, no. 7, pp. 1349–1352, 2011.
[164] R. S. and H. Y. Y. and T. J. and S. K. and K. De Meyer, "The Study of Effective Work Function Modulation by As Ion Implantation in TiN/TaN/HfO 2 Stacks," *Jpn. J. Appl. Phys.*, vol. 46, no. 4L, p. L320, 2007.
[165] M. Kadoshima *et al.*, "Effective-Work-Function Control by Varying the TiN Thickness in Poly-Si/TiN Gate Electrodes for Scaled High-k CMOSFETs," *IEEE Electron Device Lett.*, vol. 30, no. 5, pp. 466–468, 2009.
[166] H. K. and M. X. and Y. H. and W. Wenwu, "Modulation of the effective work function of TiN metal gate for PMOS application," *J. Semicond.*, vol. 34, no. 8, p. 86002, 2013.
[167] C. C. Fulton, G. Lucovsky, and R. J. Nemanich, "Electronic properties of the Zr–ZrO2–SiO2–Si(100) gate stack structure," *J. Appl. Phys.*, vol. 99, no. 6, p. 63708, Mar. 2006.
[168] Y. C. and K. K. and T. N. and K. N. and A. Toriumi, "Quantitative Characterization of Band-Edge Energy Positions in High- k Dielectrics by X-ray Photoelectron Spectroscopy," *Jpn. J. Appl. Phys.*, vol. 52, no. 2R, p. 21101, 2013.
[169] Y. Yamashita, H. Yoshikawa, T. Chikyow, and K. Kobayashi, "Bias-voltage application in a hard x-ray photoelectron spectroscopic study of the interface states at oxide/Si(100) interfaces," *J. Appl. Phys.*, vol. 113, no. 16, p. 163707, Apr. 2013.
[170] Y. Yamashita, H. Yoshikawa, T. Chikyo, and K. Kobayashi, "Direct observation of bias-dependence potential distribution in metal/HfO2 gate stack structures by hard x-ray photoelectron spectroscopy under device operation," *J. Appl. Phys.*, vol. 115, no. 4, p. 43721, Jan. 2014.
[171] K. J. Yang and C. Hu, "MOS capacitance measurements for high-leakage thin dielectrics," *IEEE Trans. Electron Devices*, vol. 46, no. 7, pp. 1500–1501, 1999.
[172] K. Ahi and M. Anwar, "Developing terahertz imaging equation and enhancement of the resolution of terahertz images using deconvolution," 2016, vol. 9856, no., p. 98560N–9856–18.

List of publications

[PK-1] Pushpendra Kumar, Charles Leroux, Blend Mohamad, Alain Toffoli, Giovanni Romano, Xavier Garros, Gilles Reimbold, Florian Domengie, Carlos Suarez Segovia, G. Ghibaudo, “Effect of La and Al addition used for VT shift on the BTI reliability of HfON-based FDSOI MOSFETs,” IEEE IRPS, pp. 2B-2.1-2B-2.7, 2017.

[PK-2] Pushpendra Kumar, Florian Domengie, Charles Leroux, Patrice Gergaud, Gerard Ghibaudo, “Impact of CMOS TiN metal gate process on microstructure and its correlation with electrical properties,” accepted at the MRS Fall Meeting, 2018.

[PK-3] Pushpendra Kumar, Charles Leroux, Florian Domengie, Eugenie Martinez, Virginie Loup, Denis Guiheux, Yves Morand, Jean-Michel Pedini, Claude Tabone, Frederic Gaillard , Gerard Ghibaudo, “Development of X-ray Photoelectron Spectroscopy under bias and its application to determine band energies and dipoles in the HKMG stack,” accepted at IEEE IEDM, 2018.

[PK-4] Gilles Reimbold, Charles Leroux, Carlos Suarez Segovi, Blend Mohamad, Pushpendra Kumar, Xavier Garros, Florian Domengie, Philippe Blaise and Gerard Ghibaudo, "Dipoles in Gate-stack/FDSOI structure" ECS Conference, MA2017-02(14): 844, 2017.

List of Patents

[PK-5] Pushpendra Kumar, Charles Leroux, Eugenie Martinez, “Procédé de caractérisation d’un empilement de couches au moyen de mesures électriques et de spectrométrie de photoélectrons induits par rayons X sous tension, empilement de couches pour la mise en œuvre du procédé, procédé de fabrication de l’empilement de couches selon l’invention,” 31 July 2018

[PK-6] Pushpendra Kumar, Florian Domengie, “TiN metal gate with unique orientation to reduce threshold voltage,” 2017

Annex

In section 6.4.1, we presented the fitting procedures for Ti-2p, Hf-4f and Si-2p XPS signals of the HKMG stack. In fitting these signals, various area and binding energy constraints between the different components and their doublets were applied. First, the reference for the range of binding energies, used for different components in the fitting software, was taken from previous works on similar gate stacks [gaumer][rachid]. Secondly some fixed constraints were applied, that are reported in Table A1.

