



HAL
open science

Etude de la fiabilité porteurs chauds et des performances des technologies CMOS 0.13 μm - 2nm

Thierry Di Gilio

► To cite this version:

Thierry Di Gilio. Etude de la fiabilité porteurs chauds et des performances des technologies CMOS 0.13 μm - 2nm. Micro et nanotechnologies/Microélectronique. Université de Provence - Aix-Marseille I, 2006. Français. NNT: . tel-00117263v2

HAL Id: tel-00117263

<https://theses.hal.science/tel-00117263v2>

Submitted on 29 Jan 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

présentée à

L'université de Provence



pour obtenir

LE GRADE DE DOCTEUR

Spécialité : Systèmes Complexes.

Ecole doctorale : Physique, Modélisation et Sciences pour l'Ingénieur.

par

Thierry DI GILIO

"Etude de la fiabilité porteurs chauds et des performances des technologies CMOS 0.13 μm - 2nm"

Soutenue publiquement le 20 Octobre 2006 devant la commission d'examen:

Rapporteurs : Carole Plossu
Henri Happy

Examineurs : Rachid Bouchakour
Alain Bravaix (Encadrant)
Michel Lannoo (Directeur de Thèse)

Cette thèse a été préparée au **L**aboratoire de **M**atériaux ET **M**icro-electronique de **P**rovence (L2MP-UMR CNRS 6137) au sein de l'**I**nstitut **S**upérieur de l'**E**lectronique et du **N**umérique (ISEN) à Toulon.

"La théorie, c'est quand on sait tout et que rien ne fonctionne.
La pratique, c'est quand tout fonctionne et que personne ne sait pourquoi.
Si la pratique et la théorie sont réunies, rien ne fonctionne et on ne sait pas pourquoi."

Albert Einstein.

Remerciements

Je remercie tout particulièrement **Alain BRAVAIX** pour la rigueur et la qualité de son encadrement. Les nombreuses discussions que nous avons pu avoir tout au long de ses travaux m'ont toujours permis d'éclaircir ma vision des choses. Je remercie également **Didier GOUENHEIM** pour sa disponibilité, ses connaissances sur la physique du semi-conducteur et son enthousiasme à les transmettre.

Je tiens également à remercier **Michel LANNOO** de m'avoir accueilli dans le Laboratoire de Matériaux et de Micro-électronique de Provence (L2MP), et ce, dès mon stage de Maîtrise, et par la suite pour mon stage de DEA puis ma Thèse. Je dois également remercier L'Institut Supérieur d'Electronique et du Numérique (ISEN) qui a mis à ma disposition ses locaux pour le déroulement de mes travaux de recherche.

Sur un plan plus personnel je voudrais remercier **ma Famille** pour son soutien. Il me faut également remercier ici l'ensemble des doctorants de l'ISEN et amis qui ont largement contribué à créer une ambiance chaleureuse et conviviale au sein de l'équipe : **Fabien Chaillan, Anne Collard-Bovy, Simon Desbief, Mmike Fournigault, Fabrice Guigues, François Rudolf, Vincent Telandro, Alexandre Tramoni**, ainsi que les enseignants-chercheurs : **Philippe Courmontagne, Edith Kussener, Lionel Patrone**, et mes amis **Marc Bescond, Solenn Mezerette, Pascal Sati, Arnaud Pauc, Christine Castelli . . .**

Merci aussi l'ensemble du personnel administratif de l'ISEN et de la Maison des Technologies, tout particulièrement **Véronique Casiez, Régine Matheron, Béatrice Negre, Valérie Gallene, Jean-Louis Campelloni**, mais aussi le **BDE de l'ISEN**, l'association **ISEN Espérance**.

Enfin je remercie la société **STMicroelectronics** (Central R&D Crolles, France) pour son soutien et sa collaboration, ainsi que la **Région PACA** et la **Délégation Provence du CNRS** pour le financement de cette thèse.

Table des matières

Table des matières	i
Introduction	1
I Principe de fonctionnement et modélisation des transistors MOSFET	9
Introduction	9
I.1 La capacité MOS	10
I.1.1 Régimes de fonctionnement	10
I.1.2 Charge totale dans le semi-conducteur Q_{SC}	12
I.1.3 Capacité totale C	14
a) Calcul de la capacité	14
b) Mesure de la capacité différentielle	17
I.1.4 Défauts dans la structure MOS	18
a) Défauts d'interface	18
b) Charges mobiles	19
c) Charges fixes	19
d) Charges piégées dans l'oxyde	19
e) Conséquences	20
I.2 Le transistor MOSFET	22
I.2.1 Modélisation du courant de Drain	23
a) Régime linéaire	23
i. Calcul de Q_{SC}	24
ii. Calcul de Q_D	24
iii. Calcul de I_{DS}	25
b) Régime de saturation	27
I.3 Effets liés à la réduction des dimensions des transistors MOSFETs	30
I.3.1 Effet de la géométrie du canal	30
a) Effet canal court	30
b) Effet canal court inverse	32
c) Effet canal étroit	33
I.3.2 Effet DIBL	34

I.3.3	Réduction de la mobilité des porteurs	35
a)	Collisions avec les phonons	35
b)	Interactions coulombiennes	36
c)	Collisions sur la rugosité de surface	36
d)	Expression de la mobilité effective	36
e)	Saturation de la vitesse des porteurs avec le champ électrique latéral	37
f)	Lien avec la tension de saturation $V_{Dsat}(\xi_{sat})$	38
I.3.4	Modulation de la longueur du canal	39
I.3.5	Application à la modélisation du courant de Drain	41
I.4	Extraction des principaux paramètres	44
I.4.1	Extraction de la tension de seuil	44
I.4.2	Extraction du facteur de forme et du dopage	46
I.4.3	Extraction du facteur de réduction de mobilité	47
I.4.4	Détermination des dimensions effectives du canal	48
a)	Extraction de ΔL_G	49
b)	Extraction de ΔW_G	49
I.4.5	Détermination de la valeur du pic de champ latéral	50
I.4.6	Calcul de la résistance série	51
	Conclusion	55
	Références du Chapitre I	57
II	Techniques d'analyses de la dégradation du transistor MOSFET	61
	Introduction	61
II.1	Analyse par mesures courant-tension I-V	64
II.1.1	Suivi de la variation du courant de Drain au cours du temps	64
a)	Régime linéaire	64
b)	Régime saturé	66
II.1.2	Variation de la tension de seuil	67
II.1.3	Variation de la transconductance	69
II.1.4	Variation de la pente sous seuil	70
II.2	Analyse des défauts par pompage de charges	73
II.2.1	Pompage de charges standard à deux niveaux	74
a)	Principe	74
b)	Théorie de la statistique de génération-recombinaison	74
c)	Expression du courant pompé I_{CP}	76
II.2.2	Mesures expérimentales et interprétations	82
a)	Mesures à fréquence variable	82
b)	Mesures à amplitude constante	85

c)	Mesures à V_{Gbas} ou V_{Ghaut} fixe	86
II.2.3	Evolution au cours d'un stress	88
a)	Analyse de défauts CP standard	88
b)	Technique CP Différentielle	90
II.2.4	Contraintes et limitations	91
II.3	Autres méthodes d'analyse	95
II.3.1	Pompage de charge à trois niveaux	95
II.3.2	Profilage de la zone de défauts	96
II.3.3	Grille flottante	98
	Conclusion	102
Références du Chapitre II		103
III Mécanismes de vieillissement des transistors MOSFETs soumis aux injections de porteurs chauds		107
	Introduction	107
III.1	Le régime porteurs chauds : généralités	108
III.1.1	Injection d'électrons chauds	108
III.1.2	Injection de trous chauds	112
III.1.3	Influence de l'épaisseur de l'isolant de Grille	114
III.2	Expérimentations : stress statiques.	116
III.2.1	Description des échantillons	117
III.2.2	Oxydes épais : $12nm$	120
a)	Transistors NMOS, cas standard : maximum du courant substrat	120
b)	Transistors PMOS, cas standard : maximum du courant de Grille électronique	126
III.2.3	Oxydes moyens : $6.5nm$	131
a)	Cas des transistors NMOS	132
b)	Cas des transistors PMOS	137
III.2.4	Oxydes minces : $2.1nm$	139
a)	Etude des transistors NMOS	142
i.	Stress HE.	142
ii.	Stress IB.	144
iii.	Stress VG+.	145
b)	Etude des transistors PMOS.	148
i.	Stress HH.	150
ii.	Stress IB.	152
iii.	Stress VG-.	154
c)	Spécificité du cas HH dans le PMOS	155
d)	Comparaison des dispositifs PMOS <i>Low Leakage</i> et <i>High Speed</i>	160

III.3	Stress alternés	166
III.3.1	Influence des pièges neutres dans les oxydes épais	166
III.3.2	Oxydes minces	167
III.4	Synthèse	169
	Conclusion	171
Références du Chapitre III		173
IV	Techniques d'estimation de la durée de vie des dispositifs MOSFET	177
	Introduction	177
IV.1	Expression de la durée de vie	178
IV.1.1	Première modélisation du vieillissement des dispositifs MOSFET	179
IV.1.2	Détermination pratique de la durée de vie (Transistor NMOS)	182
IV.1.3	Expression de la durée à l'aide du " <i>Lucky Electron Model</i> "	184
a)	Durée de vie en fonction de I_{BS} (transistor NMOS)	184
b)	Durée de vie en fonction de I_{GS} (transistors NMOS)	187
IV.1.4	Equivalence pour les transistors PMOS	189
a)	Durée de vie en fonction de I_{BS}	189
b)	Durée de vie en fonction de I_{GS}	191
IV.1.5	Synthèse sur les modèles LEM et LHM	194
IV.1.6	Expression de la durée de vie dans le cas des injections uniformes	196
a)	Durée de vie en fonction de la tension de Grille	197
b)	Durée de vie en fonction du courant de Grille	198
IV.2	Durée de vie des transistors à oxydes ultra minces soumis aux injections de porteurs chauds	200
IV.2.1	Les oxydes minces et le concept des " <i>porteurs chanceux</i> "	200
IV.2.2	Modèle de séparation des porteurs chauds et froids	200
IV.3	Influence du choix du moniteur	205
	Conclusion	212
Références du Chapitre IV		213
	Conclusion générale	215
	Annexes	217
A	Régime d'inversion faible avant saturation	217
A.1	Calcul des charges Q_{SC} , Q_D et Q_n	217
A.2	Expression du courant I_{DS}	219
	Annexes	223

B Courant de fuite à travers l'oxyde de grille mince	223
B.1 Courant tunnel Fowler-Nordheim	224
B.2 Courant tunnel direct	226

Symbole	Unité	Définition
n_0	cm^{-3}	Concentration d'électrons dans le Silicium loin de l'interface.
p_0	cm^{-3}	Concentration de trous dans le Silicium loin de l'interface.
$n(y)$	cm^{-3}	Concentration d'électrons dans le Silicium à la distance y de l'interface.
$p(y)$	cm^{-3}	Concentration de trous dans le Silicium à la distance y de l'interface.
n_i	cm^{-3}	Concentration intrinsèque d'électrons dans le Silicium.
q	C	Valeur relative de la charge électronique (1.6×10^{-19}).
ϕ_F	V	Potentiel du substrat dû au dopage du Silicium.
ϕ_{ms}	V	différence des travaux de sortie ente métal et semi-conducteur ($\phi_m - \phi_s$).
ψ_S	V	Potentiel à l'interface Si-SiO ₂ .
$\psi(y)$	V	Potentiel dans le silicium à la distance y de l'interface.
k	$J.K^{-1}$ $eV.K^{-1}$	Constante de Boltzmann ($1,3806503 \times 10^{23}$) ($8,61738 \times 10^{-5}$).
β	eV	q/kT .
E_c	eV	Niveau d'énergie de la bande de conduction loin de l'interface.
E_v	eV	Niveau d'énergie de la bande de valence loin de l'interface.
E_F	eV	Niveau d'énergie pour le niveau de Fermi.
E_i	eV	Niveau d'énergie intrinsèque, milieu de E_c et E_v loin de l'interface.
E_{cs}	eV	Niveau d'énergie de la bande de conduction à l'interface.
E_{vs}	eV	Niveau d'énergie de la bande de valence à l'interface.
E_{is}	eV	Niveau d'énergie intrinsèque, milieu de E_{cs} et E_{vs} à l'interface.
V_{DS}	V	Tension appliquée au Drain avec la Source référence des potentiels.
V_{FB}	V	Tension de bande plate.
V_{DSat}	V	Valeur de saturation de la tension de Drain avec la Source référence des potentiels.
V_{GS}	V	Tension appliquée à la Grille avec la Source référence des potentiels.
V_{BS}	V	Tension appliquée au Substrat avec la Source référence des potentiels.
V_{OX}	V	Différence de potentiel au bornes de l'isolant de Grille.
V_{DD}	V	Tension d'alimentation nominale pour la technologie considérée.
I_{DS}	A	Courant mesuré au Drain avec la Source référence des potentiels.
I_{DSat}	A	Courant de Drain saturé avec la Source référence des potentiels.
I_{GS}	A	Courant mesuré à la Grille avec la Source référence des potentiels.
I_{BS}	A	Courant mesuré au Substrat avec la Source référence des potentiels.
V_{Th}	V	Tension de seuil des transistors MOSFETs.
W_{eff}	cm	Largeur effective de la Grille (inférieure à la largeur dessinée).
L_{eff}	cm	Longueur effective de la Grille (inférieure à la longueur dessinée).
T_{ox}	cm	Epaisseur de l'oxyde de Grille.

TAB. 1 – Définitions des symboles et notations employés dans ce manuscrit.

Symbole	Unité	Définition
$\phi_{b,h}$	eV	Hauteur de barrière tunnel pour les trous.
$\phi_{b,e}$	eV	Hauteur de barrière tunnel pour les électrons.
$\phi_{i,h}$	eV	Coefficient d'ionisation pour les trous.
$\phi_{i,e}$	eV	Coefficient d'ionisation pour les électrons.
$\phi_{it,h}$	eV	Énergie nécessaire à un trou pour générer un état d'interface.
$\phi_{it,e}$	eV	Énergie nécessaire à un électron pour générer un état d'interface.
$\lambda_{e,h}$	nm	Libre parcours moyen des électrons/trous.
τ	s	Durée de vie des transistor MOSFETs.
X_J	cm	Profondeur des jonctions de Drain et de Source.
$N_{D/A}$	cm^{-3}	Densité de dopants (Donneur/Accepteur).
C_{OX}	$F.cm^{-2}$	Capacité de l'oxyde de Grille.
C_{sc}	$F.cm^{-2}$	Capacité du semi-conducteur (Silicium).
Q_{OX}	$C.m^{-2}$	Charge dans l'oxyde de Grille.
Q_{sc}	$C.m^{-2}$	Charge dans le semi-conducteur (Silicium).
Q_{it}	$C.m^{-2}$	Charge piégée sur les états d'interface.
Q_{ot}	$C.m^{-2}$	Charge piégée dans l'oxyde.
Q_m	$C.m^{-2}$	Charge mobile dans l'oxyde.
N_{it}	cm^{-2}	Nombre de charges piégées à l'interface.
N_m	cm^{-2}	Nombre de charges mobiles dans l'oxyde.
D_{it}	$eV^{-1}.cm^{-2}$	Nombre de charges piégées à l'interface par niveau d'énergie.
ξ, ξ_m	$V.cm^{-1}$	Champ électrique latéral dans le canal.
F_{ox}	$V.cm^{-1}$	Champ électrique à travers l'oxyde.
J	$A.cm^{-2}$	Densité de courant (J=Courant/Surface).
ϵ_0	$F.cm^{-1}$	Permittivité électrique du vide (8.84×10^{-14}).
ϵ_{ox}	s.u.	Permittivité relative de l'oxyde (SiO ₂).
ϵ_{Si}	s.u.	Permittivité relative du Silicium.
μ_0	$V.cm^{-1}.s^{-1}$	Mobilité des porteurs en champ nul.
μ_{eff}	$V.cm^{-1}.s^{-1}$	Mobilité effective des porteurs.
R_{SD}	Ω	Somme des Résistances séries à la Source et au Drain.
R_{Canal}	Ω	Résistance électrique du canal.
G_m	S ou Ω^{-1}	Transconductance (dI_{DS}/dV_{GS}).
$\sigma_{n/p}$	cm^{-2}	Section efficace de capture d'électron/trous.
$c_{n/p}$	$cm^{-3}.s^{-1}$	Coefficient de capture d'électron/trou d'un piège.
$e_{n/p}$	$cm^{-3}.s^{-1}$	Coefficient d'émission d'électron/trou d'un piège.
v_T	$cm.s^{-1}$	Vitesse thermique des porteurs ($1/2mv_T^2 = 3/2kT$)
$m_{n,p}^*$	kg	Masse effective des électrons/trou dans le milieu considéré.

TAB. 2 – Définitions des symboles et notations employés dans ce manuscrit (suite).

Introduction

La performance et la fiabilité sont deux notions antagonistes mais indissociables. C'est la performance, qui depuis le début de l'électronique (devenue micro puis nano-électronique), a été le moteur d'un formidable développement. En parallèle la fiabilité, parce que nécessaire à la production de masse, tempère ces évolutions. Ainsi, ces dernières années, on a vu croître la part de budget allouée aux études de fiabilité, mais nettement moins que les coûts de recherche et développement. On peut illustrer ceci par quelques faits historiques marquants. Prenons comme exemple l'informatique, dont l'évolution est intrinsèquement liée à celle de l'électronique. En 1938, Konrad Zuse invente un ordinateur qui fonctionne grâce à des relais électromécaniques : le Z3. Cet ordinateur est le premier à utiliser le binaire au lieu du décimal. En 1937, Howard Aiken met au point un ordinateur programmable mesurant 17 m de long et 2.5 mètres de hauteur, permettant de calculer 5 fois plus vite que l'homme : c'est le Mark I d'IBM. Il est constitué de 3300 engrenages, 1400 commutateurs reliés par 800 km de fil électrique. En 1943, le premier ordinateur ne comportant plus de pièces mécaniques est créé grâce à J. Mauchly et J. Presper Eckert: l'ENIAC (Electronic Numerical Integrator And Computer). Il est composé de 18000 lampes à vide, et occupe une place de 1500 m². Il fut utilisé pour des calculs ayant servi à mettre au point la bombe H. Sa consommation électrique avoisine quelques centaines de kilowatts, et sa fiabilité n'a d'égale que celle des lampes qui le constituent, soit quelques heures¹. Plusieurs améliorations notables, basées sur cette technologie, seront apportées par la suite. Mais c'est réellement avec l'invention du premier transistor en 1947 par John Bardeen, Walter Brattain et William Shockley dans les laboratoires de Bell Telephone, que l'électronique moderne trouve ses racines. Ce transistor bipolaire (Fig. 1), d'abord en Germanium, puis en Silicium (1954) pour des raisons de coûts et de procédés de fabrication, ouvre la voie à une longue série d'innovations. Le transistor à effet de champ en 1959: plus proche de la triode que ne l'est le transistor bipolaire, celui-ci est composé d'une électrode appelée Grille (G) qui module la conductance entre une zone dite Source (S) et une autre dite Drain (D). En 1960, le premier transistor dit planar: Jean Hoerni fabrique un transistor plat à l'aide de gaz dopant positivement, négativement ou bien transformant le silicium en silice (oxyde de silicium SiO_2)

1. On peut citer ici comme anecdote la première erreur de calcul, due à un court-circuit provoqué par un insecte, *bug* en anglais, qui est resté le terme consacré.

qui est un isolant.

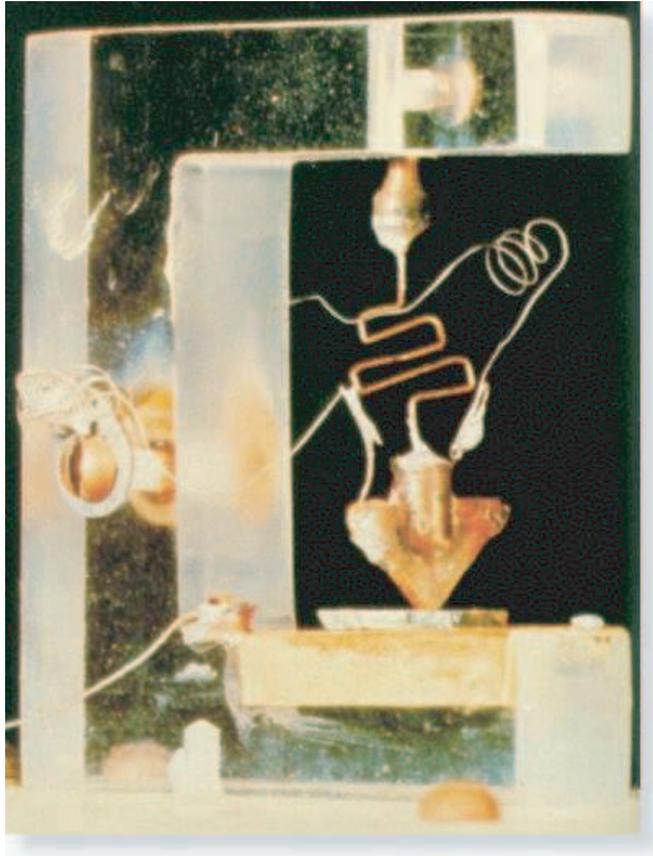


FIG. 1 – *Premier transistor bipolaire en germanium, 1947.*

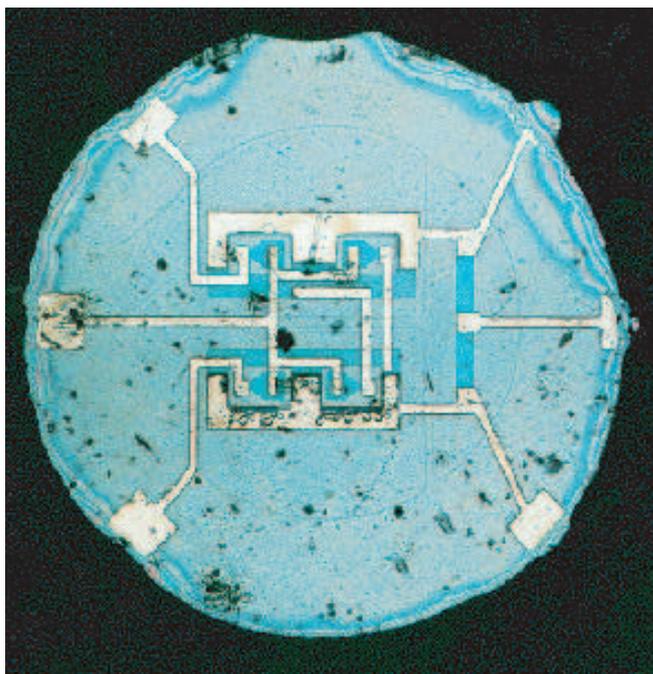


FIG. 2 – *Premier circuit intégré planar, 1961.*

Cette idée sera reprise quelques mois plus tard par Robert NOYCE qui intégrant la technologie *planar* mettra au point des procédés toujours utilisés aujourd'hui pour fabriquer les circuits intégrés. Les débuts de la micro-électronique moderne coïncident avec la mise au point de la technologie Metal-Oxide-Semi-conducteur (MOS), en 1970. Cette technologie permet de fabriquer des transistors plus petits et plus rapides. Une course à la densité, à la vitesse et à la consommation commence... On parle alors de SSI (Small Scale Integration) : 30 à 80 transistors, MSI (Medium Scale Integration) , LSI (Large Scale Integration) , VLSI (Very Large Scale Integration). C'est en 1971, que la compagnie Intel produit son premier microprocesseur, le *i4004*, dont la puissance est équivalente à celle de l'ENIAC. Un des co-fondateurs de de cette entreprise, Gordon Moore, prévoit dès 1965 que le nombre de transistors par circuit intégré doublera tous les deux ans². La loi est restée vérifiée jusqu'à la fin des années 1970, pour tendre vers un doublement tous les 18 mois actuellement (Fig. 3).

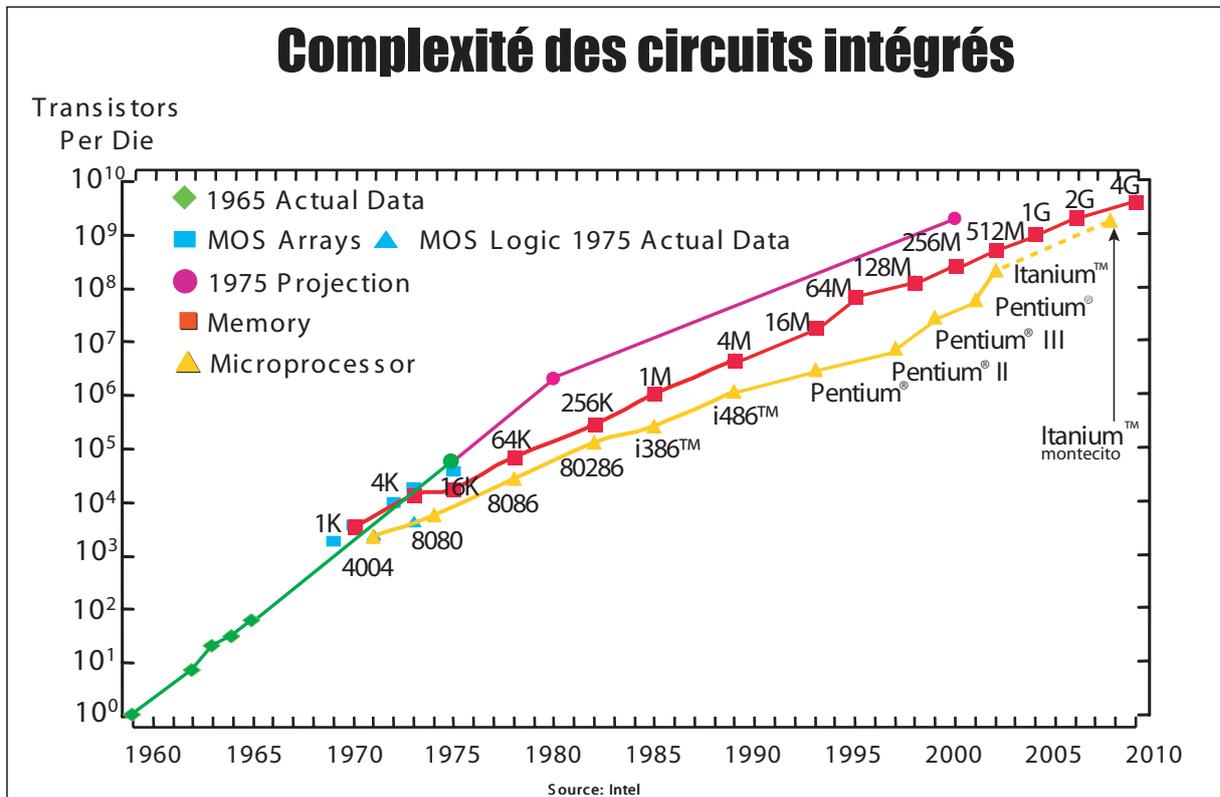


FIG. 3 – Illustration de la loi de Moore avec l'évolution du nombre de transistors dans les générations successives de processeurs Intel (<http://www.intel.com>).

Cette croissance de la densité passe par un processus de miniaturisation des dispositifs, qui tente de maintenir les caractéristiques électriques constantes. Plusieurs façons de procéder ont

2. G. E. Moore, "Cramming more components onto integrated circuits", *Electronics*, Vol 38, No 8, April 19, 1965.

vu le jour, le principal problème de la miniaturisation étant la dépendance directe et surtout indirecte des caractéristiques électriques vis à vis des paramètres physiques contrôlables. Ceci entraîne de nombreux effets parasites qui modifient les performances et la consommation d'énergie des systèmes.

Paramètre	Réduction à champ électrique constant	Réduction à tension constante	Réduction à champ électrostatique constant
Longueur de grille	1/K	1/K	1/K
Largeur de Grille	1/K	1/K	1/K
Epaisseur d'oxyde	1/K	1/K	1/K
Profondeur de jonction	1/K	1/K	1/K
Dopage	K	K^2	K^2/λ
Tension de Drain	1/K	1	1/ λ
Courant de Drain	1/K	K	K/λ^2
Tension de seuil	1/K	1	1/ λ
Temps de propagation	1/K	1/K	1/K
Tension d'alimentation	1/K	1	1/ λ
Capacité de Grille	1/K	1/K	1/K
Nombre de transistors	K^2	K^2	K^2
Taille de puce	1	1	1
Puissance consommée	1	K^3	K^3/λ^3

TAB. 3 – *Lois de dessin utilisées pour réduire les dimensions des technologies déjà existantes.*

Le premier schéma complet de lois d'intégration a été présenté par Dennard en 1974³. La méthode s'appelle la réduction à champ électrique constant (voir le Tab. 3). Afin de réduire la zone de désertion de porteurs, les champs internes, les courants, et les capacités, ainsi que toutes les dimensions sont mesurés par un facteur K. Selon la variable, le paramètre peut être multiplié, ou divisé par K. De cette manière, les effets parasites ont été évités dans une certaine mesure. L'inconvénient principal de cette méthode est qu'il n'est souvent pas possible de modifier les paramètres dans les proportions requises. Par l'exemple, le dopage de substrat a une limite supérieure de 10^{18} cm^{-3} . Ainsi, si la limite est déjà atteinte, il est impossible d'augmenter les doses de dopants. La réduction de la tension de seuil pose également quelques problèmes particulièrement délicats. Suivant le type d'application pour lesquelles ils sont employés, les tensions de seuil des dispositifs varient. En effet, les circuits logiques rapides utilisent des transistors rapides dits High Speed (HS) ayant de faibles tensions de seuil, alors que les circuits analogiques

3. R. H. Dennard, F. H. Gaensslen, H-N. Yu, V. L. Rideout, E. Bassous, A. R. LeBlanc, "Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions", *IEEE Journal of Solid-State Circuits*, Vol. SC-9, pp.256-268, 1974.

nécessitent des transistors avec peu de fuite de courant dit Low Leakage (LL). Clairement, la réduction à champ constant est le fil conducteur des technologues, même s'il reste difficile à suivre.

Différentes façons de procéder vont suivre, comme la réduction à tension constante (Tab. 3). La méthode à tension constante tente de suivre la réduction de la tension d'alimentation imposée par l'industrie. Dans la réduction à champ électrique constant, les tensions d'alimentation sont réduites d'un facteur K . Par l'intermédiaire de "roadmaps", l'industrie a convenu, des années à l'avance, de ce que seront les tensions d'alimentation, laissant ainsi un délai suffisant pour concevoir les circuits d'alimentation (régulateur/convertisseur de tension et de courant). La réduction à tension constante est donc une méthode plus adaptée que l'évolution à champ électrique constant⁴. Le principal inconvénient de cette méthode est l'augmentation des champs électriques créés dans le dispositif, due à la conservation des tensions d'alimentation. De plus, ceci amène une dégradation de la mobilité des porteurs, à la génération de porteurs énergétiques ou "porteurs chauds", et à d'autres problèmes de fiabilité (perçage, claquage d'oxyde). Enfin, cette méthode engendre plus de puissance consommée et donc des techniques de refroidissement plus lourdes que pour l'évolution à champ électrique constant.

Une troisième méthode a été proposée: la diminution à champ électrostatique constant, ou à tension quasi-constante (voir le Tab. 1). Dans cette méthode, les dimensions sont mesurées par le même facteur K , mais les potentiels sont mesurés par un facteur différent⁵ λ tel que $\lambda = K^{0.5}$. Cette méthode est un autre compromis entre la réalité et la méthode à champ électrique constant. Le facteur λ est utilisé lorsque les tensions ne peuvent pas être réduites par K . Ceci laisse les lignes de champ inchangées et réduit les effets parasites. Bien que cette méthode corrige la plupart des défauts pratiques des deux précédentes, elle reste une méthode théorique et sert seulement de bon point de départ pour les concepteurs de dispositifs. Davantage de tests et d'optimisations seront toujours nécessaires pour mettre au point les "recettes" de fabrication pour une technologie particulière.

La caractérisation électrique d'un dispositif est d'une grande importance et exige une attention particulière dans le développement des technologies MOSFET avancées. Le choix de la longueur du canal, l'épaisseur d'oxyde de grille, le dopage de substrat et la technologie de Source/Drain déterminent en grande partie les performances du dispositif. En concevant des dispositifs plus petits, on doit également considérer l'impact de la réduction des dimensions sur la fiabilité des circuits intégrés. L'étude de la fiabilité d'une technologie est basée sur le suivi des

4. S. Wolf, "Silicon Processing for the VLSI Era, Volume III, The Submicron MOSFET", Lattice Press, California, 1995.

5. G. Baccarini, M. R. Wordeman, R. H. Dennard, "Generalized Scaling Theory and Its Application to a Micrometer MOSFET Design", *IEEE Transactions on Electron Devices*, Vol. ED-31, pp. 452-462, 1984.

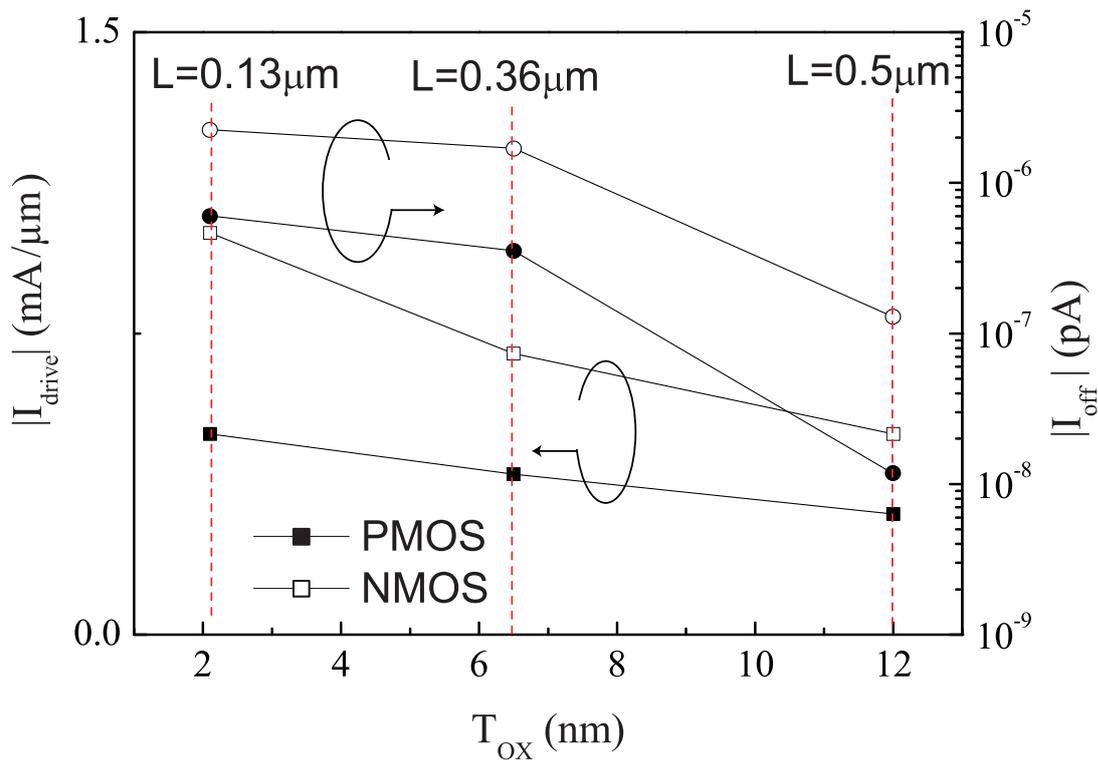


FIG. 4 – Evolution du courant de Drain en mode passant I_{on} et bloqué I_{off} avec la réduction des dimensions

performances des circuits dans le temps, et il est de la responsabilité des spécialistes en fiabilité de s'assurer (par des modélisations du comportement et des vieillissements accélérés) que la vie d'un dispositif réduit est acceptable. Il y a beaucoup de mécanismes potentiels de détérioration dans les technologies actuelles MOSFET sur silicium, telles que la dégradation par porteurs chauds, le claquage d'oxyde de Grille, les défauts d'interconnexion dus à l'électromigration, ou encore les problèmes liés à la température. Ces mécanismes peuvent conduire à des modes de défaillance qui représentent une limitation sévère à l'intégration des filières CMOS, en raison des champs électriques et des densités de courant trop élevés. Cette course à la miniaturisation est finalement affaire de compromis : l'amélioration des performances se fait souvent au détriment d'autres paramètres. On peut par exemple comparer le Tab. 4 et la Fig. 4 qui montrent d'une part la réduction des temps de commutation des portes logiques (gain de performance) et d'autre part l'évolution des courants de Drain I_{DRIVE} ($V_{GS} = V_{DS} = V_{DD}$), I_{off} ($V_{GS} = 0V, V_{DS} = V_{DD}$) et I_{on} ($V_{GS} = V_{DD}, V_{DS} = 0V$) en fonction de T_{ox} . On peut constater une augmentation significative du courant de fuite I_{off} , qui multiplié par les quelques millions de transistors qui composent désormais les circuits logiques, donnent une consommation au repos assez conséquente.

$T_{ox}(nm)$	$W(\mu m)$	$L(\mu m)$	$V_{DD}(V)$	$V_{TN}(V)$	$C_{OX}(fF)$	$\tau_{Ch}(ns)$
12nm	10	0.5	5V	1	14.38	39.5
6.5nm	10	0.36	3.3V	0.7	20.7	4.19
2.1nm	10	0.13	1.2V	0.3	21.36	1.6

TAB. 4 – Tableau de l'évolution avec la technologie du temps de descente d'une porte CMOS, calculé pour une capacité de charge $\overline{C}_L = 2/3C_{OX}$, τ_{Ch} est donné par $\tau_{Ch} \equiv \overline{C}_L \frac{V_{DD}}{I_{DRIVE}}$ où I_{DRIVE} est le courant mesuré pour $V_G = V_D = V_{DD}$.

Le but de cette étude est l'analyse de la fiabilité et des performances de la filière CMOS $0.13\mu m$. Dans cette optique, nous allons commencer par établir les équations de bases des structures MOS, pour ensuite en affiner la modélisation en introduisant les effets de la miniaturisation, et enfin décrire les principales méthodes de caractérisations indispensables pour l'élaboration et la validation des modèles. Dans le second chapitre du manuscrit seront abordées les techniques d'analyses de la fiabilité des transistors MOSFETs, en fonction des caractéristiques courant-tension, ou encore à l'aide de la technique dite "pompage de charges" (CP). Ceci nous amènera à l'étude des mécanismes de dégradation rencontrés dans la filière $0.13\mu m$ présentant un oxyde de Grille de $2nm$. L'apparition de nouveaux modes de dégradation ainsi que leur impact sur les techniques de détermination de la durée de vie, qui feront l'objet du chapitre IV.

CHAPITRE I

Principe de fonctionnement et modélisation des transistors MOSFET

Introduction

La miniaturisation incessante des technologies CMOS répond à des impératifs de performances et de rentabilité, moteurs de l'industrie de la micro-électronique mondiale. Cette miniaturisation va de pair avec l'augmentation du nombre de transistors MOSFET par puce, phénomène bien connu, puisque Gordon Moore, co-fondateur de INTEL Corporation, avait prévu cette loi de croissance dès 1965, quatre ans seulement après la fabrication du premier circuit intégré planaire. On est proche aujourd'hui du milliard de transistors par puce, et la complexité des architectures sur silicium est telle, que la conception assistée par ordinateur, requiert une modélisation poussée du comportement électrique des dispositifs MOS. La réduction des géométries des transistors s'accompagne du développement de nouveaux procédés de fabrication qui ne sont pas sans générer des effets parasites qui influencent le fonctionnement des composants. Ceci a pour conséquence d'accroître la complexité des modèles prédictifs, qui sont ainsi ajustés et optimisés à chaque nouvelle génération de procédés de fabrication. On peut citer les onze générations de modèle SPICE (1-11) qui se succédèrent avant de laisser place aux modèles plus complexes BSIM (de Berkeley) et MM9 (développé par Philips). Ces modèles s'établissent à partir des propriétés électriques globales du composant, obtenues à l'aide de leurs caractéristiques I-V ajustées. Dans cette partie, nous allons décrire les équations de base d'une structure capacitive MOS, pour arriver à la modélisation du transistor. Les effets parasites les plus couramment rencontrés seront ensuite définis.

I.1 La capacité MOS

Le transistor MOSFET repose sur le principe de la capacité MOS à laquelle on adjoint les deux diodes qui constituent le Drain et la Source. Nous allons donc dans cette partie, décrire le comportement des structures MOS vis à vis des charges en présence. Nous allons nous attacher à faire une description classique et conventionnelle de cette structure, et mettrons en évidence la nécessité de recourir à une modélisation plus fine pour les capacités à oxydes de grille ultra-fins.

I.1.1 Régimes de fonctionnement

La capacité d'une structure MOS est équivalente à la mise en série de la capacité de l'oxyde avec la somme des capacités présentes dans le silicium. Dans les technologies actuelles, la grille n'est plus un métal mais du poly-silicium dont le fort dopage ($> 10^{20} \text{ cm}^{-3}$) lui confère les mêmes propriétés qu'un métal.

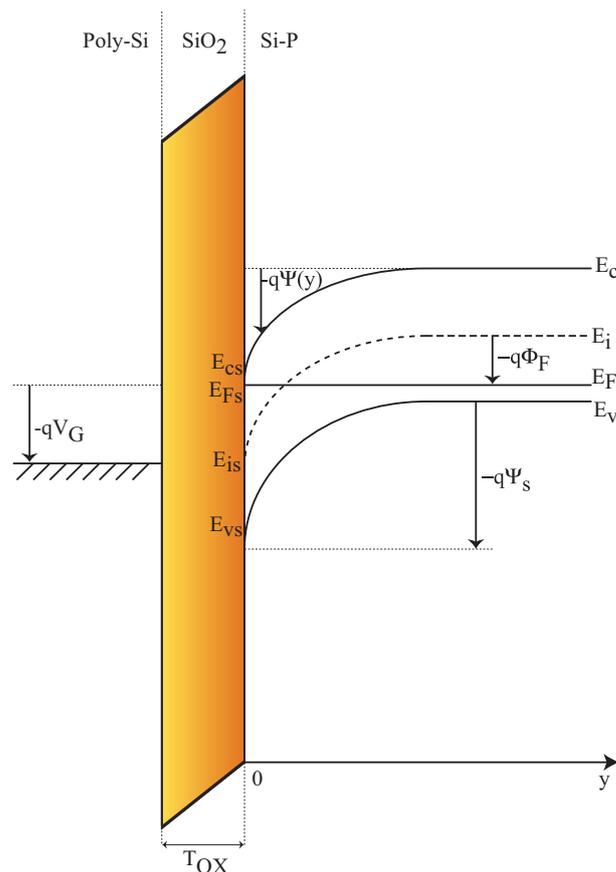


FIG. I.1 – Diagramme de bande d'une capacité MOS à substrat P

La Fig. I.1 représente le diagramme des bandes d'énergie d'une structure Métal Oxyde Semi-conducteur pour un transistor MOS à canal N en régime d'inversion. Le niveau de Fermi E_F est au dessous du minimum de la bande de conduction, permettant au matériaux de disposer d'une

grande quantité d'électrons. Le bas de la bande de conduction est représenté par E_c , et le haut de la bande de valence par E_v . E_i est le milieu de la bande interdite du silicium. A l'interface, tous ces niveaux d'énergie sont indicés "s" (E_{vs} , E_{cs} , E_{is}); la courbure des bandes, conséquence de la polarisation extérieure, est donnée par $\psi(y)$.

Le potentiel de volume du substrat, ϕ_F , indique la différence entre le milieu de la bande interdite et le niveau de Fermi, il a pour expression [1] :

$$\phi_F = \frac{kT}{q} \ln \left(\frac{N_A}{n_i} \right) = \frac{1}{\beta} \ln \left(\frac{N_A}{n_i} \right) = -\frac{1}{q} (E_F - E_i) \quad (\text{I.1})$$

où k est la constante de Boltzmann, T la température ambiante absolue (en kelvin), N_A le dopage du substrat (de type P) supposé uniforme dans le volume où se forme la couche d'inversion, et q la valeur absolue de la charge de l'électron.

En régime d'inversion faible, le potentiel de surface ψ_s varie entre $\phi_F - V_{BS}$ et $2\phi_F - V_{BS}$, où V_{BS} est la différence de potentiel entre la source et le Substrat. Les densités de porteurs majoritaires et minoritaires s'expriment à l'aide du niveau de Fermi :

$$n(y) = n_i \exp \left(\frac{E_F - E_i(y)}{kT} \right) \quad (\text{I.2})$$

$$p(y) = n_i \exp \left(\frac{E_i(y) - E_F}{kT} \right) \quad (\text{I.3})$$

En utilisant (I.1) et considérant que dans le volume à l'équilibre les concentrations sont données par :

$$n_0 = n_i \exp \left(\frac{-q\phi_F}{kT} \right) \quad (\text{I.4})$$

$$p_0 = n_i \exp \left(\frac{q\phi_F}{kT} \right) \quad (\text{I.5})$$

on obtient :

$$n(y) = n_0 \exp (\beta (\psi(y))) \quad (\text{I.6})$$

$$p(y) = p_0 \exp (-\beta \psi(y)) \quad (\text{I.7})$$

où $\beta = q/kT$.

I.1.2 Charge totale dans le semi-conducteur Q_{SC}

La densité de charge totale ρ du semi-conducteur dépend des densités de charges libres n/p et de charges fixes dues aux impuretés des dopants ionisés du substrat N_D/N_A . Ceci se traduit par :

$$\rho = q [p - n + N_D - N_A] \quad (\text{I.8})$$

Dans le volume du substrat, le semi-conducteur est à l'équilibre, et respecte la condition de neutralité :

$$\rho(y \rightarrow \infty) = p_0 - n_0 + N_D - N_A = 0 \quad (\text{I.9})$$

Il vient $p_0 - n_0 = N_A - N_D$. En utilisant les densités de porteurs libres (I.2) et (I.3), (I.9) se transforme en :

$$\rho = -q \{ n_0 (\exp^{\beta\psi(y)} - 1) - p_0 (\exp^{-\beta\psi(y)} - 1) \} \quad (\text{I.10})$$

Dans la zone désertée, le champ électrique ξ et la densité à l'équilibre ρ sont reliés par l'équation de Poisson :

$$\frac{d^2\psi}{dy^2} = -\frac{d\xi}{dy} = -\frac{\rho}{\epsilon_{si}} \quad (\text{I.11})$$

où $\epsilon_{Si} = \epsilon_{sc}\epsilon_0$ et ϵ_{sc} est la permittivité relative du semi-conducteur. On utilise (I.10) et (I.11) se transforme :

$$\frac{d^2\psi}{dy^2} = \frac{q}{\epsilon_{Si}} \{ n_0 (\exp^{\beta\psi(y)} - 1) - p_0 (\exp^{-\beta\psi(y)} - 1) \} \quad (\text{I.12})$$

Avec $N_A = p_0$ et $n_0 = n_i^2/N_A$ pour un substrat de type p, une première intégration de l'équation de Poisson amène à [1] :

$$\frac{d\psi}{dy} = -\xi(y) = \pm \sqrt{\frac{2kT}{\epsilon_{si}}} \sqrt{n_0 (\exp^{\beta\psi(y)} - \beta\psi(y) - 1) + p_0 (\exp^{-\beta\psi(y)} + \beta\psi - 1)} \quad (\text{I.13})$$

On prendra le signe + pour $\psi < 0$, (c'est à dire pour $Q_{sc} > 0$ et le signe - pour $\psi > 0$ ($Q_{sc} < 0$). De (I.13), on déduit l'expression du champ électrique ξ à l'interface ($y = 0$) soit pour $\psi(0) = \psi_S$:

$$\xi_S = \pm \sqrt{\frac{2kT}{\epsilon_{si}}} \sqrt{n_0 (\exp^{\beta\psi_S} - \beta\psi_S - 1) + p_0 (\exp^{-\beta\psi_S} + \beta\psi_S - 1)} \quad (\text{I.14})$$

On mettra le signe + pour $\psi > 0$ et un - pour $\psi < 0$.

Le calcul de la densité de charge totale du semi-conducteur se fait en utilisant le théorème de Gauss. Pour cela, on considère comme surface d'intégration un cylindre fermé, de section unitaire, d'axe Oy , et ayant une base dans le plan de l'interface Si-SiO₂ et l'autre dans le volume du semi-conducteur au delà de la zone désertée (i.e. dans la région neutre). $\xi = 0$ dans la région neutre. Sur l'autre face latérale du cylindre le champ est parallèle (axe Ox : *Source-Drain*) entraînant un flux nul et permettant d'écrire :

$$Q_{sc} = \pm \sqrt{2\epsilon_{si}kT} \sqrt{n_0 (\exp^{\beta\psi_S} - \beta\psi_S - 1) + p_0 (\exp^{-\beta\psi_S} + \beta\psi_S - 1)} \quad (\text{I.15})$$

avec cette fois-ci un signe + pour $\psi > 0$ et un - pour $\psi < 0$.

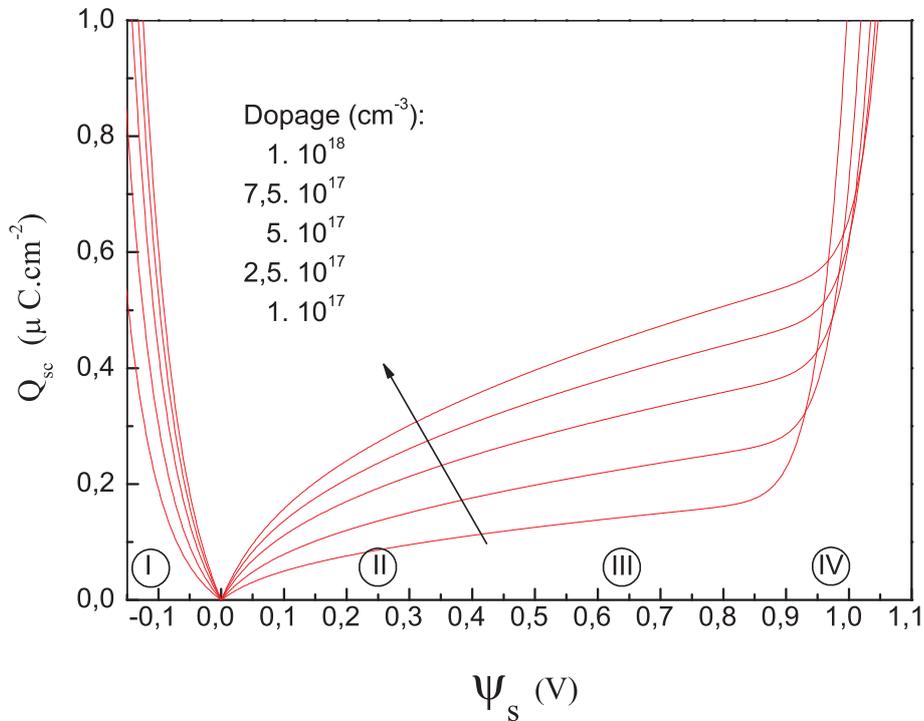


FIG. I.2 – Calcul de la variation de la charge du semi conducteur Q_{sc} avec (I.15) en fonction du potentiel de surface ψ_s suivant le dopage du substrat, pour une capacité de type P.

Les variations de la charge du semi-conducteur sont représentées sur la figure I.2, où l'on voit clairement les 4 zones de fonctionnement de la structure :

- $\psi_S < 0$: Accumulation des majoritaires près de l'interface (Région I).
- $\psi_S = 0$: Concentrations des porteurs (majoritaires et minoritaires) constantes dans tout le semi-conducteur et égales aux concentrations des régions neutres.
- $\psi_S \in]0; \phi_F[$: Déplétion (Région II).
- $\psi_S \in [\phi_F; 2\phi_F[$: Création de la couche d'inversion. On a une agglomération de minoritaires à l'interface et l'inversion reste faible (Région III).
- $\psi_S \geq 2\phi_F$: Inversion forte (Région IV).