Constraints used	Ti-2p	Hf-4f	Si-2p
Doublet Area ratio	0.32	0.75	0.5
Doublet BE difference	TiN-2p = 5.9 eV TiON-2p = 5.4 eV TiO_2 = 5.3 eV	1.71 eV	0.6 eV
Components BE difference	No constraint	HfOSi - HfN = 2.07 eV HfO_2 - HfN = 1.09 eV	No constraint

Table A1: Binding energy and area ratio constraints used for the HKMG XPS signals between different components and their doublets

Résumé: Cette thèse concerne l'étude des procédés de fabrication des grilles HKMG des technologies FDSOI 14 et 28 nm sur les performances électriques des transistors MOS. Elle a porté spécifiquement sur l'aspect fiabilité et la maîtrise du travail de sortie effectif (WF_{eff}), au travers de la diffusion des additifs comme le lanthane (La) et l'aluminium (Al). Ce travail combine des techniques de caractérisation électriques et physico-chimiques et leur développement. L'effet de l'incorporation de ces additifs sur la fiabilité et la durée de vie du dispositif a été étudié. Le lanthane dégrade les performances de claquage TDDB et de dérives suite aux tests aux tensions négatives. L'introduction d'aluminium améliore le claquage TDDB, mais dégrade les dérives aux tensions positives. Ces comportements ont été reliés à des mécanismes physiques. Par ailleurs, la diffusion de ces additifs dans l'empilement de grille a été étudiée pour différents matériaux high-k en fonction de la température et de la durée de recuit de diffusion. Les doses d'additifs ont pu être ainsi mesurées, comparées et corrélées au décalage de travail de sortie effectif de grille. On a également étudié, les influences des paramètres du procédé de dépôt de grille TiN sur leur microstructure et les propriétés électriques du dispositif, identifiant certaines conditions à même de réduire la taille de grain ou la dispersion d'orientation cristalline. Toutefois, les modulations obtenues sur le travail de sortie effectif de grille dépendent plus du ratio Ti/N, suggérant un changement du dipôle à l'interface SiO_2 / high-k. Enfin, une technique éprouvée de mesure de spectroscopie à rayon X sous tension a pu être mise en place grâce des dispositifs spécifiques et une méthodologie adaptée. Elle permet de mesurer les positions relatives des bandes d'énergie à l'intérieur de l'empilement de grille. Cette technique a démontré que le décalage du travail de sortie effectif induits par des additifs (La or Al) ou par des variations d'épaisseur de grille métallique TiN provient de modifications du dipôle à l'interface SiO_2/ high-k.

Mots-clés: grille métallique, fiabilité, XPS sous tension, travail de sortie effectif, RF-PVD, microstructure, lanthane, aluminium, TiN, diffusion, dipôle, bandes d'énergie, NBTI, PBTI, TDDB, CV, CMOS, FDSOI, high-k

---

Abstract: This Ph.D. thesis is focused on the impact of the 14 and 28 nm FDSOI technologies HKMG stack processes on the electrical performance of MOS transistors. It concerns specifically the reliability aspect and the engineering of effective workfunction (WF_{eff}), through diffusion of lanthanum (La) and aluminum (Al) additives. This work combines electrical and physicochemical characterization techniques, and their development. The impact of La and Al incorporation, in the MOS gate stack, on reliability and device lifetime has been studied. La addition has a significant negative impact on device lifetime related to both NBTI and TDDB degradations. Addition of Al has a significant negative impact on lifetime related to PBTI, but on the contrary improves the lifetime for TDDB degradation. These impacts on device lifetime have been well correlated to the material changes inside the gate oxides. Moreover, diffusion of these additives into the HKMG stack with annealing temperature and time has been studied on different high-k materials. The diffused dose has been compared with the resulting shift in effective workfunction (WF_{eff}), evidencing clear correlation. In addition, impact of TiN metal gate RF-PVD parameters on its crystal size and orientation, and device electrical properties has been studied. XRD technique has been used to obtain the crystal size and orientation information. These properties are significantly modulated by TiN process, with a low grain size and a unique crystal orientation obtained in some conditions. However, the WF_{eff} modulations are rather correlated to the Ti/N ratio change, suggesting a change in the dipole at SiO_2/high-k interface. Lastly, using specific test structures and a new test methodology, a robust and accurate XPS under bias technique has been developed to determine the relative band energy positions inside the HKMG stack of MOS devices. Using this technique, we demonstrated that WF_{eff} shift induced by La and Al or by variations in gate thickness originates due to modifications of the dipole at SiO_2/high-k interface.

Keywords: metal gate, reliability, XPS under bias, effective workfunction, microstructure, RF-PVD, lanthanum, aluminum, TiN, diffusion, dipole, band energy, NBTI, PBTI, TDDB, CV, CMOS, FDSOI, high-k