I.1.3 Capacité totale C

a) Calcul de la capacité

La capacité MOS idéale (sans défaut) est équivalente à la mise en série de C_{ox} et C_{sc} , respectivement les capacités de l'oxyde et du semi-conducteur ; ce qui s'écrit :

$$\frac{1}{C} = \frac{1}{C_{OX}} + \frac{1}{C_{SC}} \quad (\text{I.16})$$

où C_{SC} , qui se calcule en dérivant (I.15) par rapport à ψ_S , et C_{OX} sont donnés par :

$$C_{SC} = \sqrt{\frac{q\epsilon_{Si}\beta p_0}{2}} \frac{p_0 |1 - \exp^{-\beta\psi_S}| + n_0 |1 - \exp^{\beta\psi_S}|}{\sqrt{p_0 (\exp^{-\beta\psi_S} + \beta\psi_S - 1) + n_0 (\exp^{\beta\psi_S} - \beta\psi_S - 1)}} \quad (\text{I.17})$$

$$C_{OX} = \frac{\epsilon_{OX}}{T_{ox}} \quad (\text{I.18})$$

D'autre part, la conservation de la charge dans la structure permet d'écrire en l'absence de défauts :

$$(V_{GS} - \psi_S - \phi_{ms})C_{OX} = -Q_{SC}(\psi_S) \quad (\text{I.19})$$

$$V_{GS} = \psi_S + \phi_{ms} - \frac{Q_{SC}}{C_{OX}} \quad (\text{I.20})$$

La résolution numérique de (I.20) donne la relation qui lie la tension de grille au potentiel de surface. Ceci est représenté sur la figure I.3 pour différentes valeurs du dopage.

Lorsque le potentiel de surface est nul, les bandes d'énergies sont plates, et en l'absence d'états d'interface et de charges dans l'oxyde (cas de la Fig. I.3), la tension de bande plate V_{FB} est égale à la différence des travaux de sortie du semi-conducteur et du métal ϕ_{ms} . La capacité

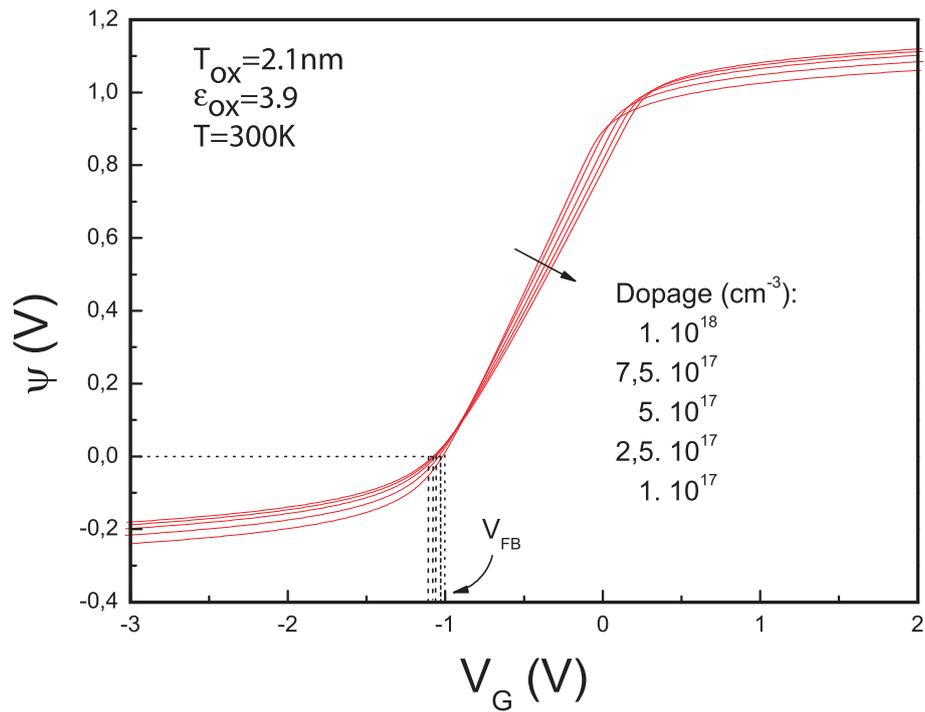


FIG. I.3 – Simulation de la relation entre le potentiel de surface et la tension appliquée sur la grille pour différentes valeurs du dopage du substrat.

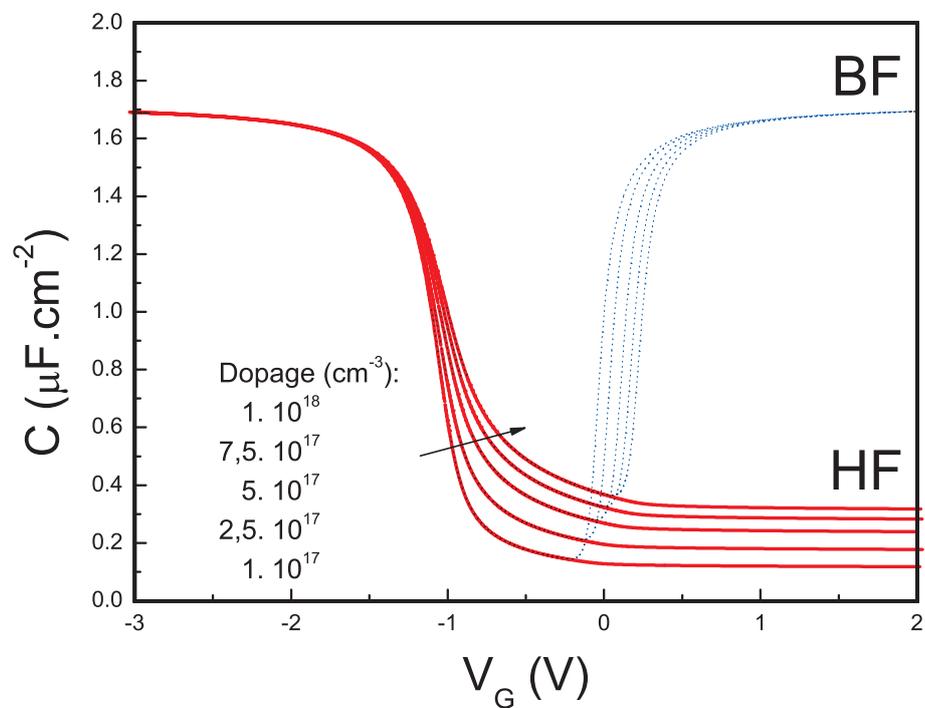


FIG. I.4 – Caractéristique simulée $C(V_{GS})$ pour différentes valeurs du dopage du substrat et une épaisseur d'oxyde $T_{ox} = 2.1 \text{ nm}$.

totale de la structure peut ainsi être associée à la tension de grille (Fig. I.4). Les différents régimes de fonctionnement sont :

- **Accumulation** ($\psi_S < 0$). La capacité du semi-conducteur est maximale et

$$\frac{1}{C_{OX}} \gg \frac{1}{C_{SC}} \quad \Rightarrow \quad C \approx C_{OX} \quad (\text{I.21})$$

- **Bandes plates** ($\psi_S = 0$). C_{SC} s'exprime comme ϵ_{Si}/L_D , autrement dit, la capacité du semi-conducteur est équivalente à une capacité d'épaisseur L_D [1] :

$$L_D = \sqrt{\frac{\epsilon_{Si} kT}{q^2 N_A}} \quad (\text{I.22})$$

qui est la *longueur de Debye* (longueur maximale sur laquelle les inhomogénéités peuvent perturber sensiblement la répartition spatiale des porteurs majoritaires). On a donc :

$$\frac{1}{C_{FB}} = \frac{1}{C_{OX}} + \frac{L_D}{\epsilon_{Si} S} \quad (\text{I.23})$$

Ce régime est obtenu pour $V_{GS} = V_{FB} = \phi_{ms}$.

- **Déplétion** ($0 < \psi_S < \phi_F$). Une zone de charge d'espace désertée de porteurs libres et d'épaisseur W :

$$W = \sqrt{\frac{2\epsilon_{Si}}{qNd} |\psi_S|} \quad (\text{I.24})$$

se forme sous l'oxyde et permet d'approximer :

$$C_{SC} = \frac{\epsilon_{Si} S}{W} \quad (\text{I.25})$$

- **“Mid gap”** ($\psi_S = \phi_F$). C'est la transition entre déplétion et inversion faible.
- **Inversion faible** ($\phi_F < \psi_S < 2\phi_F$) La couche d'inversion commence à se former et la charge du semi-conducteur est proportionnelle à $\sqrt{\psi_S}$.
- **Inversion forte** ($\psi_S \geq 2\phi_F$). La concentration des minoritaires présents à l'interface n'est plus négligeable et domine la charge de déplétion ; la capacité du semi-conducteur est donnée par :

$$C_{SC} = \sqrt{\frac{q\epsilon_{Si}\beta p_0}{2}} \frac{p_0 |1 - \exp^{-\beta\psi_S}| + n_0 |1 - \exp^{\beta\psi_S}|}{\sqrt{p_0 (\exp^{-\beta\psi_S} + \beta\psi_S - 1) + n_0 (\exp^{\beta\psi_S} - \beta\psi_S - 1)}} \quad (\text{I.26})$$

Afin de comparer les capacités simulées aux mesures, nous devons introduire le principe de mesure de la capacité différentielle.

b) Mesure de la capacité différentielle

Nous devons considérer que la capacité totale s'exprime par la variation de la charge stockée sur la Grille due à une variation de V_{GS} . Ceci se traduit par :

$$C = \frac{dQ_G}{dV_{GS}} \quad (\text{I.27})$$

Pour effectuer la mesure on applique sur la Grille un petit signal AC autour d'un point de polarisation DC. Durant la phase où le signal AC amène la structure légèrement au delà du point de polarisation, des paires électron-trous sont thermiquement générés dans la ZCE, à sa surface, ainsi que dans la zone neutre du substrat, donnant naissance la densité de courant J_{therm} :

$$J_{therm} = \frac{qn_i y_d}{\tau_g} \quad (\text{I.28})$$

où τ_g est le temps nécessaire pour générer une paire électron-trou par un processus thermique [2] et y_d l'épaisseur de la ZCE. En parallèle à ce phénomène, il existe dans le silicium une densité de courant de déplacement égale à :

$$J_{dep} = C \frac{dV}{dt} \quad (\text{I.29})$$

Pour que les porteurs minoritaires puissent suivre le signal AC imposé par la Grille, il faut que la condition $J_{dep} \leq J_{therm}$ soit vérifiée. Cette dernière se traduit par :

$$\frac{dV}{dt} \leq \frac{qn_i y_d}{\tau_g C_{OX}} \quad (\text{I.30})$$

Selon que (I.30) est vrai ou non, c'est à dire à haute (HF) ou basse (BF) fréquence on doit considérer les deux cas suivants :

$$HF : C_{SC} = \sqrt{\frac{q\epsilon_{Si}\beta p_0}{2}} \frac{|1 - \exp^{-\beta\psi_S}|}{\sqrt{\exp^{-\beta\psi_S} + \beta\psi_S - 1}} \quad (\text{I.31})$$

$$BF : C_{SC} = \sqrt{\frac{q\epsilon_{Si}\beta p_0}{2}} \frac{p_0 |1 - \exp^{-\beta\psi_S}| + n_0 |1 - \exp^{\beta\psi_S}|}{\sqrt{p_0 (\exp^{-\beta\psi_S} + \beta\psi_S - 1) + n_0 (\exp^{\beta\psi_S} - \beta\psi_S - 1)}} \quad (\text{I.32})$$

Un troisième cas reste à considérer quand les variations de palier de la composante DC du signal sont trop rapides pour générer la charge d'inversion. La couche d'inversion ne se forme pas et on parle de déplétion Profonde (DP).

I.1.4 Défauts dans la structure MOS

On appellera défauts les sites électriquement actifs du silicium ou de l'oxyde, pouvant perturber localement les densités de charges ou le champ vu par les porteurs. Sur la Fig.I.5 on peut distinguer les différents types de défauts qui peuvent apparaître dans la structure MOS. Il faut tenir compte des états d'interface (pièges de charges positives/négatives ou pièges neutres) le plus souvent constitués par des liaisons pendantes $\text{Si}\equiv\text{Si}_3$ appelées centre P_b dans la littérature [3, 4]; mais aussi des défauts de l'oxyde (Charges fixes près de l'interface, des ions mobiles dans le volume ou encore pièges de charges) [5].

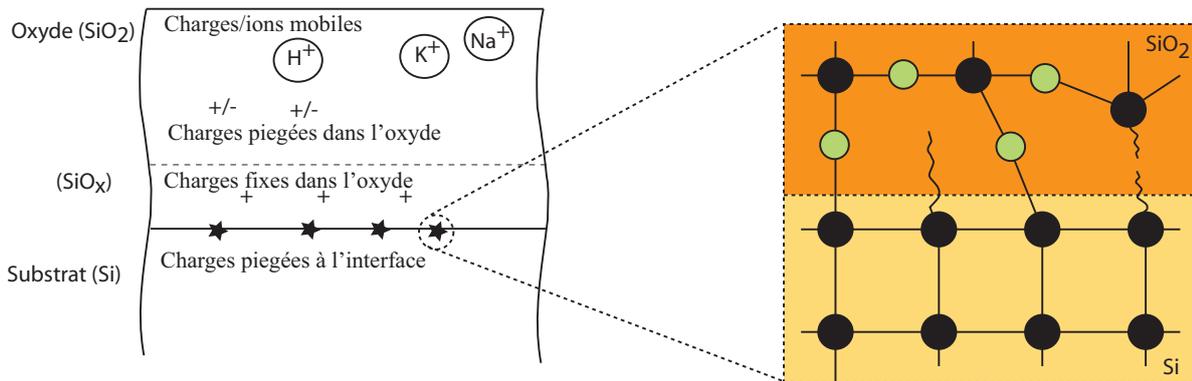


FIG. I.5 – Schéma représentant une structure Métal Oxyde Semi-conducteur avec les différents pièges de charges.

a) Défauts d'interface

Les pièges d'interface peuvent apparaître pendant les procédés de fabrication, les défauts sont la conséquence de la croissance thermique de l'oxyde (SiO_2) sur le silicium (Si). Les pièges sont alors matérialisés par une liaison pendante ou une lacune interstitielle due au mauvais alignement de SiO_2 sur Si. Ces défauts d'interface peuvent également être créés par des mécanismes physiques durant le fonctionnement de la structure : impacts de porteurs énergétiques générés dans le volume du silicium, irradiations extérieures, hautes températures... Dans la statistique de Shockley, Read et Hall (SRH) [6], les défauts d'interface sont caractérisés par leur niveau d'énergie et leur section efficace de capture/émission. On note $Q_{it}(\text{C.cm}^{-2})$ la charge piégée sur les défauts d'interface, $N_{it}(\text{cm}^{-2})$ le nombre de charges piégées et $D_{it}(\text{eV}^{-1}.\text{cm}^{-2})$ le nombre de pièges par niveau d'énergie. Ces défauts sont en contact avec la couche d'inversion et peuvent se vider et se remplir suivant la valeur du potentiel de surface. Ils ont une nature amphotère : ils peuvent capturer des électrons ou des trous (vis-à-vis des électrons, les pièges accepteurs sont chargés négativement si occupés et neutres sinon ; les pièges donneurs sont neutres si occupés et chargés positivement sinon). Ils peuvent être mis en évidence et quantifiés par des mesures capacité-tension (CV), par des mesures de pompages de charges à deux

ou trois niveaux (CP)[16, 8, 9], ou encore des mesures de type DLTS (“*Deep level transient spectroscopy*”)[10].

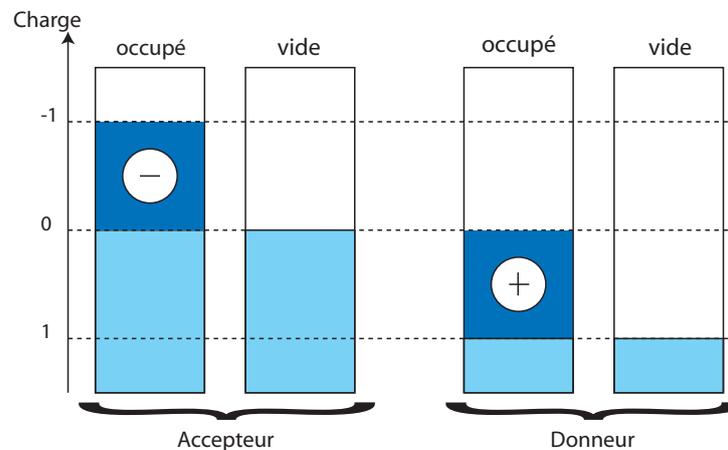


FIG. I.6 – représentation des pièges suivant leur type vis à vis des électrons.

b) Charges mobiles

Les charges mobiles dans l’oxyde sont essentiellement des impuretés ionisées venant de contaminations ioniques telles que N_A^+ , K^+ , Li^+ , H^+ . Les métaux lourds et certains ions négatifs peuvent devenir mobiles à haute température ($T > 500^\circ C$). On peut caractériser ces charges mobiles par des mesures $C(V)$ à différentes températures, modifiant ainsi leur mobilité et donc leur transport dans l’oxyde. Plusieurs techniques connues permettent de les mettre en évidence : CV (*capacitance-voltage* [11], TVS (*Triangular voltage sweep* [12]...). La charge mobile et la densité des espèces ionisées sont notés $Q_m(C.cm^{-2})$ et $N_m(cm^{-2})$.

c) Charges fixes

Les charges fixes sont généralement positives et ont pour origine les défauts structuraux du SiO_2 induits par le processus d’oxydation du silicium. Elles sont localisées près de l’interface (quelques Å, cela dépend de T_{ox}) mais ne répondent généralement pas aux variations du potentiel de surface. Leur quantité dans l’oxyde ($Q_f(C.cm^{-2})$, $N_f(cm^{-2})$) est liée à l’orientation du silicium et aux paramètres de fabrication (température de croissance et de recuit [13]). Enfin les procédés de fabrication récents pour les oxydes de grille ultra-fins rajoutent une nitruration en fin de durcir l’isolant, cette technique induit des charges positives supplémentaires mais réduit parallèlement la quantité états d’interface [14].

d) Charges piégées dans l’oxyde

Ces charges peuvent être positives ou négatives et trouvent leurs origines dans les phénomènes d’ionisation par irradiation (réversible par traitement thermique $T < 500^\circ C$), ou par

injection de porteurs plus ou moins énergétiques (mode d'avalanche, injection de porteurs chauds, courant tunnel). Enfin, pour ce type de charges, on utilise les notations $Q_{ot}(C.cm^{-2})$ et $N_{ot}(cm^{-2})$.

e) Conséquences

Les équations établies dans les paragraphes précédents, ne tiennent pas compte de la présence de l'ensemble de ces charges. Dans la mesure où les densités N_m et N_f peuvent être contrôlées par la qualité des méthodes de fabrication, nous ne les incluons pas dans la suite. En revanche, N_{it} et $N_{OX} = N_f + N_m + N_{ot} \approx N_{ot}$ sont des paramètres critiques.

La première conséquence est la modification de l'équation de la conservation de la charge (I.20) :

$$(V_{GS} - \psi_S - \phi_{ms})C_{OX} = -Q_{SC}(\psi_S) - Q_{it}(\psi_S) \quad (I.33)$$

N_{it} peut être obtenu par $Q_{it}(\psi_S) = qN_{it}(\psi_S)$ et :

$$N_{it}(\psi_S) = \int_{gap} N_{it}^+(E) (1 - f_s(E, \psi_S)) dE + \int_{gap} N_{it}^-(E) (f_s(E, \psi_S)) dE \quad (I.34)$$

où N_{it}^+ et N_{it}^- sont respectivement les densités de pièges donneurs et accepteurs et f_s la fonction de remplissage de la statistique de Fermi-Dirac :

$$f_s = \frac{1}{1 + \exp\left(\frac{E - E_{FS}}{kT}\right)} \quad (I.35)$$

La seconde conséquence est la modification locale de la valeur du champ électrique vertical $F_{OX}(y)$ défini par :

$$F_{OX} = \frac{V_{OX}}{T_{ox}} = \frac{V_{GS} - \psi_S - \phi_{ms} - Q_{OX}/C_{OX}}{T_{ox}} \quad (I.36)$$

où $Q_{OX} = qN_{OX}(C.cm^{-2})$ est la charge correspondante aux pièges présents dans SiO_2 . A travers l'oxyde, tout plan de charges parallèle à l'interface va donc modifier localement le champ. Clairement, ceci se traduit par une dérive de la tension de bande plate qui s'exprime alors par :

$$V_{FB} = \phi_{ms} - \frac{Q_{OX}}{C_{OX}} - \frac{qN_{it}}{C_{OX}} (\psi_S - \phi_F) \quad (I.37)$$

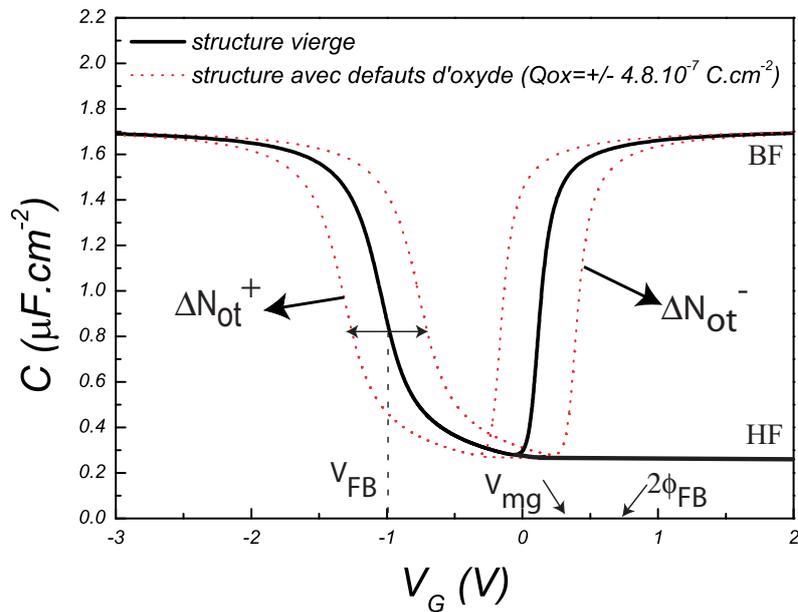


FIG. I.7 – Effet simulé de la présence de défauts dans l'oxyde sur une capacité de type P avec un oxyde de Grille de 2.1nm et un dopage de substrat de 4.10^{17}cm^{-3}

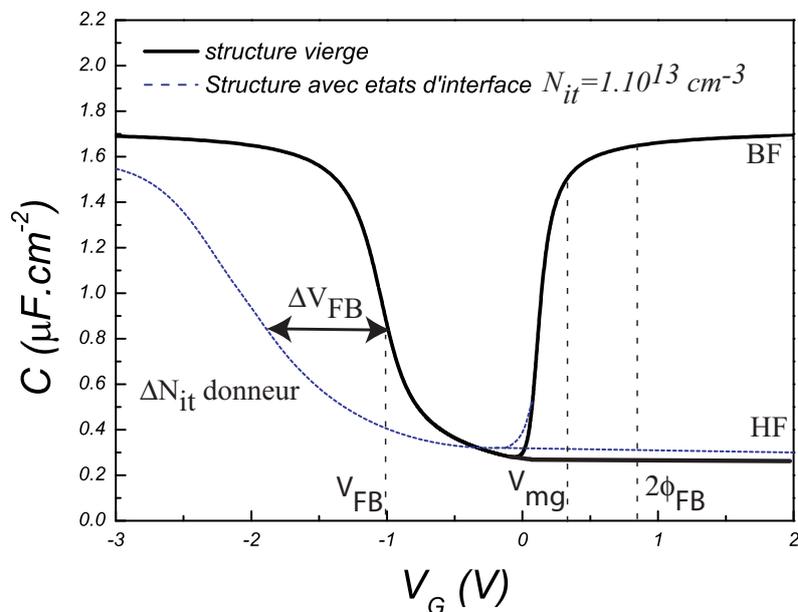


FIG. I.8 – Effet simulé de la présence de défauts à l'interface $\text{Si} - \text{SiO}_2$ sur une capacité de type P avec un oxyde de Grille de 2.1nm et un dopage de substrat de 4.10^{17}cm^{-3}

I.2.1 Modélisation du courant de Drain

a) Régime linéaire

Le courant total près de l'interface Si-SiO₂ est la somme des courants de diffusion et de conduction des porteurs libres. En inversion, on considère que la concentration de majoritaires est nulle dans le canal et on peut écrire [15] :

$$\vec{J}(x,y) = q \left(\mu_n n \vec{\xi} + D_n \vec{\nabla} n \right) = \vec{J}_n \quad (\text{I.40})$$

où μ_n est la mobilité des électrons en champ faible dans la couche d'inversion et D_n le coefficient de diffusion des électrons et $\vec{\nabla} n$ le champ électrique. Si on se place dans l'approximation graduelle de Shockley [16], c'est à dire lorsque le transistor fonctionne en mode non saturé ou encore lorsque le canal n'est pas pincé, on considère que les lignes de champ sont parallèles à l'interface, ce qui se traduit par :

$$\left| \frac{\partial^2 \psi}{\partial y^2} \right| \gg \left| \frac{\partial^2 \psi}{\partial x^2} \right| \quad (\text{I.41})$$

Dans ces conditions, (I.40) devient :

$$J_n = \underbrace{q\mu_n n \xi_x}_{J_1} + \underbrace{qD_n \text{grad}_x n}_{J_2} \quad (\text{I.42})$$

ξ_x étant le champ suivant l'axe x (figure I.9), J_1 est la composante de dérive (due au champ) et J_2 représente le courant de diffusion.

Considérant le coefficient d'Einstein $D_n = \mu_n kT/q$ où μ_n est la mobilité des électrons supposée constante, et sachant que ξ_x dérive d'un potentiel scalaire (ce qui se traduit par $\xi_x = -\text{grad}_x(\psi)$), (I.42) peut se simplifier en [15] :

$$J_n = -q\mu_n n \frac{d\phi_c}{dx} \quad (\text{I.43})$$

où les relations en potentiel sont données par :

$$\begin{aligned} \phi_c(x=0) &= -V_{BS} \\ \phi_c(x=L_G) &= V_{DS} - V_{BS} \\ \phi_c(0 < x < L_G) &= V(x) - V_{BS} \end{aligned}$$

Le courant total s'obtient en intégrant la densité J_n sur toute l'épaisseur de la couche d'inversion y_i du transistor de largeur W_G :

$$I_{DS} = -W_G \int_{y_i}^0 J_n dy = W_G \int_{y_i}^0 q\mu_n n \frac{d\phi_c}{dx} dy = W_G \mu_n \frac{d\phi_c}{dx} \int_{y_i}^0 n dy \quad (\text{I.44})$$

de plus I_{DS} étant constant tout le long du canal, il vient :

$$\int_{L_G}^0 I_{DS} dx = I_{DS} L_G = -W_G \mu_n \int_{-V_{BS}}^{V_{DS}-V_{BS}} Q_n d\phi_c \quad (\text{I.45})$$

où $Q_n = \int_{y_i}^0 n dy$. D'autre part l'influence de V_{BS} étant identique en tout point du canal sur les concentrations de porteurs libres, (I.45) se réécrit :

$$I_{DS} = -\frac{\mu_n W}{L_G} \int_0^{V_{DS}} Q_n(V) dV = -\frac{\mu_n W}{L_G} \int_0^{V_{DS}} (Q_{SC} - Q_D)(V) dV \quad (\text{I.46})$$

Q_{SC} peut être considéré comme la charge totale du semi-conducteur à laquelle on doit soustraire la charge Q_D de la zone désertée sous la zone d'inversion. Le calcul du courant relatif à la charge d'inversion Q_n se ramène donc au calcul de Q_{SC} et Q_D .

i. Calcul de Q_{SC} La charge Q_{SC} s'obtient à partir de (I.20) dans laquelle il faut intégrer l'influence de la polarisation appliquée au drain et éventuellement au substrat :

$$(V_{GS} - \underbrace{\phi_c + V_{BS} - 2\phi_F - \phi_{ms}}_{\psi_s}) C_{OX} + Q_{OX} + Q_{SC} = 0 \quad (\text{I.47})$$

$$Q_{SC} = -C_{OX} \left(V_{GS} - \phi_c - \phi_{ms} + \frac{Q_{OX}}{C_{OX}} - 2\phi_F \right) \quad (\text{I.48})$$

ii. Calcul de Q_D La densité de la charge de la zone désertée vaut $\rho = -qN_A$ et l'équation de Poisson permet d'écrire :

$$\frac{d\xi}{dy} = -q \frac{N_A}{\epsilon_{Si}} \Rightarrow \xi(y) = -q \frac{N_A}{\epsilon_{Si}} (y - y_d) \quad (\text{I.49})$$

où y_d est l'épaisseur de la zone désertée. On déduit ainsi l'expression du potentiel :

$$\psi(y) = \frac{qN_A}{2\epsilon_{Si}} (y - y_d)^2 \quad (\text{I.50})$$

En surface on a donc :

$$\xi_s = -q \frac{N_A}{\epsilon_{Si}} y_d^2 \quad (\text{I.51})$$

$$\psi_S = \frac{q N_A}{2 \epsilon_{Si}} y_d^2 \quad (\text{I.52})$$

$$(\text{I.53})$$

Si l'on tient compte du fait qu'en inversion forte on a $\psi_S = \phi_c - V_{BS} + 2\phi_F$ et que y_d est maximal ; il vient :

$$y_{dmax} = \sqrt{\frac{2\epsilon_{Si}}{qN_A} (\phi_c - V_{BS} + 2\phi_F)} \quad (\text{I.54})$$

La charge de la zone désertée est alors $Q_D = -qN_A y_{dmax}$, et devient :

$$Q_D = -\sqrt{2N_A \epsilon_{Si} (\phi_c - V_{BS} + 2\phi_F)} \quad (\text{I.55})$$

iii. Calcul de I_{DS} Pour obtenir l'expression de I_{DS} il faut maintenant reporter (I.48) et (I.55) dans (I.46), donnant :

$$I_{DS} = -\frac{\mu_n W_G}{L_G} C_{OX} \int_0^{V_{DS}} \left[V_{GS} - V - V_{FB} - 2\phi_F - \gamma (V - V_{BS} + 2\phi_F)^{1/2} \right] dV \quad (\text{I.56})$$

Soit encore après intégration :

$$I_{DS} = -\frac{\mu_n W_G}{L_G} C_{OX} \left[V \left(V_{GS} - \frac{V}{2} - V_{FB} - 2\phi_F \right) - \frac{2}{3} \gamma (V - V_{BS} + 2\phi_F)^{3/2} \right]_0^{V_{DS}} \quad (\text{I.57})$$

avec

$$V_{FB} = \phi_{ms} - \frac{Q_{OX}}{C_{OX}} \quad (\text{I.58})$$

$$\gamma = \frac{\sqrt{2qN_A \epsilon_{Si} \epsilon_0}}{C_{OX}} \quad (\text{I.59})$$

On arrive finalement à l'expression du courant de Drain en régime linéaire :

$$I_{DS} = \frac{W_G}{L_G} \mu_n C_{ox} \left[V_{DS} (V_{GS} - V_{DS} - V_{FB} - 2\phi_F) - \frac{2}{3} \gamma (V_{DS} - V_{BS} + 2\phi_F)^{3/2} - (-V_{BS} + 2\phi_F)^{3/2} \right] \quad (\text{I.60})$$

Pour des valeurs de V_{DS} suffisamment faibles, typiquement $V_{DS} \ll -V_{BS} + 2\phi_F$, un développement limité au second ordre conduit à :

$$(V_{DS} - V_{BS} + 2\phi_F)^{3/2} \approx (-V_{BS} + 2\phi_F)^{3/2} + \frac{3}{2} V_{DS} (-V_{BS} + 2\phi_F)^{1/2} + \frac{3}{4} V_{DS}^2 (-V_{BS} + 2\phi_F)^{-1/2} \quad (\text{I.61})$$

Le courant I_{DS} se réécrit alors sous la forme simplifiée :

$$I_{DS} = \frac{W_G}{L_G} \mu_n C_{ox} V_{DS} \left[\underbrace{V_{GS} - V_{FB} - 2\phi_F - \gamma (-V_{BS} + 2\phi_F)^{1/2}}_{V_{Th}} - V_{DS} \frac{1}{2} \underbrace{\left(1 + \frac{\gamma}{2} (-V_{BS} + 2\phi_F)^{-1/2} \right)}_{\alpha} \right] \quad (\text{I.62})$$

ce qui permet d'exprimer le courant de Drain d'un transistor MOS en régime linéaire sous sa forme synthétique :

$$I_{DS} = \frac{W_G}{L_G} \mu_n C_{ox} \left[V_{GS} - V_{Th} - \frac{\alpha}{2} V_{DS} \right] V_{DS} \quad (\text{I.63})$$

avec les paramètres standards de fonctionnement du transistor MOSFET :

$$V_{Th} = V_{FB} + 2\phi_F + \gamma (-V_{BS} + 2\phi_F)^{1/2} \quad (\text{I.64})$$

$$\alpha = \left(1 + \frac{\gamma}{2} (-V_{BS} + 2\phi_F)^{-1/2} \right) \quad (\text{I.65})$$

b) Régime de saturation

Lorsque la tension de Drain augmente, la zone de charge de charge d'espace (ZCE) s'étend et réduit la charge d'inversion à la pointe du Drain. Le transistor rentre en mode de saturation lorsque V_{DS} est suffisamment grand pour que $Q_{inv}(x = L_{eff})$ soit quasiment nulle. Ceci se produit pour V_D égal à V_{DSat} . On a alors :

$$Q_n = -C_{ox} \left[V_{GS} - V_{DSat} - V_T - 2\phi_F - \gamma \sqrt{V_{DSat} - V_{BS} + 2\phi_F} \right] \approx 0 \quad (\text{I.66})$$

Cette équation quadratique en V_{DSat} se résout en :

$$V_{DSat} = V_{GS} + V_{fb} - 2\phi_F + \frac{\gamma^2}{2} \left(1 - \sqrt{1 + \frac{4}{\gamma^2} (V_{GS} - V_{fb} - V_{BS})} \right) \quad (\text{I.67})$$

L'expression (I.67) de V_{DSat} , bien qu'exacte, ne permet pas d'établir une expression pratique à utiliser pour la détermination du courant de Drain. Il existe une autre approche pour déterminer la tension de Drain de saturation. Il suffit de considérer que pour $V_{DS} > V_{DSat}$ le courant de Drain ne varie pratiquement plus (cas idéal) ce qui revient à résoudre :

$$G_d = \left. \frac{\partial I_{DS}}{\partial V_{DS}} \right|_{V_{DSat}} = 0 \quad (\text{I.68})$$

Avec (I.63), on arrive ainsi à :

$$V_{DSat} = \frac{V_{GS} - V_T}{\alpha} \quad (\text{I.69})$$

Ceci permet de retrouver l'expression de I_{DSat} couramment utilisée :

$$I_{DSat} = \frac{W_G}{L_G} \mu_n C_{ox} \frac{(V_{GS} - V_T)^2}{2\alpha} \quad (\text{I.70})$$

Les figures I.10, I.11, I.12 et I.13 montrent les caractéristiques $I_{DS}(V_{GS})$ et $I_{DS}(V_{DS})$ calculées (avec (I.63) et (I.70)) et mesurées, pour un transistor à canal long ($W_G/L_G = 10/10\mu m$) et un transistor à canal court ($W_G/L_G = 10/0.13\mu m$). On peut y observer une divergence entre le modèle et la mesure lorsque les tensions de Grille et de Drain augmentent. Sur les petites géométries, l'écart s'accroît, ce qui met en évidence la nécessité d'affiner les équations en incluant les effets liés à la diminution des dimensions des transistors MOSFET. Dans la section suivante, nous allons présenter ces effets, en partant de leurs origines physiques, pour arriver à une quantification mathématique permettant d'inclure leurs impacts sur le comportement électrique des dispositifs. Nous allons en particulier décrire la réduction de la mobilité des porteurs

dans le canal sous l'influence du champ électrique, la modulation de la longueur du canal liée à l'extension de la zone de charge d'espace de la jonction Drain-Substrat.

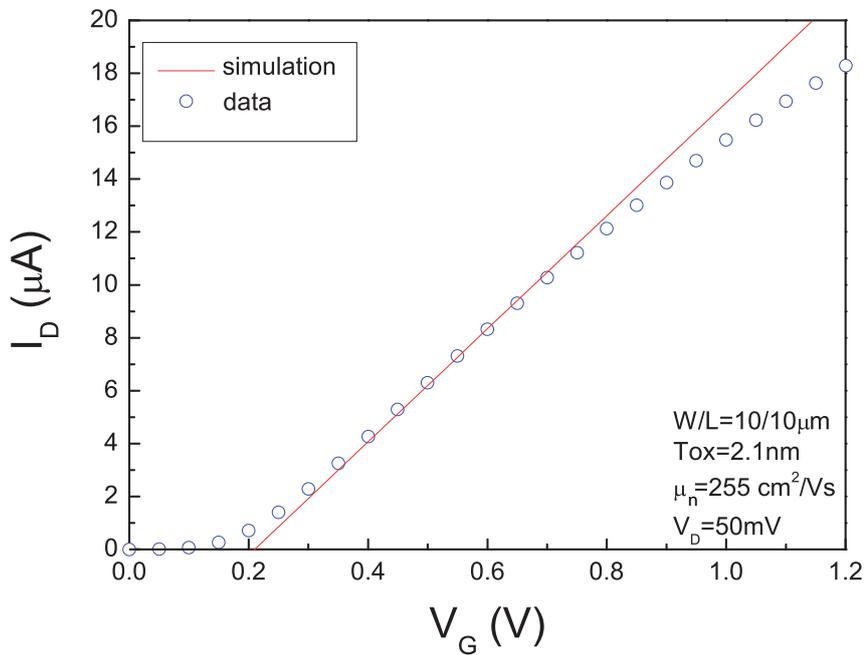


FIG. I.10 – Comparaison entre la modélisation et la mesure du courant de Drain en régime linéaire pour un transistor MOS à canal N long ($W_{eff}/L_{eff} = 10/10 \mu m$) et oxyde ultra mince $T_{ox} = 2.1 nm$.

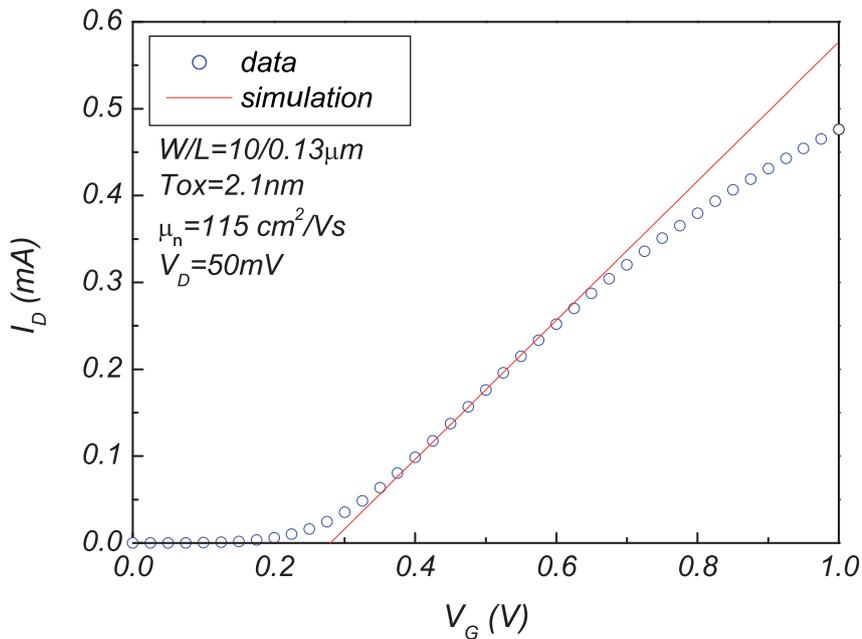


FIG. I.11 – Comparaison entre la modélisation et la mesure du courant de Drain en régime linéaire pour un transistor MOS à canal N court ($W_{eff}/L_{eff} = 10/0.13 \mu m$) et oxyde ultra mince $T_{ox} = 2.1 nm$.

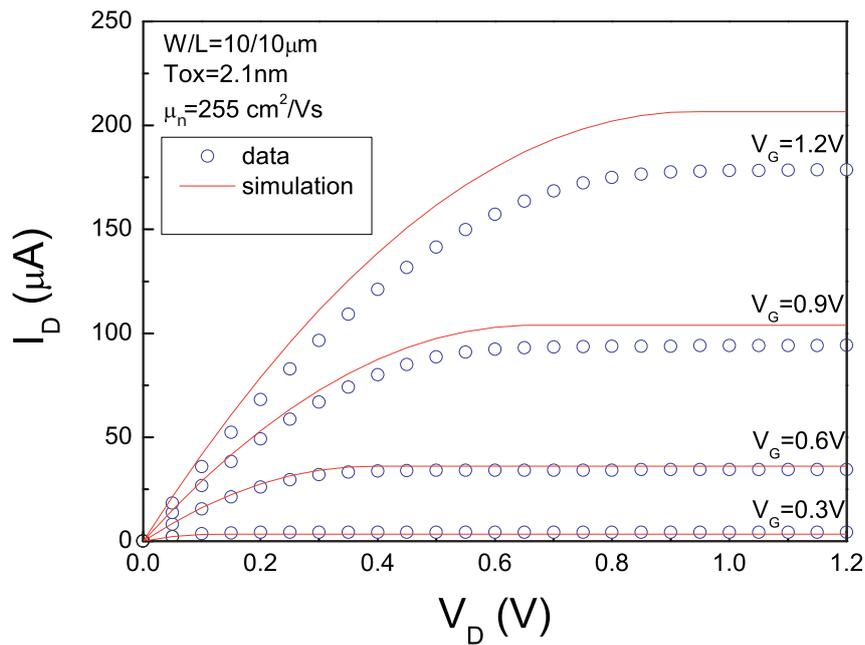


FIG. I.12 – Comparaison entre la modélisation et la mesure du courant de drain en fonction de la tension de Drain pour différentes valeurs de la tension de Grille pour un transistor MOS à canal N long ($W_{eff}/L_{eff} = 10/10\mu m$) et oxyde ultra mince $T_{ox} = 2.1nm$.

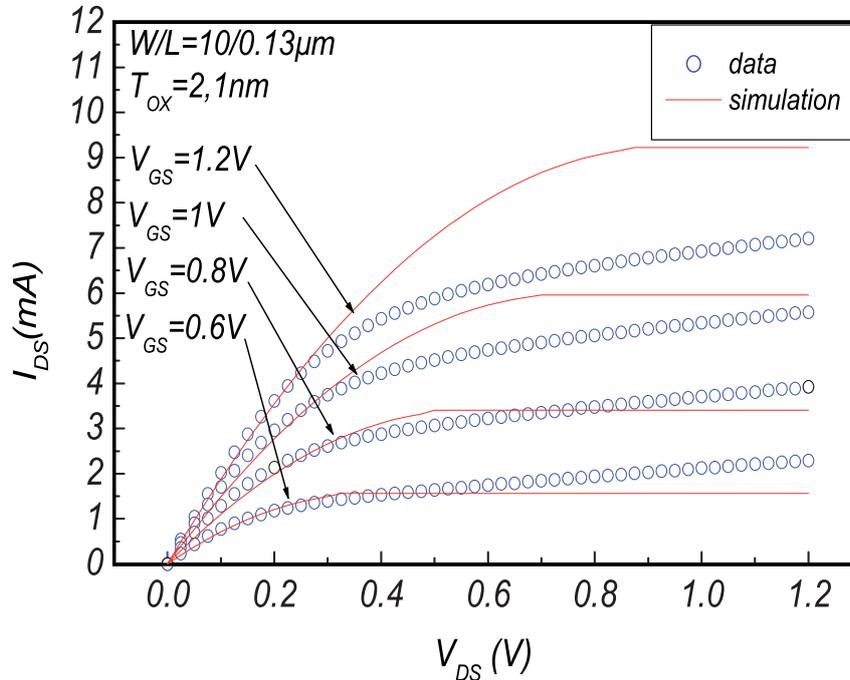


FIG. I.13 – Comparaison entre la modélisation et la mesure du courant de drain en fonction de la tension de Drain pour différentes valeurs de la tension de Grille pour un transistor MOS à canal N court ($W_{eff}/L_{eff} = 10/0.13\mu m$) et oxyde ultra mince $T_{ox} = 2.1nm$.

I.3 Effets liés à la réduction des dimensions des transistors MOSFETs

La réduction des dimensions des transistors utilisés dans les circuits CMOS, implique l'ajustement des modèles existants, afin de tenir compte des effets parasites. Ceux-ci sont induits par :

1. la réduction des dimensions et la perte de contrôle des charges vis à vis du potentiel de Grille
2. la complexité croissante des transistors avec la modification des structures de Source et de Drain, des caissons, la présence de "halo" et de "pockets".

Ainsi on peut citer les plus répandus dans les technologies actuelles : l'effet canal court "Short Channel Effect"(SCE), la réduction de la hauteur de barrière induite par la polarisation du Drain "Drain induced Barrier Lowering" (DIBL), la réduction de la mobilité des porteurs avec l'augmentation du champ latéral $\vec{\xi}$ et du champ vertical F_{OX} , la modulation de la longueur du canal en forte saturation "Channel Length Modulation "(CLM). Ces effets vont être abordés dans cette partie.

I.3.1 Effet de la géométrie du canal

Une variation importante des paramètres de fonctionnement du transistor MOS est constatée avec la réduction des dimensions W_G et L_G . Cet effet est l'illustration de l'évolution des procédés de fabrication et du niveau d'intégration. Les variations du dopage entre la surface (N_{surf}) et le volume du semi-conducteur (N_{dop}), on observera l'effet canal court (SCE pour "short channel effect") ou canal court inverse (RSCE pour "reverse short channel effect") :

$$\begin{aligned} N_{surf} > N_{vol} &\rightarrow \text{SCE} \\ N_{surf} < N_{vol} &\rightarrow \text{RSCE} \end{aligned}$$

Enfin, selon la méthode retenue pour l'isolation latérale des transistors, on pourra observer l'effet canal étroit (NCE pour "narrow channel effect").

a) Effet canal court

L'effet SCE se traduit par la réduction de la valeur de la tension de seuil avec la diminution de la longueur du canal (pour une technologie donnée). Ce phénomène s'explique par le fait que pour un canal court, la forme de la zone de déplétion n'est pas la même que pour un canal long [17] ; ceci est illustré sur la figure I.16. En effet, en inversion forte on peut considérer la tension de seuil en négligeant les effets des défauts ($Q_{OX} = 0$), qui peut s'exprimer en fonction de la forme de la zone déplétée avec :

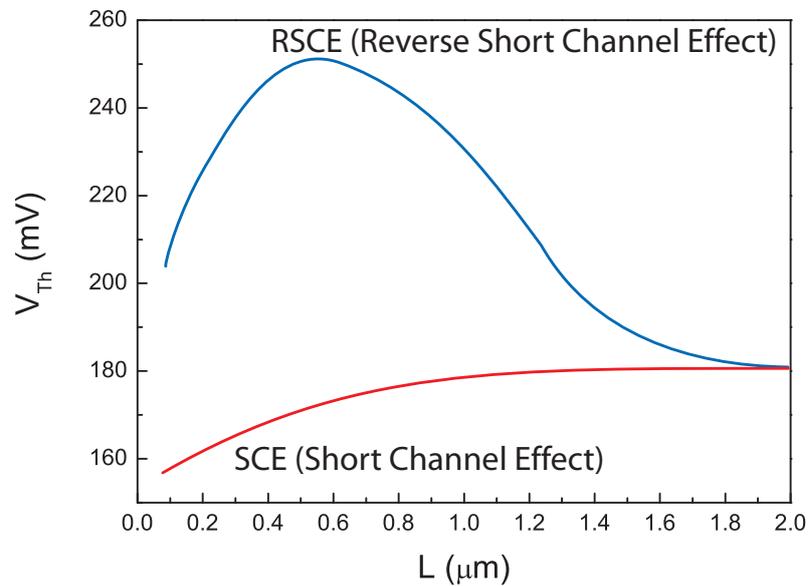


FIG. I.14 – Représentation schématique de la tension de seuil en fonction de la longueur du canal

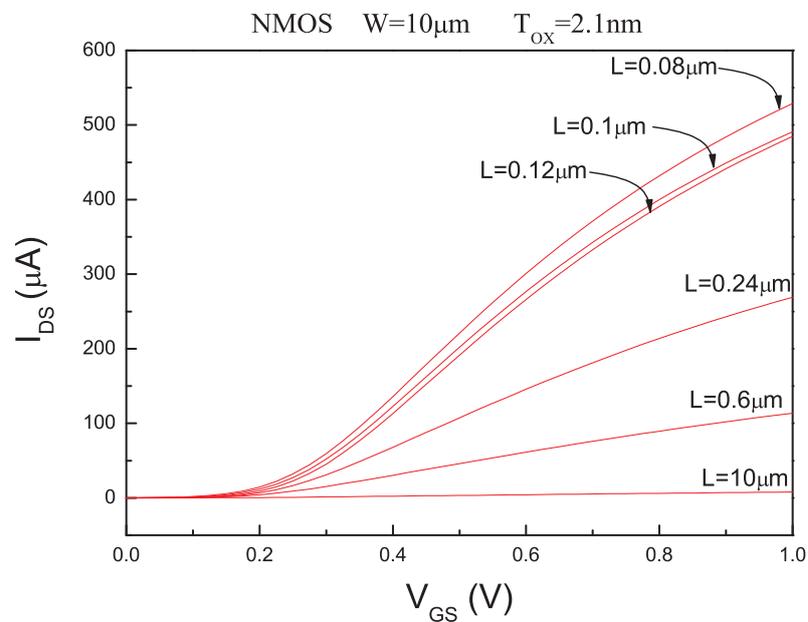
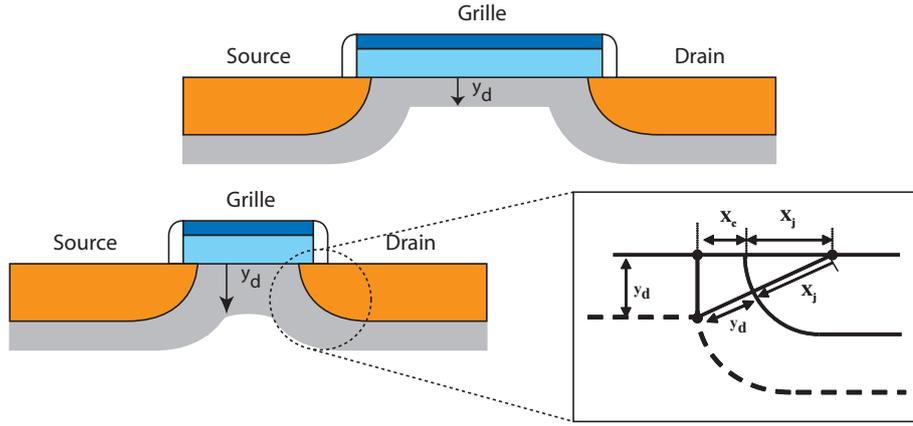


FIG. I.15 – Evolution de la caractéristique $I_{DS} - V_{GS}$ pour les transistor NMOS en fonction de la longueur du canal, W étant fixée


 FIG. I.16 – Effet de la réduction de L_{eff} sur la forme de la zone de charge d'espace

$$Q_D = qN_D\nu_D = qN_D y_d \frac{(L - L')}{2L} \quad (\text{I.71})$$

où ν_D est le volume occupé par la zone de charge d'espace et :

$$L' = L_G - 2X_c \quad (\text{I.72})$$

$$X_c = X_j \left(\sqrt{1 + \frac{2y_d}{X_j}} - 1 \right) \quad (\text{I.73})$$

Où X_j est la profondeur de jonction. Pour un canal court, le rapport entre l'épaisseur de la ZCE contrôlée par les Source/Drain et celle tenue par la Grille, est plus important que pour un canal long. Ceci entraîne d'après (I.53) un potentiel de surface plus grand dans ces régions, rendant l'interface plus attractive pour les porteurs minoritaires du substrat. Pour deux longueurs différentes et une même tension de grille on aura donc un courant de Drain plus important sur le canal court du fait que V_{Th} est diminué [18]. Ceci est illustre par la Fig. I.15. La charge de la zone désertée s'exprime finalement comme :

$$\begin{aligned} Q_D &= qN_D \sqrt{\frac{2\epsilon_{Si}}{qN_A} (2\phi_F - V_{BS})} \frac{(L_G - L')}{2L_G} \\ &= C_{OX} \gamma F_S \sqrt{2\phi_F - V_{BS}} \end{aligned} \quad (\text{I.74})$$

avec

$$F_S = \frac{(L_G - L')}{2L_G} = 1 - \frac{X_c}{L_G} = 1 - \frac{X_j}{L_G} \left(\sqrt{1 + \frac{2y_d}{X_j}} - 1 \right) \quad (\text{I.75})$$

b) Effet canal court inverse

L'effet canal court inverse, "Reverse Short Channel Effect" (RSCE) a pour conséquence de donner à la relation $V_{Th}(L_G)$ une forme en cloche, centrée sur la longueur nominale de la

technologie MOS étudiée (figure I.14). Il peut être décrit classiquement par (I.73), où l'influence de la forme de la ZCE sur la charge de déplétion totale est de moins en moins importante avec l'augmentation de L_{eff} . On l'explique aussi par l'influence des profils de dopage au niveau de la Source et du Drain dans les structures LDD [19], où les gradients de dopages importants doivent être pris en compte pour le calcul de Q_D .

c) Effet canal étroit

L'effet canal étroit (NCE pour "narrow channel effect") correspond à l'augmentation de la tension de seuil avec la réduction de W_G . Cette tendance est aussi due à la forme de la ZCE suivant sa largeur : elle peut déborder de l'oxyde quand W_G est grand, sinon, pour des oxydes étroits, elle présente un écrasement dû à l'épaississement du SiO_2 sur le bord, on nomme ces structures "LOCOS" (LOCAL Oxidized Silicon). La figure I.17 illustre schématiquement l'épaississement de l'oxyde sur les bords de la structure dans le cas LOCOS. Sur cette même figure, est présenté le principe de la tranchée d'oxyde, dit STI ("Shallow-Trench-Isolation"), cette dernière méthode, plus avancée, améliore l'isolation et permet le rapprochement des transistors pour un gain de place ; elle réduit en outre l'effet NCE.

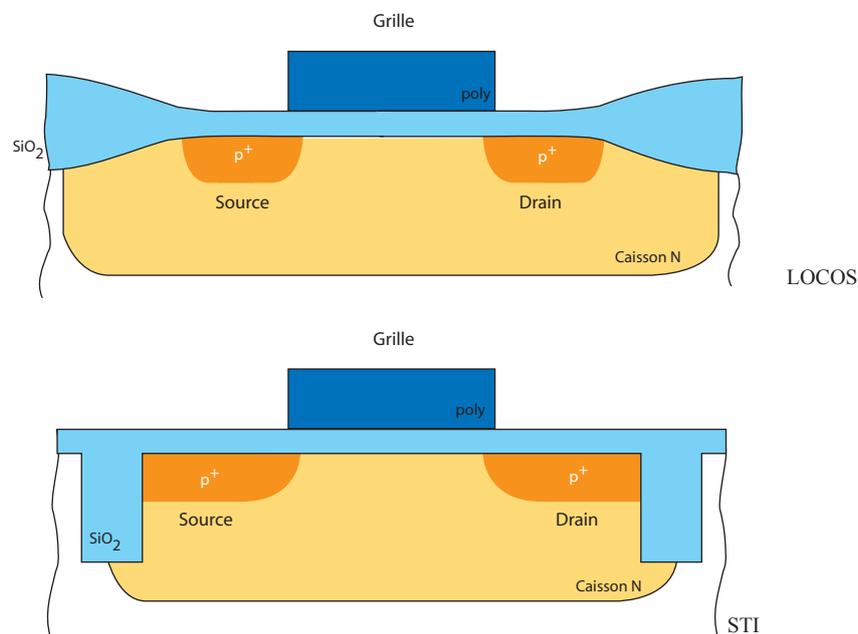


FIG. I.17 – Schéma de principe des méthodes d'isolation latérale LOCOS et STI

On traduit cet effet par le facteur F_N dans l'expression de V_{Th} :

$$F_N = \frac{\pi \epsilon_0 \epsilon_{Si}}{C_{OX} W_G} \quad (I.76)$$

La figure I.18 représente la zone de charge d'espace induite par la courbure particulière de l'oxyde imposée par la technique LOCOS.

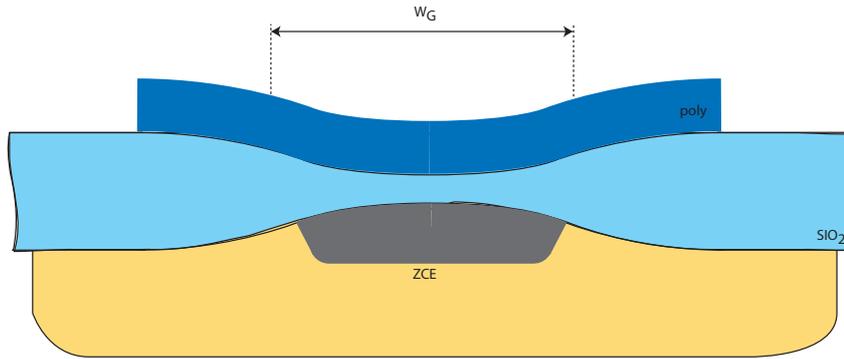


FIG. I.18 – Illustration de l'épaissement de l'oxyde sur les bords (LOCOS).

I.3.2 Effet DIBL

Les résultats précédents sont obtenus pour de faibles valeurs de la tension de Drain. Si on augmente cette dernière, la zone de charge d'espace croît davantage à la jonction Drain-Substrat. Ceci se traduit par un abaissement de la hauteur de barrière induit par la polarisation du Drain. Comme pour l'effet SCE on peut calculer la charge de la zone déplétée qui augmente avec V_{DS} . On peut exprimer V_{Th} comme une fonction de V_{DS} [20] :

$$V_{Th} = V_T - \sigma V_{DS} \quad (I.77)$$

$$\sigma = \frac{\epsilon_0 \epsilon_{Si}}{\pi C_{OX} L_G} \quad (I.78)$$

Le coefficient dépend de la longueur du canal, ceci apparaît bien sur la figure I.19 où V_{Th} est tracé en fonction de la tension de Drain, pour différentes valeurs de L_{eff} . Les trois effets précédents affectent essentiellement le paramètre de tension de seuil et peuvent être rassemblés dans une même formulation de V_{Th} pour transistor NMOS :

$$V_{Th} = V_{FB} + 2\phi_F + \gamma_N \sqrt{2\phi - V_{BS}} \underbrace{F_S}_{SCE} + \underbrace{F_N(2\phi - V_{BS})}_{NCE} - \underbrace{\sigma V_{DS}}_{DIBL} \quad (I.79)$$

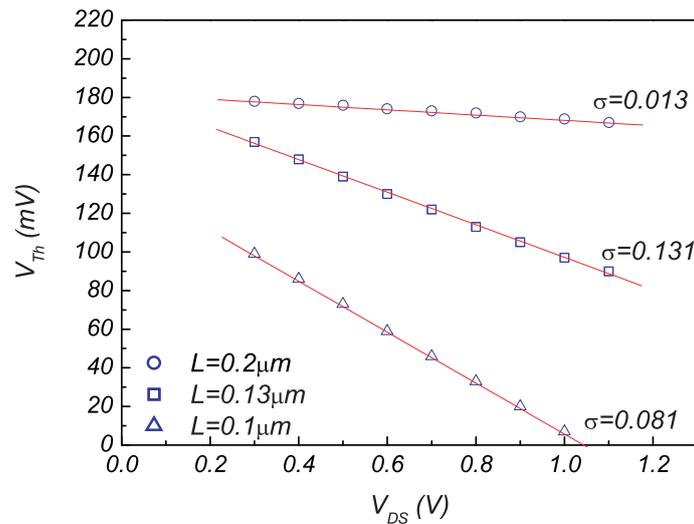


FIG. I.19 – Variation de la tension de seuil en fonction de la polarisation du Drain.

I.3.3 Réduction de la mobilité des porteurs

La mobilité des porteurs dans le silicium peut être sensiblement réduite par leur environnement immédiat : quantité de dopants, température [21]. Les champs électriques vertical (\vec{F}_{OX}) et latéral ($\vec{\xi}$) ont chacun un rôle important dans la dégradation de la mobilité des porteurs. On attribue la réduction par F_{OX} à trois phénomènes (figure I.20) :

- les interactions avec les phonons
- les interactions coulombiennes
- les interactions avec la rugosité de surface

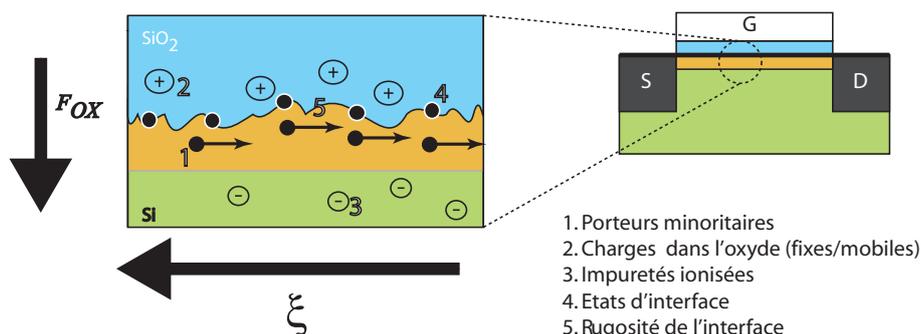


FIG. I.20 – Représentation des effets responsables de la réduction de la mobilité.

a) Collisions avec les phonons

A une température donnée, les porteurs sont en interaction avec les atomes du réseau cristallin qui vibrent autour de leur position d'équilibre. Cette interaction se traduit par le gain ou

la perte de vibration sous forme de quanta d'énergie appelés phonons [22]. A basse température ($T < 100K$), le réseau émet des phonons dit acoustiques, et les collisions avec ceux-ci provoquent des chocs quasi-élastiques. Pour des températures plus élevées ($300^\circ K$) et à faible champ, ce sont les phonons optiques qui prédominent et on assiste à des collisions inélastiques. Ces collisions ont pour effet de réduire la mobilité des porteurs.

b) Interactions coulombiennes

Les défauts de la structure cristalline précédemment cités, comme les états d'interface N_{it} , les charges fixes Q_f , ou encore les impuretés ioniques, induisent des modifications locales du potentiel vu par les porteurs de la couche d'inversion. Ainsi, leur mobilité varie inversement proportionnellement à la densité d'impuretés. Ces interactions peuvent également avoir lieu avec les charges piégées dans l'oxyde ou dans le silicium, près de l'interface. Dans ce cas, la couche créée en forte inversion peut faire écran pour les porteurs relativement éloignés de l'interface. Enfin ce type d'interaction, est plus important à basse température, lorsque les collisions sur les phonons sont moins prédominantes.

c) Collisions sur la rugosité de surface

Lorsque la couche d'inversion est importante, c'est à dire lorsque le champ électrique devient fort (quelques MV/cm), l'influence des interactions coulombiennes et les collisions avec les phonons devient faible. En revanche, un troisième phénomène, dû à la rugosité de l'interface, vient dégrader la mobilité. Les porteurs, fortement "plaqués" à l'interface, se voient freinés par les défauts structuraux liés à la modification cristallographique entre le SiO_2 sur le substrat Si et par les états d'interface.

Finalement pour les températures d'utilisation courante, autour de 300K, et en inversion forte, la mobilité des porteurs minoritaires est fortement réduite par la rugosité de l'interface du système SiO_2 - Si (voir figure I.20). Cette réduction se modélise empiriquement avec un facteur de réduction appliqué à la mobilité en champ faible μ_n , appelé θ .

d) Expression de la mobilité effective

Afin de tenir compte de cette variation de mobilité, on introduit la mobilité effective μ_{eff} dans l'équation du courant. Pour les transistors à canal long on introduit un facteur θ pour obtenir [23][24][25]:

$$\mu_{eff} = \frac{\mu_0}{1 + \theta(V_{GS} - V_{Th})} \quad (I.80)$$

Pour les technologies plus avancées avec des oxydes de grille présentant une rugosité d'interface plus marquée, on peut rajouter un deuxième facteur dans l'expression (I.80) pour tenir

compte de l'effet plus prononcé du champ vertical. La mobilité effective devient dans ce cas une expression quadratique de $(V_{GS} - V_{Th} - V_{DS}/2)$ [26] :

$$\mu_{eff} = \frac{\mu_0}{1 + \theta_1(V_{GS} - V_{Th} - \alpha \frac{V_{DS}}{2}) + \theta_2(V_{GS} - V_{Th} - \alpha \frac{V_{DS}}{2})^2} \quad (\text{I.81})$$

Enfin si le champ latéral est important et présente une forte influence sur la mobilité des porteurs du canal, on utilise une expression de la mobilité effective dans laquelle les influences des champs latéral et vertical sont décorrélées (BSIM 2) [27] :

$$\mu_{eff} = \frac{\mu_0}{1 + \theta_1(V_{GS} - V_{Th}) + \theta_2(V_{GS} - V_{Th})^2 + \theta_C V_{DS}} \quad (\text{I.82})$$

e) Saturation de la vitesse des porteurs avec le champ électrique latéral

La polarisation transversale qui existe le long de la couche d'inversion induit un champ électrique responsable du transport des porteurs minoritaires entre la Source et le Drain. La saturation du niveau de courant est essentiellement due à la saturation de la vitesse des porteurs dans le canal. Celle-ci s'exprime en fonction de la mobilité comme $\vec{v} = \mu \vec{\xi}$, μ est la mobilité. Le champ électrique s'exprime classiquement [28] par une relation en cosinus hyperbolique :

$$\xi(x) = \xi_{sat} \cosh\left(\frac{x - x_{sat}}{l}\right) \quad (\text{I.83})$$

$$l = \sqrt{\frac{\epsilon_{Si}}{\epsilon_{OX}} T_{ox} x_j} \quad (\text{I.84})$$

où ξ_{sat} est le champ de saturation au point de pincement x_{sat} (qui apparaît pour $V_{DS} = V_{Dsat}$ en $x_{sat} = l_{eff}$) et l est la longueur effective de la région de saturation de la vitesse des porteurs [29]. Le champ critique est relié la vitesse maximale des porteurs par :

$$v_{sat} = \mu_{sat} \xi_{sat} \quad (\text{I.85})$$

La valeur maximale du champ E_m s'exprime ainsi en fonction de la tension de saturation et de la valeur du champ critique :

$$E_m = \sqrt{\left(\frac{V_{DS} - V_{Dsat}}{l}\right)^2 + \xi_{sat}^2} \quad (\text{I.86})$$

f) Lien avec la tension de saturation $V_{Dsat}(\xi_{sat})$

Pour les transistors à canal court, la modélisation du Pic de champ doit se faire à l'aide d'une formulation plus précise de la tension de saturation. On peut considérer que le courant, d'une manière générale, est $I_{Dsat} = Q_{dw}\nu_{sat}$ où Q_{dw} est la densité de charge de la couche d'inversion par unité de surface et de temps, récoltée au drain :

$$Q_{dw} = W_G C_{OX} (V_{GS} - V_{Th} - \alpha_n V_{DS}) \quad (\text{I.87})$$

Par ailleurs, au point de pincement ($x_d = l_{eff}$) on peut écrire :

$$I_{DS} = \frac{\mu_{eff} C_{OX} W_G}{L_{eff}} \left[(V_{GS} - V_{Th}) V_{Dsat} - \frac{\alpha_n}{2} V_{Dsat}^2 \right] \quad (\text{I.88})$$

Il est alors possible d'obtenir l'expression de V_{Dsat} en égalisant (I.87) et (I.88) [27] :

$$\frac{\mu_{eff} C_{OX} W_{eff}}{L_{eff}} \left[(V_{GS} - V_{Th}) V_{Dsat} - \frac{\alpha_n}{2} V_{Dsat}^2 \right] = W_{eff} C_{OX} (V_{GS} - V_{Th} - V_{DS}) \nu_{sat} \quad (\text{I.89})$$

Cette équation quadratique en V_{Dsat} a comme solutions :

$$V_{Dsat} = \frac{(V_{GS} - V_{Th})}{\alpha_n} + \xi_{sat} L_{eff} \pm \sqrt{\left(\frac{V_{GS} - V_{Th}}{\alpha_n} \right)^2 + (\xi_{sat} L_{eff})^2} \quad (\text{I.90})$$

avec $\xi_{sat} = \nu_{sat}/\mu_{sat}$. En posant $a = (V_{GS} - V_{Th})/\alpha_n$ et $b = \xi_{sat} L_{eff}$, (I.90) se développe sous la forme :

$$V_{Dsat} = a + b - \sqrt{a^2 + b^2} \quad (\text{I.91})$$

Un développement limité de Taylor au premier ordre de la racine conduit à :

$$\begin{aligned} V_{Dsat} &\approx a + b - (a + b) \left(1 - \frac{ab}{(a + b)^2} \right) \\ &= \frac{ab}{a + b} \\ &= \frac{(V_{GS} - V_{Th}) \xi_{sat} L_{eff}}{(V_{GS} - V_{Th}) + \alpha_n x_{i_{sat}} L_{eff}} \end{aligned} \quad (\text{I.92})$$

I.3.4 Modulation de la longueur du canal

Le phénomène de la modulation effective du canal est lié la position du point de pincement dans le canal. la Fig. I.21 met en évidence cet effet : pour des tensions de drain supérieures à la tension de saturation, le point de pincement se déplace vers la source, et la longueur électrique de la couche d'inversion diminue. La modélisation du raccourcissement de L_{eff} permet d'obtenir l'expression (I.84) à l'aide de la résolution de l'équation de Poisson à une dimension :

$$\Delta L = X_D \left[\sqrt{V_{DS} - V_{Dsat} + \phi_D} - \sqrt{\phi_D} \right] \quad (I.93)$$

avec :

$$X_D = \sqrt{\frac{2\epsilon_0\epsilon_{si}}{qN_A}} \quad (I.94)$$

$$\phi_D = \frac{\epsilon_0\epsilon_{si}\xi_{sat}}{2qN_A} \quad (I.95)$$

où ξ_{sat} est le champ critique au point de pincement approximé par $\xi_{sat} = V_{Dsat}/L_{eff}$.

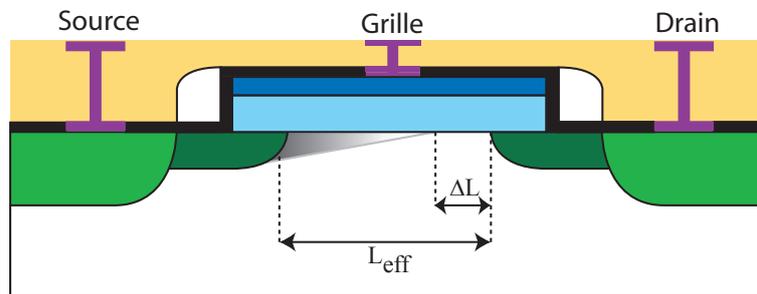


FIG. I.21 – Illustration de la modulation de la longueur du canal pour $V_{DS} > V_{Dsat}$.

L'erreur introduite par l'approximation de ξ_{sat} devient non négligeable lorsqu'on s'intéresse à des dispositifs de longueurs $L_{eff} \leq 0.15\mu m$ présentant des oxydes de grilles ultra-fins $T_{ox} \leq 3nm$. On peut alors exprimer le champ latéral sous forme d'un cosinus hyperbolique où le champ critique est calculé à partir de la vitesse de saturation des porteurs [30] :

$$\xi_{sat} = \frac{2\nu_{sat}}{\mu_0} \quad (I.96)$$

L'expression (I.86) du pic de champ latéral devient :

$$E_m = \sqrt{\left(\frac{V_{DS} - V_{Deff}}{l}\right)^2 + \xi_{sat}^2} \quad (I.97)$$

où V_{Deff} est une tension de saturation effective issue de Bsim 3v3 [31] :

$$V_{Deff} = V_{Dsat} - \frac{1}{2} \left[V_{Dsat} - V_{DS} - \delta_s + \sqrt{(V_{Dsat} - V_{DS} - \delta_s)^2 + 4\delta_s V_{Dsat}} \right] \quad (I.98)$$

δ_s est un paramètre fixé, qui permet d'ajuster la transition entre les régimes linéaire et saturé. Dans cette modélisation, le calcul de la distance du point de pincement au drain se fait par :

$$\Delta L_G = l \ln \left[\frac{(V_{DS} - V_{deff})/l + E_m}{\xi_{sat}} \right] \quad (I.99)$$

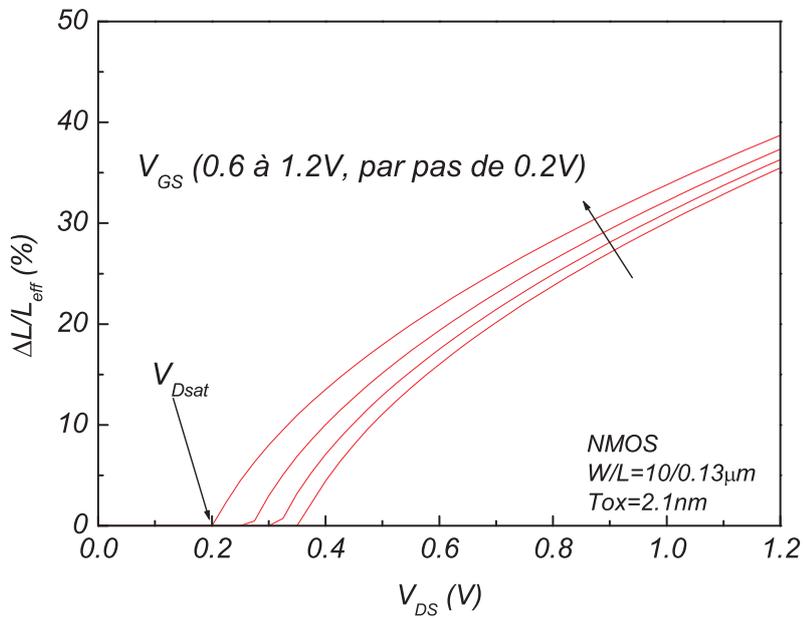


FIG. I.22 – Illustration de l'effet CLM.

Sur la figure I.22, on constate la variation logarithmique de ΔL avec la tension de Drain pour V_{GS} fixé à V_{DD} , l'extension de la zone de charge d'espace peut atteindre jusqu'à 35% de la longueur physique du canal.

I.3.5 Application à la modélisation du courant de Drain

Afin d'évaluer l'importance de ces effets sur le comportement électrique des transistors, nous avons comparé simulations et mesures pour des transistors à canal long et court. Pour le canal long, l'introduction de la mobilité effective à un coefficient θ_1 (I.80) a permis de corriger les divergences observées sur les figures I.10 et I.12. En revanche, le canal court a nécessité plus de raffinement dans la modélisation : prise en compte de la réduction de mobilité à trois coefficients θ_1 , θ_2 et θ_C (I.82), rajout des effets CLM et DIBL :

$$I_{DS} = \frac{\mu_0 C_{OX} W_{eff} (V_{GS} - V_{Th} - \frac{\alpha}{2} V_{DS}) V_{DS}}{1 + \theta_1 (V_{GS} - V_{Th}) + \theta_2 (V_{GS} - V_{Th})^2 + \theta_C V_{DS}} \cdot \frac{1}{L_{eff} \left(1 - \frac{\Delta L}{L_{eff}}\right)} \quad (\text{I.100})$$

Les résultats sont présentés sur les figures I.23, I.24, I.25 et I.26. On constate de bons résultats pour le canal long en mode linéaire et saturé. La modélisation des régimes linéaire et saturé sur le canal court donne de bons résultats, en revanche, la transition entre les deux modes présente un écart avec la mesure. Ce problème peut être corrigé en introduisant des paramètres empiriques permettant d'améliorer le raccordement entre les caractéristiques simulées des différents régimes de fonctionnement.

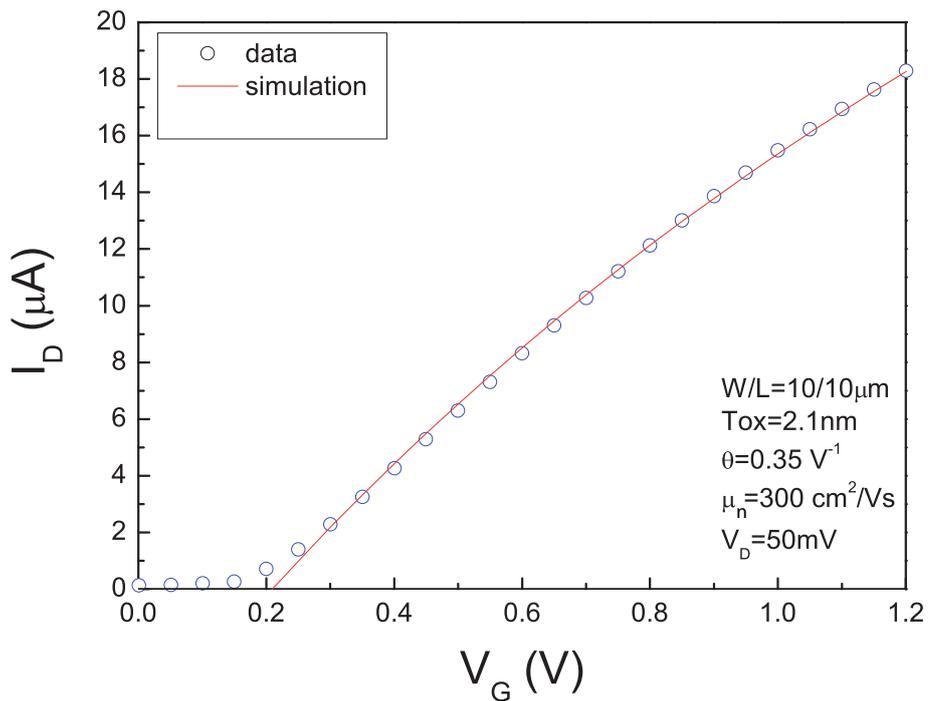


FIG. I.23 – Modélisation du courant de Drain suivant V_{GS} avec un facteur de réduction de mobilité en régime linéaire à un coefficient pour un canal N long ($W/L = 10/10\mu\text{m}$) et oxyde ultra mince $T_{ox} = 2.1\text{nm}$.

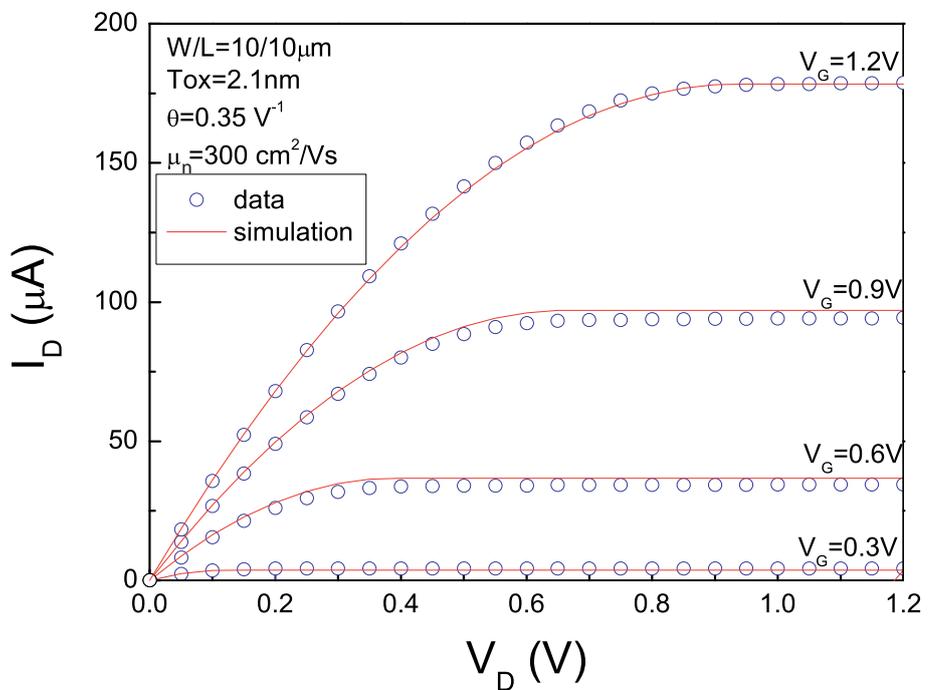


FIG. I.24 – Modélisation du courant de Drain suivant V_{DS} en régime saturé avec réduction de mobilité à un coefficient pour un canal N long ($W/L = 10/10\mu\text{m}$) et oxyde ultra mince $T_{ox} = 2.1\text{nm}$.

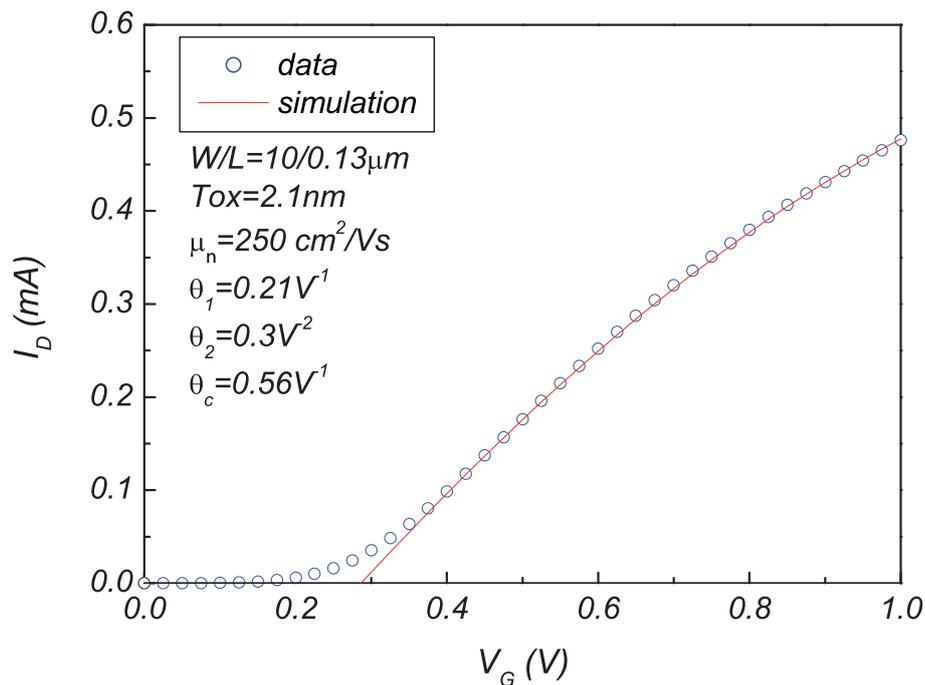


FIG. I.25 – Modélisation du courant de Drain suivant V_{GS} pour un canal N court ($W/L = 10/0.13\mu\text{m}$) et oxyde ultra mince $T_{ox} = 2.1\text{nm}$, en régime linéaire avec une réduction de mobilité à trois coefficients, incluant les effets CLM et DIBL.

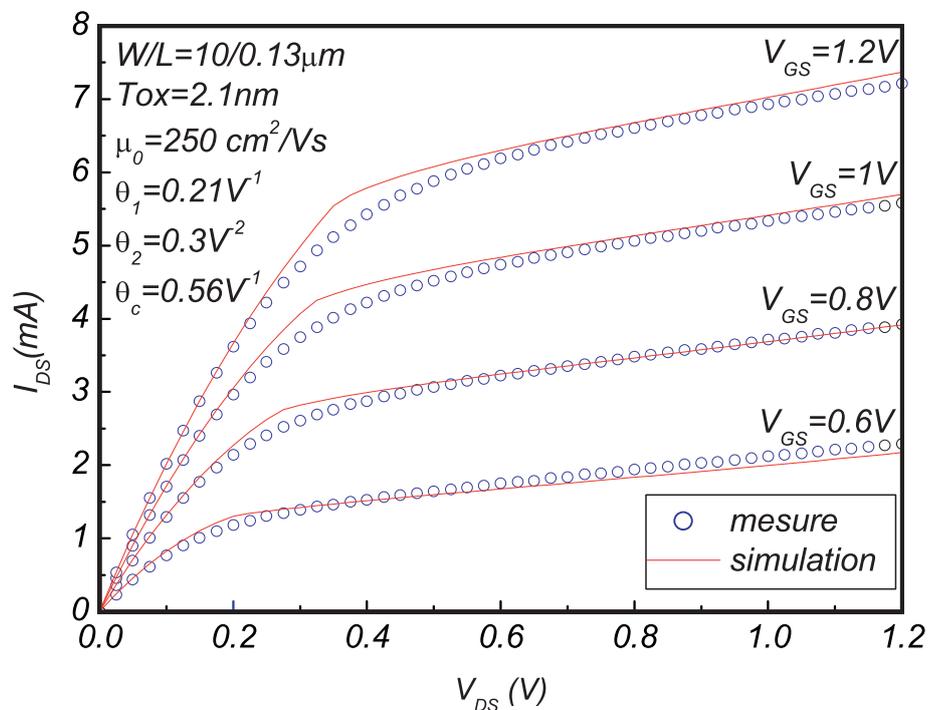


FIG. I.26 – Modélisation du courant de Drain suivant V_{DS} pour un canal N court ($W/L = 10/0.13\mu\text{m}$) et oxyde ultra mince $T_{ox} = 2.1\text{nm}$, en régime linéaire avec une réduction de mobilité à trois coefficients, incluant les effets CLM et DIBL.

I.4 Extraction des principaux paramètres

L'évaluation des performances d'un dispositif MOSFET s'obtient à l'aide des caractéristiques DC du transistor. On peut réaliser des mesures capacité-tension (C-V) ou courant-tension (I-V). Dans le premier cas, étudié plus haut, on peut remonter aux paramètres tels que le dopage N_{sub} , l'épaisseur de l'isolant de grille T_{ox} , sa permittivité relative ϵ_{OX} , la tension de bande plate V_{FB} , la densité de charges piégées dans l'oxyde et à l'interface N_{OX} et N_{it} ou encore au dopage $N_{a,d}$. La caractérisation I-V permet la détermination des paramètres de fonctionnement comme la tension de seuil V_{Th} en linéaire et en saturé, le facteur de forme γ , le (ou les) coefficient de réduction de la mobilité θ (ou θ_1 et θ_2), le courant de conduction I_{DS} , la mobilité en champ faible μ_0 , la longueur-largeur effective du canal L_G-W_G .

I.4.1 Extraction de la tension de seuil

La tension de seuil en régime linéaire se détermine en modélisant le courant de sortie $I_{DS}(V_{GS})$ par une droite. Plus précisément, on utilise sa tangente au point d'inflexion, soit au maximum de la transconductance du transistor définie par :

$$g_m = \left. \frac{\partial I_{DS}}{\partial V_{GS}} \right|_{V_{DS}, V_{BS}=Cte} \quad (I.101)$$

Ce raisonnement revient à négliger le facteur de réduction mobilité θ et les résistances d'accès

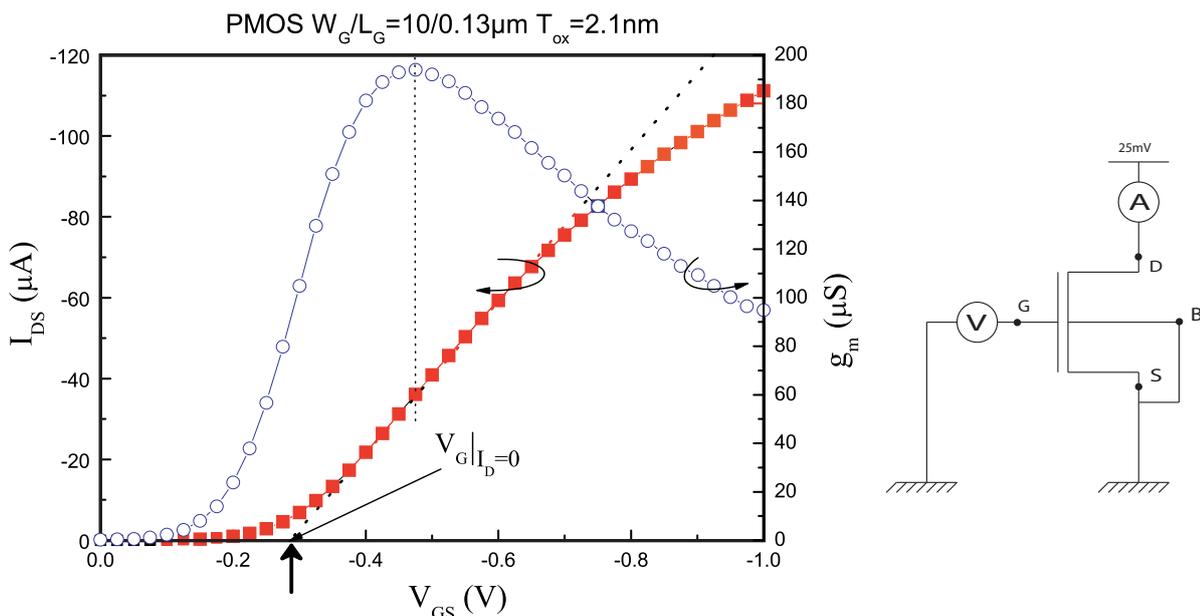


FIG. I.27 – Tracé du courant de Drain en fonction de la tension de grille pour $V_{DS} = 25\text{mV}$, à droite figure le principe de la mesure

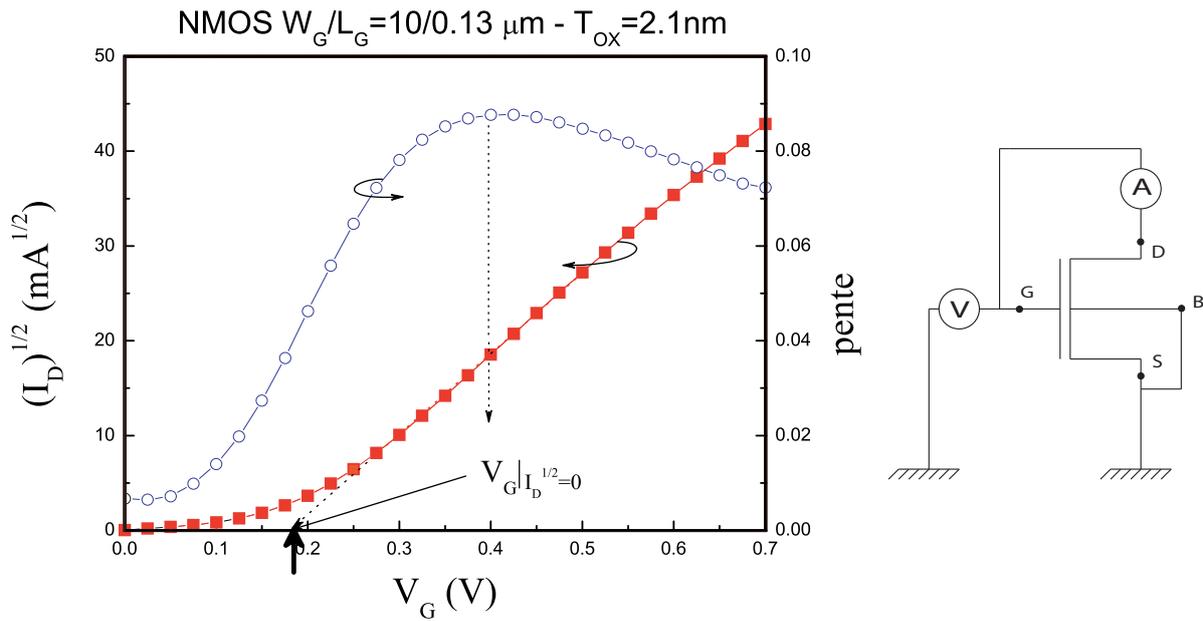


FIG. I.28 – Technique d'extraction de la tension de seuil pour le régime saturé, à droite le principe de la mesure

R_{SD} . On cherche V_{GS} solution de :

$$I_{DS} = KV_{DS} \left((V_{GS} - V_{Th}) - \frac{\alpha V_{DS}}{2} \right) = 0 \quad \Rightarrow \quad V_{Th} = V_{GS}|_{I_{DS}=0} - \frac{\alpha V_{DS}}{2} \quad (\text{I.102})$$

où $K = \mu_0 C_{OX} W_G / L_G$. Cette méthode se réalise à partir des relevés $I_{DS}(V_{GS})$, V_{DS} étant fixé (figure I.27). La pente KV_{DS} de la droite modélisant le courant donne également la mobilité en champ faible. Pour un transistor PMOS avec $L_G = 0.13\mu m$, $W_G = 10\mu m$ et $T_{ox} = 2.1nm$ (figure I.27), on trouve $V_{Th} = -278mV$ et $\mu_p = 60.98cm^2V^{-1}s^{-1}$. Pour un transistor NMOS de mêmes dimensions, on trouve $V_{Th} = 259mV$ et $\mu_p = 114.81cm^2V^{-1}s^{-1}$. Ces faibles valeurs mobilité peuvent s'expliquer par la forte interaction coulombienne due au fort dopage ($N = 4.10^{17}cm^{-3}$) de ces structures ; à faible champ, elle est responsable d'une réduction marquée de la mobilité.

Pour le mode de fonctionnement saturé, on fait la mesure avec $V_{DS} = V_{GS}$ de sorte que $V_{DS} > (V_{GS} - V_{Th}) \equiv V_{Dsat}$. Le principe reste le même : on approxime le courant de sortie par une droite, la tangente au maximum de g_m . Il nous faut résoudre :

$$\sqrt{I_{DS}} = \sqrt{\frac{K}{2\alpha}} (V_{GS} - V_{Th}) = 0 \quad \Rightarrow \quad V_{Th} = V_{GS}|_{\sqrt{I_{DS}}=0} \quad (\text{I.103})$$

Sur un transistor NMOS avec $L_G = 0.13\mu\text{m}$, $W_G = 0.13\mu\text{m}$ et $T_{ox} = 2.1\text{nm}$ (figure I.28), on obtient $V_{Tsat} = 159\text{mV}$ et $\mu_n = 163.9\text{cm}^2\text{V}^{-1}\text{s}^{-1}$

	PMOS	NMOS
Régime linéaire	-280 mV	270 mV
Régime saturé	-185mV	190 mV

TAB. I.1 – Tensions de seuil des transistors de type P et N pour les régimes saturé et linéaire.

I.4.2 Extraction du facteur de forme et du dopage

La polarisation du substrat (avec une tension de signe opposé à celle appliquée sur la grille) entraîne une augmentation de l'épaisseur de la couche de déplétion. Cette augmentation a pour conséquence d'accroître la tension de seuil. On appelle ce phénomène "effet substrat" ou "effet body". Pour déterminer le dopage et le facteur de forme, on met à profit cet effet substrat. La dépendance de la tension de seuil avec la tension substrat V_B est donnée par la relation (I.62) : $V_T = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F - V_{BS}}$. On peut remarquer que :

$$\Delta V_T = V_{T(V_B \neq 0)} - V_{T(V_B = 0)} = \gamma_N \left[\sqrt{2\phi_F - V_{BS}} - \sqrt{2\phi_F} \right] \quad (\text{I.104})$$

et ainsi tracer dans un graphe $\Delta V_T = f(\sqrt{2\phi_F - V_{BS}})$. La relation est linéaire et γ est la pente de la droite obtenue.

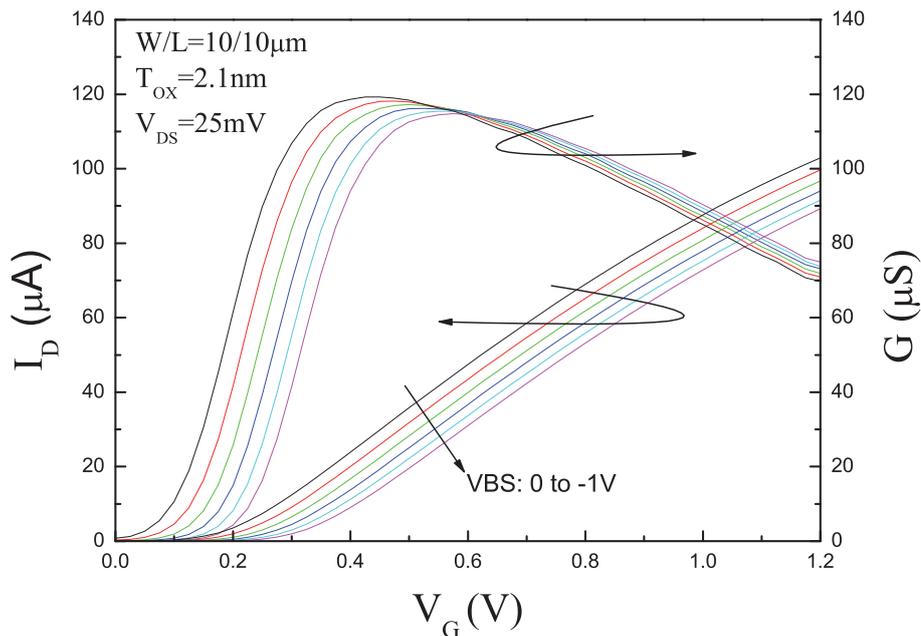


FIG. I.29 – Evolution de la caractéristique de sortie avec la tension V_{BS} .

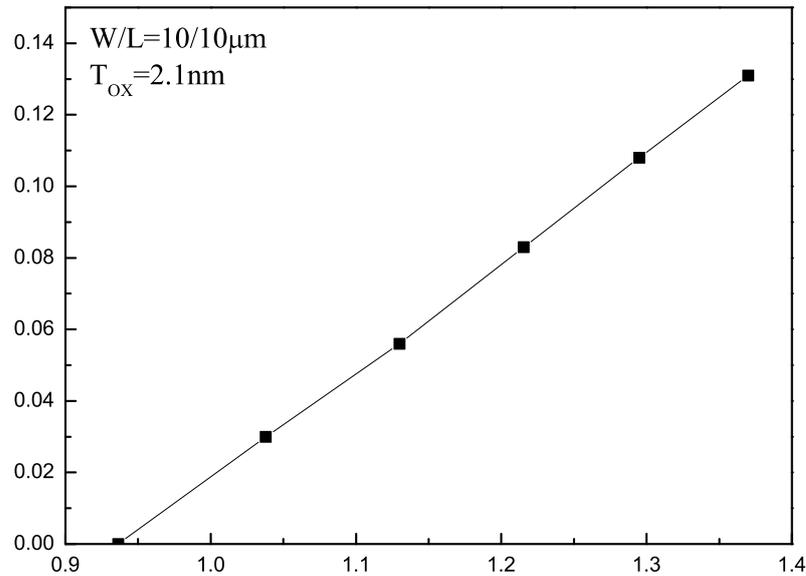


FIG. I.30 – Détermination de γ_N pour l'extraction du dopage.

Cette méthode, bien que ne donnant que la valeur moyenne du dopage en surface, permet de déterminer le facteur de forme et le dopage simultanément. Il faut procéder par dichotomie :

- on choisit une valeur N^i du dopage a priori correcte
- on calcule $\phi_F = \pm \frac{kT}{q} \ln \frac{N^i}{n_i}$
- on trace $\Delta V_T = f(\sqrt{2\phi_F} - V_{BS})$
- de $\gamma_N = \frac{\sqrt{2q\epsilon_{Si}N^f}}{C_{OX}}$ on déduit une valeur N^f du dopage
- on réitère l'opération pour faire converger le système vers la bonne valeur de N

Sur la figure I.30, on observe l'évolution de ΔV_{Th} suivant V_{BS} , les valeurs de la tension de seuil sont obtenues à partir de la figure I.29. Ici on a obtenu $\gamma_N = 0.5V^{1/2}$, et une valeur de dopage $N_A = 4.10^{17} cm^{-3}$.

I.4.3 Extraction du facteur de réduction de mobilité

Dans la majorité des cas on pourra modéliser la réduction de la mobilité des porteurs du canal par un seul coefficient. Pour extraire θ , on peut calculer g_m :

$$g_m = \frac{\partial I_{DS}}{\partial V_{GS}} = \frac{\partial}{\partial V_{GS}} \left[\frac{K(V_{GS} - V_{Th} - \alpha V_{DS}/2)}{1 + \theta(V_{GS} - V_{Th})} \right] \quad (I.105)$$

En négligeant $\alpha\theta V_{DS}/2$ devant 1, on arrive a :

$$g_m = \frac{K}{(1 + \theta(V_{GS} - V_{Th}))^2} \quad (\text{I.106})$$

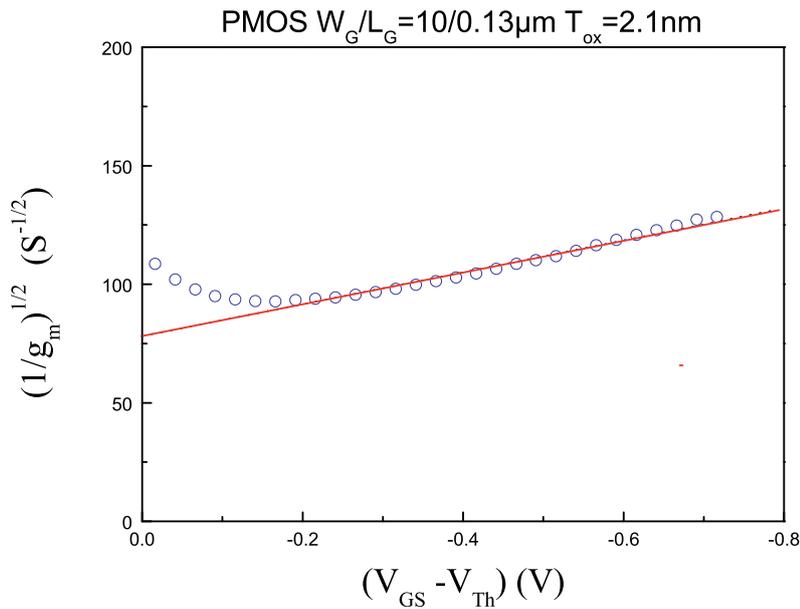


FIG. I.31 – Extraction de μ_P et θ_P avec le modèle de réduction de mobilité à un coefficient.

Il est à noter que si α n'est pas encore connu, on pourra prendre $\alpha \approx 1$, ce qui introduit une erreur de seulement quelques pour cent. Il suffit de tracer $g_m^{-1/2}$ en fonction de $(V_{GS} - V_{Th})$ (figure I.31). Ce graphe permet d'extraire la mobilité μ_p en champ faible, et le coefficient de réduction de mobilité θ . Pour un transistor à canal P de dimension $W_G = 10\mu\text{m}$ et $L_{eff} = 0.13\mu\text{m}$, on trouve $\mu_{P0} = 117\text{cm}^2\text{V}^{-1}\text{s}^{-1}$ et $\theta_P = 0.75\text{V}^{-1}$

I.4.4 Détermination des dimensions effectives du canal

Il existe systématiquement une différence entre les dimensions dessinées sur le masque de gravure et la taille réelle du canal. Ces variations sont dues aux procédés de fabrication, et sont donc identiques pour l'ensemble des dispositifs élaborés avec la même technologie. La différence ΔL_G sur la longueur électrique du canal est la conséquence à la fois du phénomène de surgravure et de la diffusion latérale des dopants des zones Source/Drain. Quand à ΔW_G il découle de la surgravure et du type d'isolation latérale (LOCOS ou STI).

a) Extraction de ΔL_G

L'extraction de la longueur électrique $L_{eff} = L_G - \Delta L_G$ peut s'effectuer en disposant d'un lot de transistors de longueur variable et de largeur fixé. Pour chaque transistor on mesure $I_{DS}(V_{GS})$ afin d'obtenir le maximum de la transconductance. Ceci permet de négliger la variation de mobilité en écrivant :

$$I_{DS} = \frac{\mu_0 C_{OX} W_G}{L_G - \Delta L_G} V_{DS} (V_{GS} - V_{Th} - \alpha V_{DS}/2) \Rightarrow \frac{1}{g_m} = \frac{1}{\mu_0 C_{OX} W_G V_{DS}} (L_G - \Delta L_G) \quad (I.107)$$

Cette relation est linéaire et le graphe de G_{max}^{-1} en fonction de la longueur dessinée L_G (figure I.32), permet d'extraire ΔL_G qui est l'abscisse à l'ordonnée et $L_{eff} = L - \Delta L$.

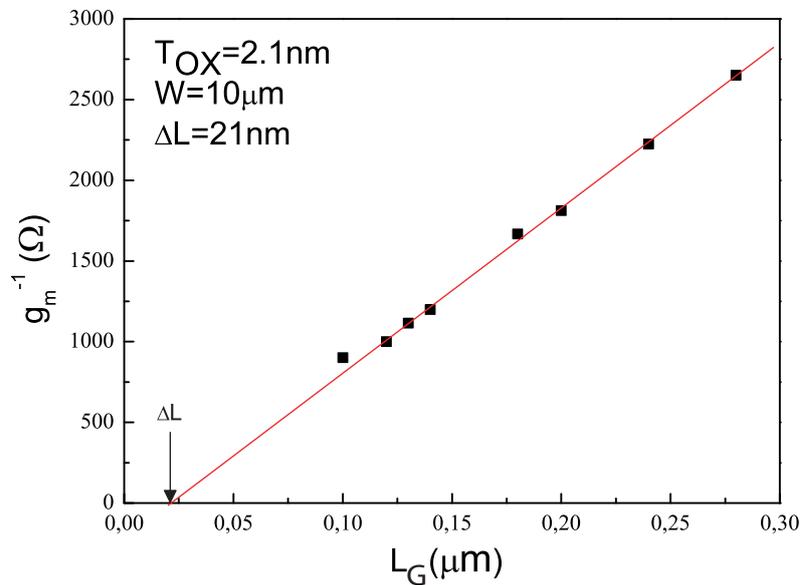


FIG. I.32 – Détermination de la longueur effective du canal.

Par cette méthode, pour des transistors avec une épaisseur d'oxyde $T_{ox} = 2.1nm$, un dopage $N_A = 4.10^{17}cm^{-3}$ et $W_G = 10\mu m$, on obtient $\Delta L_G \approx 21nm$.

b) Extraction de ΔW_G

La caractérisation de la largeur effective suit la même stratégie : à partir d'un lot de transistors ayant la longueur fixée et la largeur variable. Considérant :

$$G = \frac{\mu_0 C_{OX} V_{DS}}{L_{eff}} (W_G - \Delta W_G) \quad (I.108)$$

on trace G_{max} suivant W_G et l'abscisse à l'origine donne ΔW_G .

I.4.5 Détermination de la valeur du pic de champ latéral

Le pic de champ est relié à l'évolution du courant substrat. En effet, la valeur élevée de $\xi(x)$ autour de ξ_{sat} permet aux porteurs d'acquérir une énergie cinétique suffisante pour subir des chocs inélastiques avec le réseau cristallin. Le résultat de ces chocs est la formation de paires électron-trou : c'est le phénomène d'ionisation par impact. La majorité des paires ainsi générées est collectée au substrat.

L'équation (I.86) de E_m est donnée pour une tension très proche de la tension de saturation mais pour $V_{DS} > V_{Dsat}$, ξ_{sat}^2 devient négligeable et on peut écrire :

$$E_m = \frac{V_{DS} - V_{Dsat}}{l} \quad (\text{I.109})$$

l est donnée par [32] :

$$L_{eff} < 0.5\mu m \Rightarrow l = 0.017 T_{ox}^{1/8} L_{eff}^{1/5} X_J^{1/3} \quad (\text{I.110})$$

$$L_{eff} > 0.5\mu m \Rightarrow l = 0.22 T_{ox}^{1/3} X_J^{1/2} \quad (\text{I.111})$$

Le courant peut être relié à E_m en intégrant le coefficient d'ionisation par impact, proportionnel à $\exp^{-\beta/\xi(x)}$, entre x_{sat} et l_{eff} [28] :

$$I_B = C I_{DS} \exp\left(\frac{-\phi_{i,h/e}}{q\lambda_{h/e} E_m}\right) \quad (\text{I.112})$$

Dans ces expressions, $\phi_{i,h/e}$ est l'énergie d'ionisation par impact des trous/électrons et $\lambda_{h/e}$ le libre parcours moyen des trous/électrons (c'est à dire la distance parcourue sans subir de collisions). On trouve dans la littérature $\lambda_h = 55\text{\AA}$ et $\lambda_e = 78\text{\AA}$. Le pic de champ latéral est directement exprimé en fonction du taux d'ionisation par impact. Ce taux d'ionisation est caractérisé par les coefficients d'ionisation α_i et β_i . Ces coefficients ont été reliés aux courants de Drain et de Substrat [33] :

$$\frac{I_{BS}}{I_{DS}} = (V_{DS} - V_{Dsat}) \frac{\alpha_i}{\beta_i} \exp\left(\frac{\beta_i l_0}{V_{DS} - V_{Dsat}}\right) \quad (\text{I.113})$$

Typiquement $\alpha_i = 2.5 \cdot 10^6 \text{ cm}^{-1}$ et $\beta_i = 2 \cdot 10^6 \text{ V/cm}$ [30]. L'ionisation est maximale pour le maximum de courant substrat. Pour une tension de Drain donnée, la valeur de V_{GS} telle que le courant substrat est maximal doit donc être connue. C'est autour de cette valeur que l'on trace $I_{BS}/(I_{DS}(V_{DS} - V_{Dsat}))$ sur une échelle logarithmique suivant $(V_{DS} - V_{Dsat})^{-1}$ pour les différentes valeurs de V_{GS} .

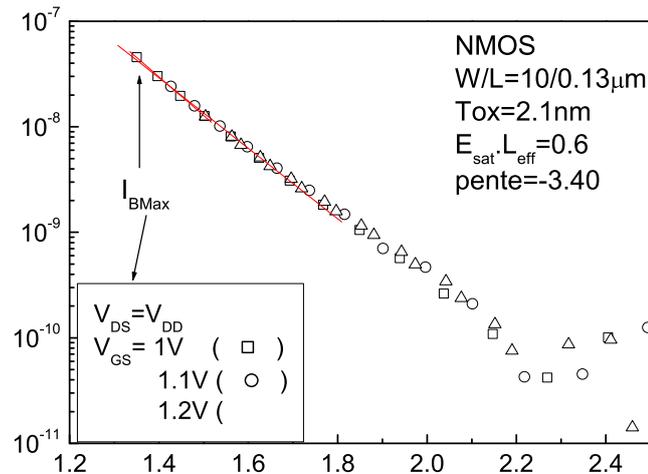


FIG. I.33 – Détermination du pic de champ latéral.

La valeur de $\xi_{sat}L_{eff}$ dans l'expression (I.92) de la tension de saturation doit être ajustée pour obtenir les mêmes caractéristiques satisfaisant à l'indépendance en V_{GS} (Fig. I.33). La valeur maximale du champ électrique latéral s'obtient à l'aide de la pente du graphe :

$$pente = \frac{-\phi_{i,e}}{q\lambda_e E_m} \quad (\text{I.114})$$

où $\phi_{i,e}$ est l'énergie minimale nécessaire à un électron pour provoquer l'ionisation par impact, et vaut $1.3eV$ [33]. Dans l'exemple de la figure I.86 on trouve ainsi une valeur du pic de champ de $206kV/cm$. Avec la même méthode on trouve pour le PMOS $119kV/cm$

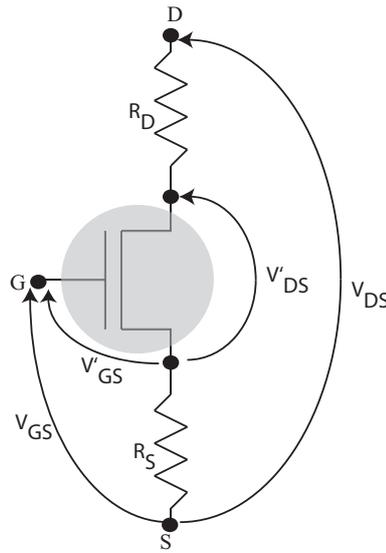
	$ E_{max} (kV/cm)$	$\phi_{i,e/h}(eV)$	pente
NMOS	210	1.3	-3.4

TAB. I.2 – Valeurs maximales du champ latéral pour un transistor NMOS de dimensions de Grille $W/L = 10/0.13\mu m$ et d'épaisseur d'oxyde $T_{ox} = 2.1nm$.

I.4.6 Calcul de la résistance série

Nous avons négligé jusqu'à présent les résistances liées aux contacts S/D. Ces résistances dites d'accès trouvent leurs origines au niveau :

1. des différentes métallisations et siliciures ($TiSi_2$, $CoSi_2$),
2. des connexions des Source et Drain,
3. de l'étalement des jonctions n^+/p (p^+/n),


 FIG. I.34 – Schéma équivalent de la loi des nœuds pour le calcul de R_{SD} .

4. de la résistivité des différentes zones, liée au dopage, à la présence de pockets ou halos. L'ensemble de ces contributions, fortement dépendantes de la complexité de la technologie étudiée, sont rassemblées sous le terme R_{SD} . La loi d'Ohm donne simplement $R_m = V_{DS}/I_{DS}$, où R_m est la résistance mesurée. Dans le cas d'un transistor non dégradé, la symétrie de la structure impose $R_S = R_D$. Pour tenir compte des résistances d'accès, on peut rajouter R_S et R_D en série avec le transistor (figure I.34) et on a :

$$V'_{GS} = V_{GS} - I_{DS}R_S \quad (\text{I.115})$$

$$V'_{DS} = V_{DS} - I_{DS}(R_S + R_D) \quad (\text{I.116})$$

On s'intéresse à l'équation du courant en régime linéaire, où la résistance est la plus forte, avec une réduction de mobilité à un seul coefficient θ :

$$I_{DS} = \frac{\mu_0 C_{OX} W_G}{L_G} \frac{V'_{DS}(V'_{GS} - V_{Th} - \frac{\alpha V'_{DS}}{2})}{1 + \theta(V'_{GS} - V_{Th})} \quad (\text{I.117})$$

Pour alléger l'écriture on utilise les notations

$$V_{GT} = V_{GS} - V_{Th} \quad (\text{I.118})$$

$$k_0 = \mu_0 C_{OX} W_G / L_G \quad (\text{I.119})$$

$$k = k_0 / (1 + \theta(V_{GT})) \quad (\text{I.120})$$

et on prend $\alpha \approx 1$. En incorporant (I.115) et (I.116) dans (I.117) et négligeant les termes de second ordre en I_{DS} on peut écrire :

$$I_{DS} = k \left(V_{GT} - I_{DS}R_S - \frac{V_{DS} - I_{DS}R_{SD}}{2} \right) (V_{DS} - I_{DS}R_{SD}) \quad (\text{I.121})$$

Pour V_{DS} faible (régime linéaire) et V_{GS} important, on a

$$V_{GT} \gg I_{DS}R_S - \frac{V_{DS} - I_{DS}R_{SD}}{2} \quad (\text{I.122})$$

et en négligeant les termes du second ordre en I_{DS} , on arrive à :

$$I_{DS} = \frac{k_0 V_{GT} V_{DS}}{1 + (\theta + k_0 R_{SD}) V_{GT}} \quad (\text{I.123})$$

On calcule alors la résistance totale :

$$R_m = R_{canal} + R_{SD} = \frac{V_{DS}}{I_{DS}} \quad (\text{I.124})$$

$$= \frac{1 + (\theta + k_0 R_{SD}) V_{GT}}{k_0 V_{GT}} \quad (\text{I.125})$$

Soit finalement :

$$R_m = \frac{1}{k_0 (V_{GS} - V_{Th})} + \frac{\theta}{k_0} + R_{SD} \quad (\text{I.126})$$

Comme la résistance totale entre les nœuds S et D est la mise en série de R_S , R_D et de la résistance du canal R_{canal} , on en déduit la résistance de la couche d'inversion :

$$R_{canal} = \frac{1}{k_0 (V_{GS} - V_{Th})} \quad (\text{I.127})$$

Pour procéder à l'extraction de la résistance des contacts, on doit tracer la résistance mesurée $R_m = V_{DS}/I_{DS}$ en fonction de $(V_{GS} - V_{Th})^{-1}$. On obtient un tracé quasi-linéaire dont l'ordonnée à l'origine vaut $\theta/k_0 + R_{SD}$. Sur la Fig. I.35 est tracée la résistance mesurée en fonction de $(V_{GS} - V_{Th})^{-1}$. L'intersection à l'ordonnée donne $\theta/k_0 + R_{SD} = 22\Omega$. Avec $\theta_N = 0.3V^{-1}$ et $k_0 = 0.038V^{-1}$ on obtient $R_{SD} = 14.1\Omega$. Le corollaire de cette méthode est le tracé de R_m comme fonction de la longueur de Grille dessinée L , pour $V_{GS} - V_{Th}$ fixé. Pour un ensemble de valeurs de $V_{GS} - V_{Th}$, on obtient un faisceau de droites dont l'intersection a pour coordonnées $X = \Delta L$ et $Y = R_{SD}$. Cette façon de procéder vient de (I.126). On obtient ainsi sur la Fig. I.36, avec le même dispositif que pour I.35, $R_m \approx 22\Omega$ et $\Delta L \approx 22nm$, en accord avec les résultats précédents.

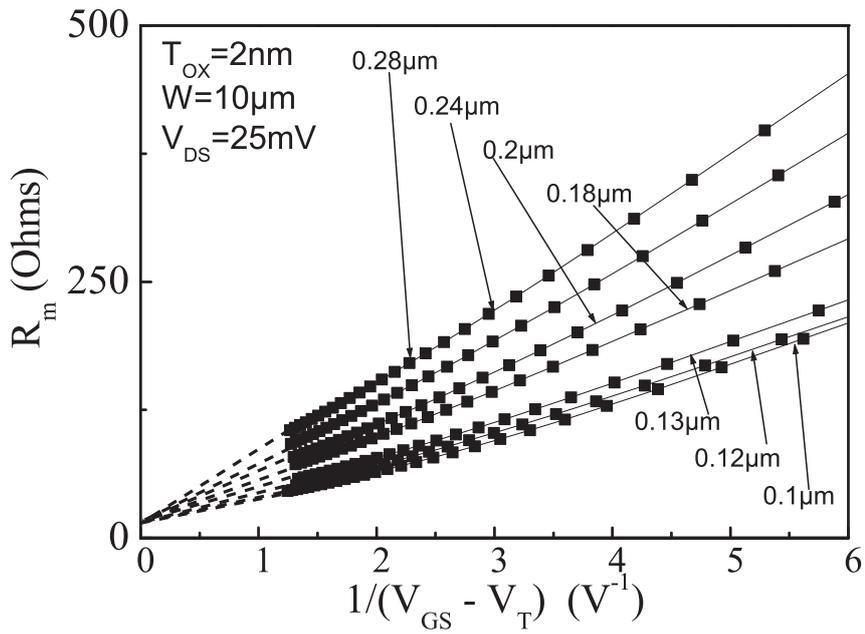


FIG. I.35 – Extraction de la résistance des contacts d'accès pour des longueurs de Grille comprises entre $0.28\mu\text{m}$ et $0.1\mu\text{m}$.

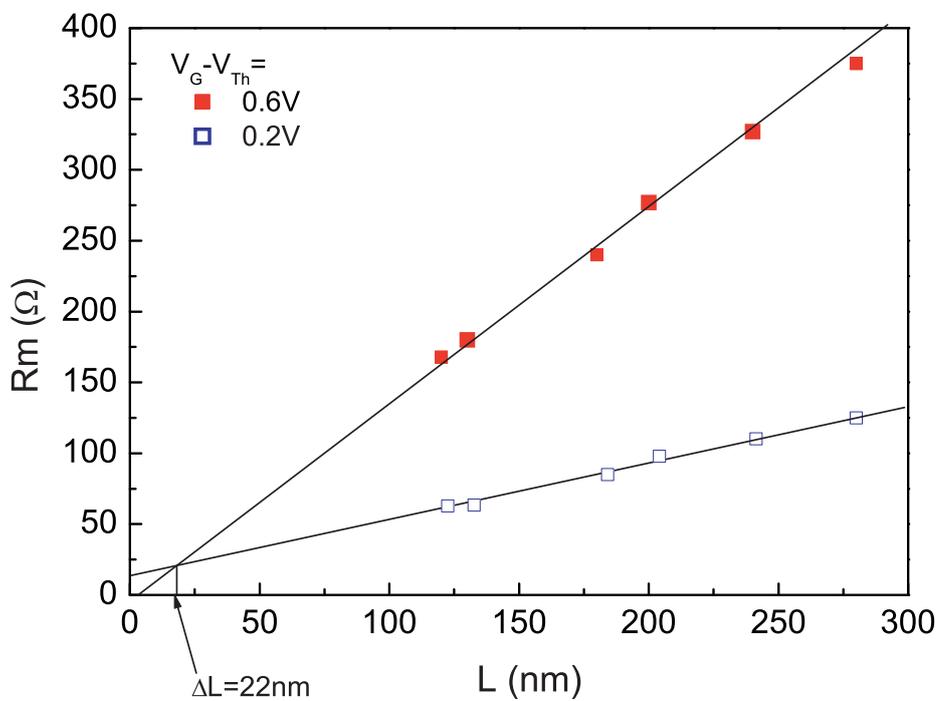


FIG. I.36 – Extraction de la résistance des contacts d'accès et de ΔL pour $V_{GS} - V_{Th} = 0.6\text{V}$ et 0.2V .

Conclusion

Dans ce premier chapitre, nous avons établi les équations de base qui permettent de décrire le fonctionnement des structures MOSFETs, avec et sans défauts. Ces équations servent à développer la modélisation du comportement électrique des transistors MOSFETS (section 2). Celles-ci constituent les premiers niveaux de modélisation SPICE, et ne prennent pas en compte les effets de la réduction d'échelle. La troisième partie de ce premier chapitre a donc exposé les effets parasites les plus courants. Quelle que soit la complexité d'un modèle, il doit être validé par les mesures électriques réalisées sur le composant étudié. La quatrième partie s'est consacrée aux méthodes d'extractions des principaux paramètres de fonctionnement du transistor. Enfin, par quelques exemples nous avons montré l'importance de l'optimisation des modèles, imposée par les évolutions des technologies. L'ensemble de ces modèles et techniques de caractérisations vont nous permettre d'analyser l'influence du vieillissement des transistors MOSFETs, sur leur performances de fonctionnement statiques.

Références du Chapitre I

- [1] S. Sze, *Physics of Semiconductor Devices, 2nd edition*, 2nd ed. New-York: Wiley and Sons, 1988. [I.1.1](#), [I.1.2](#), [a](#))
- [2] D. K. Schroder, “The concept of generation and recombinaison lifetime in semiconductors,” *IEEE Trans. Electron Devices*, pp. 1336–1338, Aout 1982. [b](#))
- [3] D. Vuillaume, D. Goguenheim, and G. Vincent, “New insights on the electronic properties of trivalent silicon defects at oxidized $\langle 100 \rangle$ silicon surface,” *Appl. Phys. Lett.*, no. 57, pp. 1206–1208, 1990. [I.1.4](#)
- [4] D. Vuillaume, D. Goguenheim, and J. C. Bourgoin, “Nature of the defects generated by electric field stress at the Si-SiO₂ interface,” *Appl. Phys. Lett.*, vol. 58, p. 490, 1991. [I.1.4](#)
- [5] D. K. Shroder, *Semiconductor material and device characterization*. Wiley interscience, John Wiley & Sons, Inc, 1990. [I.1.4](#)
- [6] W. Shockley and W. Read, “Statistics of the recombinaison of holes and electrons,” *Phys. Rev.*, vol. 87, no. 5, pp. 62–69, 1952. [a](#))
- [7] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. D. Keersmaecker, “A reliable approach to charge-pumping measurements in MOS transistors,” *IEEE Electron Device Lett.*, vol. ED-31, no. 1, pp. 42–53, 1984. [a](#)), [II.2](#), [c](#)), [c](#)), [c](#)), [II.12](#), [c](#))
- [8] J. L. Autran, F. Djahli, B. Balland, C. Plossu, and L. Gaborieau, “Three-level charge pumping on submicronic transistors,” *IEEE Electron Device Lett.*, vol. 84, no. 6, pp. 607–611, 1992. [a](#))
- [9] N. S. Saks and M. G. Ancona, “Determination of interface trap capture cross sections using three-level charge pumping,” *IEEE Electron Device Lett.*, vol. 11, no. 8, Août 1990. [a](#))
- [10] D. Goguenheim, D. Vuillaume, G. Vincent, and N. M. Johnson, “Accurate measurement of capture cross section of semiconductor insulator interface sates by a trap-filling experiment: the charge-potential feedback effect,” *Journal of applied physic*, no. 68(3), pp. 1104–1113, Août 1990. [a](#))
- [11] E. H. Snow, A. S. Grove, B. D. Deal, and C. T. Sah, “Ion transport phenomena in insulated films,” *Journal of Applied Physic*, no. 36, pp. 1664–1673, Mai 1976. [b](#))

- [12] N. J. Chou, “Application of triangular voltage sweep to mobile charges studies in mos structures,” *J Electrochen. Soc.*, no. 119, pp. 601–609, Avril 1971. **b)**
- [13] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, “Characteristics of the surface charge (Q_{ss}) of thermally oxydized silicon,” *J Electrochen. Soc.*, no. 114, pp. 266–274, Mars 1976. **c)**
- [14] M. Fadlallah, C. Petit, A. Meinertzhagen, G. Ghibaudo, O. S. M. Bidaud, and F. Guyader, “Influence of nitridation in ultra-thin oxide on the gate current degradation of N and PMOS devices,” *Microelectronics Reliability*, no. 43, pp. 1433–1438, 2003. **c)**
- [15] M. Barron, “Low level currents in insulated gate field effect transistors,” *Solid-State Electronics*, vol. 15, pp. 293–302, 1972. **a), a)**
- [16] W. Shockley, “Problems related to p-n junctions in silicon,” *Solid States Electronics*, vol. 2, no. 1, pp. 35–67, 1961. **a)**
- [17] Y. Tsididis, *Operation and modeling of The MOS Transistor*, 2nd ed. Mc Graw Hill, 1999. **a)**
- [18] T. N. Nguyen and J. D. Plummer, “Physical mechanisms responsible for short channel effect in mos devices,” in *IEEE, IEDM Technical digest*, Washington DC, 1981. **a)**
- [19] C. S. Rafferty, H.-H. Vuong, S. A. Eshraghi, M. D. Giles, M. R. Pinto, and S. J. Hillenius, “Explanation of reverse short channel effect by defect gradients,” in *IEDM technological digest*, pp. 311–314, 1993. **b)**
- [20] G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2nd ed. Mc Graw Hill, 1988. **I.3.2**
- [21] N. D. Arora, J. R. Hauser, and D. J. Roulston, “Electron and holes mobility in silicon as a function of concentration and temperature,” *IEEE Trans. Electron Devices*, vol. Ed-29, pp. 609–616, Février 1982. **I.3.3**
- [22] C. Kittel, *Introduction to solid states physics*, 6th ed. John Wiley & Sons, 1986. **a)**
- [23] P. P. Wang, “Device characteristics of short-channel and narrow-width MOSFET’s,” *IEEE Trans. Electron Devices*, vol. Ed-25, pp. 779–768, july 1978. **d)**
- [24] F. M. Klaassen, “A MOS model for computer-aided design,” *Philips Research report*, vol. 31, pp. 71–83, 1976. **d)**
- [25] F. Klaassen, *MOS device modelling*. Englewoods Cliffs, N.J.: Design of VLSI circuits for telecommunications, Y. Tsididis and P. Antognetti, Printice-Hall (editeurs), 1985. **d)**
- [26] P. K. McLarty, S. Cristoloveanu, V. M. O. Faynot, J. R. Hauser, and J. J. Wortman, “A simple parameter extraction method for ultra-thin oxide MOSFET’s,” *Solid State Electronics*, vol. 38, no. 6, pp. 1175–1177, 1995. **d)**
- [27] D. Foty, *MOSFET modeling with SPICE, Principles and practices*. Upper Saddle River: Printice Hall, 1997. **d), f)**
- [28] T. Y. Chan, P. K. Ko, and C. Hu, “Dependence of channel electric field on device scaling,” *IEEE Electron Device Lett.*, vol. ED-6, no. 10, pp. 551–553, 1985. **e), I.4.5**

-
- [29] M. C. P. H. Wong, "Approximation of length of velocity saturation region in MOSFET's," *IEEE Trans. Electron Devices*, vol. 44, no. 11, pp. 2033–2036, 1997. [e](#))
- [30] X. Z. K. Y. Lim, "An analytical effective channel-length modulation model for overshoot in submicron MOSFETs based on energy-balance formulation," *Microelectronic reliability*, no. 42, pp. 1857–1864, 2002. [I.3.4](#)
- [31] Y. Cheng, M. Chan, K. Hui, M. C. Jeng, Z. Liu, and J. H. et Al., *BSIM3v3 Manual*. CA: UC Berkeley, 1996. [I.3.4](#)
- [32] Y. A. E. Mansy and A. R. Boothroyd, "A simple two-dimensional model for IGFET operation in the saturation region," *IEEE Trans. Electron Devices*, vol. ED-24, no. 11, p. 224, 1997. [I.4.5](#)
- [33] C. Hu, S. C. Tam, F. C. Hsu, P. K. Ko, T. Y. Chan, and K. W. Terril, "Hot electron induced MOSFET degradation - model, monitor and improvement," *IEEE Trans. Electron Devices*, vol. ED-32, no. 2, fevrier 1985. [I.4.5](#), [I.4.5](#)
- [34] J. Chung, M.-C. Jeng, G. May, P. K. KO., and C. Hu, "Hot electrons in deep-submicrometer MOSFETs," in *IEDM / IEEE*, 1988. [I.4.5](#), [II.2.4](#)

CHAPITRE II

Techniques d'analyses de la dégradation du transistor MOSFET

Introduction

Les réductions d'échelles permanentes des technologies CMOS avancées, impliquent nécessairement des variations de procédés de fabrication importantes d'une génération à l'autre. Cette évolution se retrouve dans le type de dégradation généré vis-à-vis des conditions de fonctionnement du dispositif. Une attention particulière doit être portée au choix des moyens d'analyses mis en œuvre en fonction du type de défaut étudié. Afin de rendre pertinente l'étude de chaque technique, nous allons introduire ici la notion de localisation de la dégradation. Nous pouvons distinguer deux cas:

1. Le canal est dégradé de manière uniforme: les dégradations de l'interface et/ou dans l'oxyde sont répartis uniformément le long du canal, conséquence d'injections uniformes de porteurs, réalisée pour une tension de Grille importante, Drain et Source étant au même potentiel.
2. Les dégradations sont réparties dans une zone particulière du canal (en général près du Drain), à la suite d'injections localisées de porteurs chauds (HC). D'une manière générale, la génération de ces porteurs énergétiques, se produit lorsqu'on applique une tension de Drain élevée.

Ceci est illustré sur le schéma de la Fig. II.1: les dégradations peuvent être localisées dans trois zones distinctes (1: dans le canal, 2: dans la zone de recouvrement, 3: sous les espaceurs (*spacers*) d'oxyde) correspondant chacune à un mécanisme de dégradation spécifique. Ces mécanismes physiques seront décrits dans le prochain chapitre. Ici seront exposées les méthodes permettant de qualifier et localiser ces dommages dans la structure MOSFET. Nous allons considérer deux approches. En premier lieu, l'étude des caractéristiques $I - V$ est un bon indicateur du vieillissement d'un dispositif MOSFET et permet de donner une première indication sur la localisation de la dégradation. Cependant, suivant la nature des défauts et leur extension dans

le canal, il est difficile de distinguer leur rôle respectif . C'est pourquoi, dans un second temps, des mesures par pompage de charges (CP) autorisent une étude plus quantitative des défauts présents à l'interface.

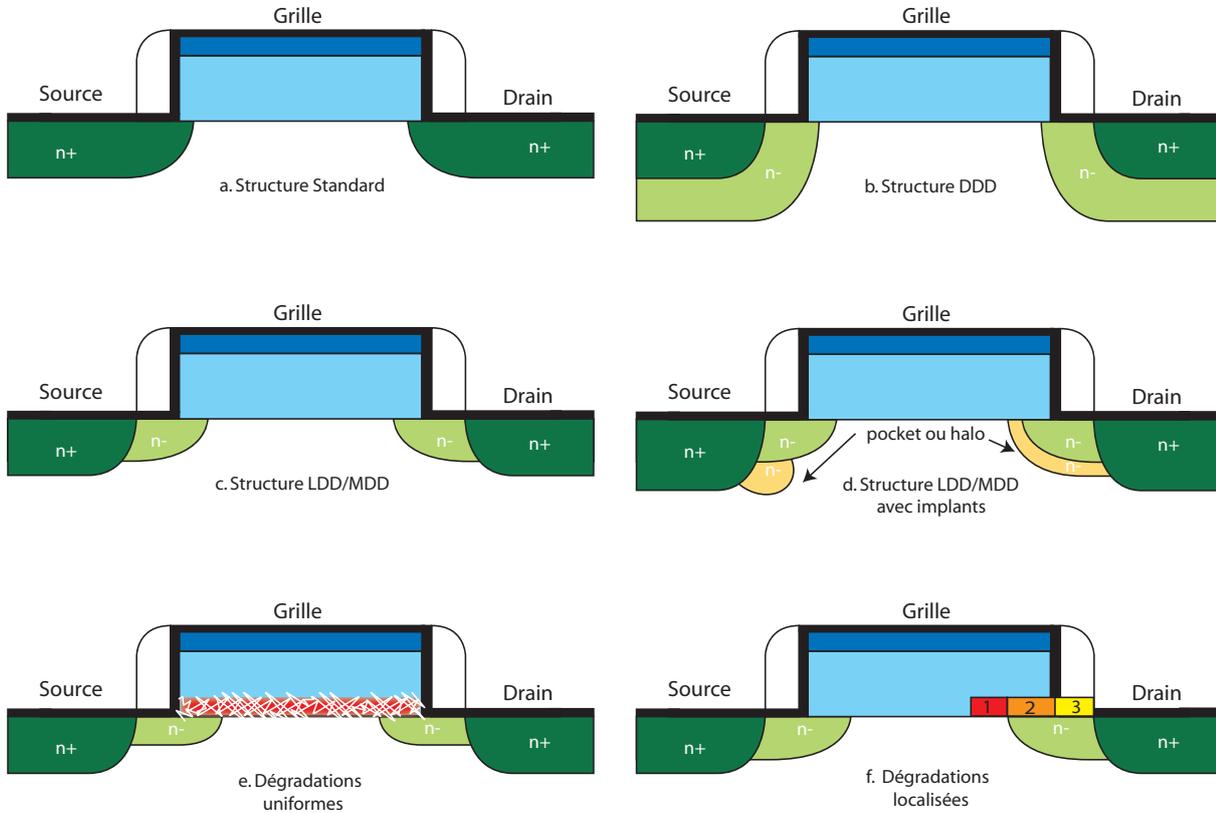


FIG. II.1 – a, b, c, d : représentation des différentes structures de Drain proposées pour réduire l'effet de l'injection des porteurs chauds, e, f : répartition des défauts suivant le mécanisme responsable des dommages.

Dans les années 1980, afin de continuer à augmenter le courant dans les structures MOS [1], les fondeurs ont dû modifier la structure conventionnelle (une seule diffusion pour la Source et le Drain), en dopant plus légèrement les zones de Drain et de Source directement en contact avec le canal, en créant ainsi des zones de diffusion dite *Lightly Doped Drain* (LDD). Ceci a pour conséquence de réduire le champ électrique latéral maximal. Sur la base du LDD a été développée l'implantation des dopants avec un large angle pour jouer sur le profil de dopage : *Large Angle Tilt Implanted Drain* (LATID) [2]. Par la suite, le *Double Dopage de Drain* (DDD) [3] a été introduit pour limiter les injections de porteurs du canal. Le DDD consiste à faire un implant plus profondément dans le substrat, en plus des zones de diffusions standards. Enfin est apparue la technologie MDD ($LG < 0.25\mu m$). La réduction de la tension d'alimentation ($V_{DD} \leq 2.5V$) a été équilibrée par l'augmentation du dopage des zones LDD pour arriver à des zones de dopage intermédiaire dites *Medium Doped Drain* (MDD) avec éventuellement, des "halo" [4] ou des "pocket" [5] qui sont des implants additionnels localisés plus profondément dans le substrat ou jusqu'à l'interface Si-SiO₂, dans le but de limiter les effets SCE et de perçage

(*punchthrough*).

Les évolutions successives de la structure de Drain sont devenues un enjeu primordial pour la fiabilité des technologies CMOS, parce qu'elle influence directement le profil du champ électrique dans le canal, clé des mécanismes d'injections de porteurs chauds [6]. Les effets nocifs de l'injection de porteurs chauds, que l'optimisation des structures de Drain [7] cherche à réduire, vont pouvoir être caractérisés :

- dans la région (1), par les techniques IV et CP préférentiellement sensibles aux dégradations dans le canal actif,
- dans la région (2), au-dessus des zones de recouvrement des diffusions secondaires (LDD-MDD), par les mêmes techniques sous certaines conditions,
- dans la région (3), sous les espaceurs, pouvant être examinées par des techniques spécifiques comme la Grille Flottante (FG).

II.1 Analyse par mesures courant-tension I-V

Elle consiste en l'étude de la dérive des principaux paramètres de fonctionnement du transistor. Parmi les bons indicateurs de vieillissement de la structure, on trouve :

- la variation relative du courant de drain pour une polarisation fixée, $\frac{\Delta I_{DS}}{I_{DS0}}$
- la variation de la tension de seuil ΔV_{Th}
- l'évolution du maximum de la transconductance G_m et sa variation relative $\frac{\Delta G_m}{G_{m0}}$ pour une tension de Drain donnée,
- la variation de la pente sous seuil ΔS ,
- ainsi que les variations des paramètres statiques tels que R_{SD}, I_{on}, I_{off} .

Toutes donnent des informations sur la présence et la nature des défauts, et sous certaines conditions leur localisation.

II.1.1 Suivi de la variation du courant de Drain au cours du temps

C'est la technique qui permet de "voir" le plus simplement et directement le niveau de dégradation de la structure dans le temps. En effet en suivant les variations relatives ($\Delta I_D / I_{D0}$) du courant de Drain pour diverses polarisations, on peut se faire une idée qualitative du vieillissement. Toutefois, il n'est pas possible d'appliquer la contrainte et de mesurer simultanément les caractéristiques $I_D - V_G$. On procède donc simplement à une alternance de stress et de mesure. Etant donné que les période de stress s'étalent sur plusieurs milliers de seconde, il est plus commode de travailler sur les échelles de temps logarithmiques, on peut ainsi se fixer des intervalles de mesures fixes dans cette échelle : 2,3,5 ou encore 8 fois par décades. On peut observer sur la Fig. II.2 l'évolution typique du courant de drain mesuré dans cet exemple pour $V_D = 25mV$ et $V_G = V_{DD}/2$ en mode direct en fonction du temps de stress. Cette évolution est linéaire dans un graphe log-log, et la pente des courbes est significative du facteur d'accélération du vieillissement du dispositif. Dans cette partie nous n'entrerons pas plus dans la description des mécanismes de dégradations, étudiés dans le chapitre suivant.

a) Régime linéaire

La dépendance de la dégradation du courant de Drain en mode linéaire en fonction de la tension de mesure V_G est une première façon de déterminer l'étalement de la région dégradée. En effet, pendant la mesure la résistance totale est la somme de la résistance du canal R_{ch} et de la résistance série R_{SD} (I.4.6):

$$R_m = R_{canal} + R_{SD} = \frac{1}{k_0(V_{GS} - V_{Th})} + \frac{\theta}{k_0} + R_{SD} \quad (II.1)$$

On peut observer deux comportements :

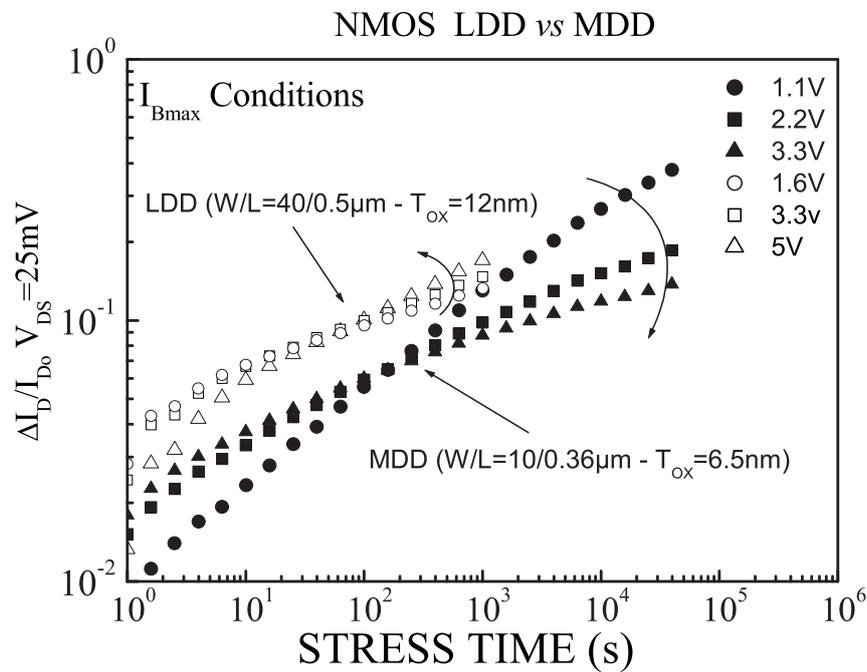


FIG. II.2 – Variation relative du courant de Drain au court d'une contrainte DC au maximum du courant substrat pour $V_{DS} = 4.5V$ (MDD) et $7V$ (LDD) appliquée à des transistors de type N de dimension $W/L = 10/0.36\mu m$, $T_{ox} = 6.5nm$ (MDD) et $W/L = 40/0.5\mu m$, $T_{ox} = 12nm$ (LDD).

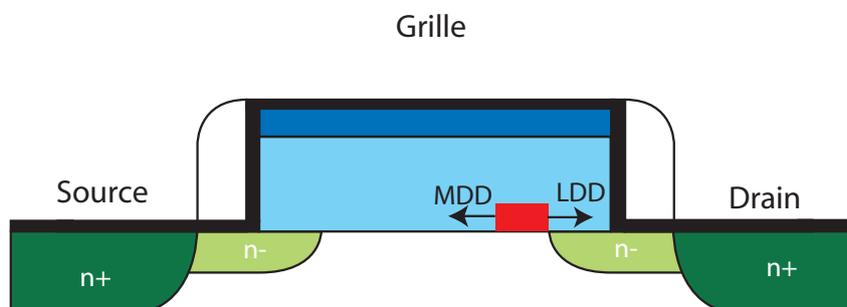


FIG. II.3 – Positionnement typique de la dégradation suivant le type de structure.

1. si les régions d'implant $n-$ sont faiblement dopées (LDD), l'influence de la résistance série à fort V_{GS} augmente, puisque R_{ch} diminue et que le point où le champ latéral est maximum se déplace vers le Drain (zones 2 et 3). Par conséquent la dégradation du courant linéaire devient plus grande avec le V_G de mesure [36, 9].
2. Si la zone $n-$ est plus fortement dopée (MDD), le point du champ latéral maximum se déplace vers le canal actif et la région dégradée s'étend vers la source (zone 1). Par conséquent l'influence des dommages devient plus grande aux faibles valeurs de V_G alors que l'impact des dommages sur la résistance de canal (assimilable à une réduction de mobilité des porteurs) devient plus importante que pour R_{SD} [9].

Les figures II.2 et II.3 illustrent ce contraste : la réduction du courant de Drain est plus faible aux petites valeurs de la tension de Grille, ce phénomène est nettement plus marqué pour la structure MDD.

b) Régime saturé

Quand des caractérisations $I_D - V_G$ en régime saturé sont mesurées en mode direct (FWD), la Source et le Drain jouent leur rôle habituel, les porteurs minoritaires vont de la Source vers le Drain. En mode inverse (REV) S et D sont échangés, les porteurs vont dans le sens opposé (de D à S). Si la caractérisation FWD est employée, une partie des dommages situés sous la ZCE sont écartés et n'affectent pas le courant et la dégradation apparaît donc amoindrie [36]. En effet, pour le régime de saturation - $V_{DS} \geq V_{Dsat}$ - en mode FWD, nous avons la ZCE qui s'étend du Drain vers le canal avec V_{DS} . En mode REV la zone de charge d'espace s'étend de la Source vers le canal avec V_{SD} . Si les mesures sont effectuées en REV, la région dégradée est située à la source et reste ainsi totalement "visible" sur les courbes $I_D - V_G$. On compare en général ces courbes avec le mode direct pour s'assurer de la localisation des défauts : $I_{DS} - V_{DS}$ par rapport $I_{SD} - V_{SD}$ avec V_{GS} comme variable.

En modulant la tension de Drain, on peut essayer de faire tendre les courbes après contrainte vers les courbes de précontrainte et ainsi calculer la longueur de la région dégradée. Considérons l'effet CLM (I.3.4):

$$I_{DS} = \frac{I_{DS}^*}{1 - \frac{\Delta L}{L_{eff}}} \quad (\text{II.2})$$

où I_{DS}^* est le courant en l'absence d'effet CLM. Après stress, lorsque le point de pincement est situé au delà de la zone dégradée dans le canal, les défauts n'affectent pas ou peu le courant de Drain. On augmente la tension de Drain V_{DS} jusqu'à ce que le courant de Drain après contrainte se superpose à I_{DS} avant contrainte. On a alors ΔL qui est égale à la longueur de la zone dégradée. La précision de cette méthode réside la connaissance du profil de dopage le long du canal à l'approche des Source et Drain.

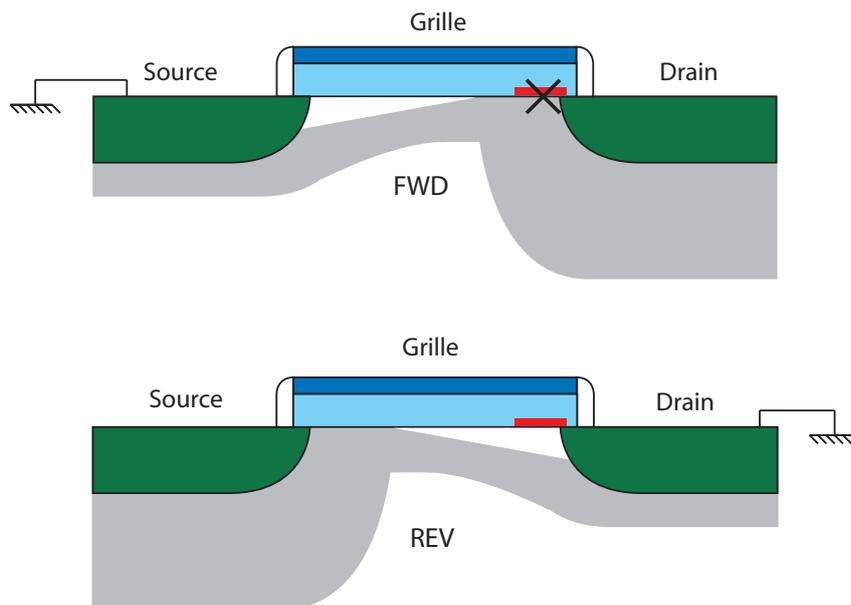


FIG. II.4 – Illustration du phénomène d'écrantage pour la caractérisation en régime saturé en présence de dégradations localisées.

II.1.2 Variation de la tension de seuil

La tension de seuil est un paramètre essentiel pour le bon fonctionnement des circuits digitaux. Comme nous l'avons vu dans le premier chapitre, elle dépend de la longueur du canal au sein d'une technologie, et permet l'extraction d'autres paramètres (dimensions effectives, résistances séries, facteur de réduction de mobilité). Rappelons les équations établies dans le premier chapitre (I.4.6):

$$V_{Th} = V_{FB} + 2\phi_F + \gamma(-V_{BS} + 2\phi_F)^{1/2} - \frac{Q_{it0}}{C_{OX}} \quad (\text{II.3})$$

$$V_{FB} = \phi_{MS} - \frac{Q_{ot}}{C_{OX}} \quad (\text{II.4})$$

avec $\gamma \approx \sqrt{2\epsilon_{si}qN_B}/C_{ox}$ et $\phi_F = \frac{kT}{q} \ln\left(\frac{N_B}{n_i}\right)$, les variations de la tension de seuil sont alors décrites par :

$$\Delta V_{Th} = -\frac{\Delta Q_{it}(2\phi_f) + \Delta Q_{ot}}{C_{ox}} \quad (\text{II.5})$$

avec ΔQ_{ot} et ΔQ_{it} sont les variations de la charge dans l'oxyde et à l'interface, données par :

$$\Delta Q_{ot} = \pm q\Delta N_{ot} \quad (\text{II.6})$$

$$\Delta Q_{it} = \pm q\Delta D_{it}(\psi_S - \phi_B) = \pm q\Delta D_{it}\cdot\phi_F \quad (\text{II.7})$$

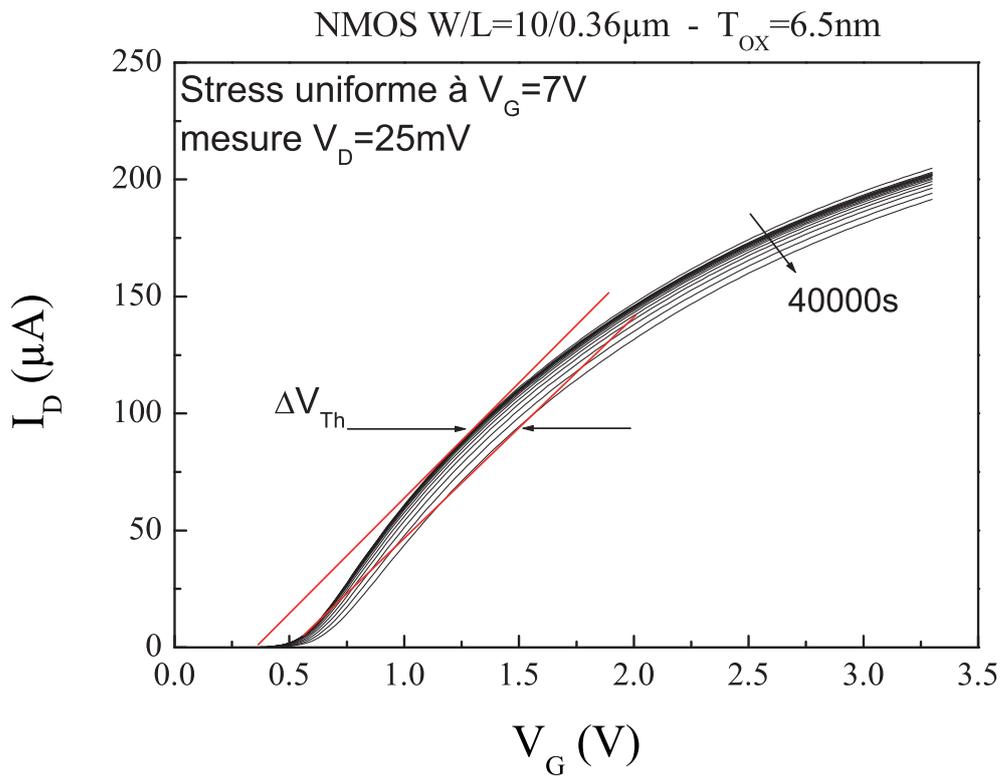


FIG. II.5 – Décalage de la tension de seuil pendant un stress uniforme à tension constante.

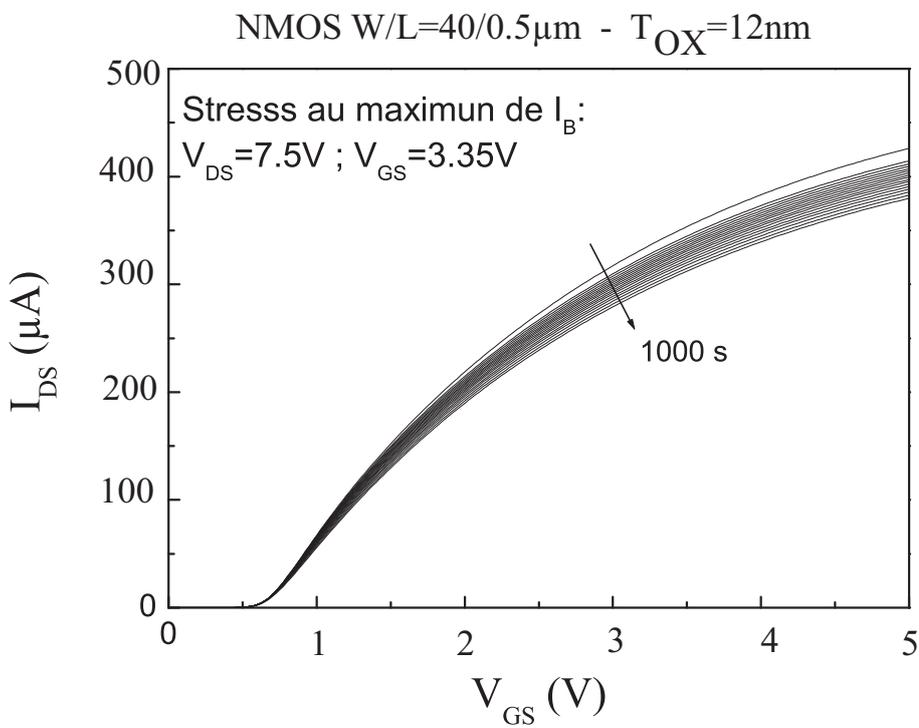


FIG. II.6 – Augmentation de la tension de seuil au cours d'un stress au maximum du courant substrat ($V_{GS} = 3.35V, V_{DS} = 7.5V$) de 1000s sur un NMOS.

Une variation de V_{Th} est associée à la présence de charges fixes dans l'oxyde et/ou à une charge piégée sur les états d'interfaces. Pour le suivi de la tension de seuil durant un vieillissement, on procède à l'extraction de la tension de seuil à un niveau de courant de Drain fixé (V_{Th} est extrait par la technique de la tangente prise à point I_D fixé et non plus au max de G_m).

Dans le cas où une contrainte uniforme est appliquée, nous observons un décalage parallèle des courbes linéaires, ce qui permet d'obtenir la variation de la charge négative piégée dans l'oxyde. Pour un NMOS (PMOS), une augmentation (diminution) de $|V_{Th}|$ traduit la présence d'une charge négative dans l'isolant, alors qu'une diminution (augmentation) de la tension de seuil révèle une charge positive.

Dans le cas de stress non uniformes, on assiste à un décalage non parallèle des courbes $I_D - V_G$ linéaires, ce qui se traduit par des dommages localisés associés à la présence d'états d'interface. Cependant, la distinction claire entre ces deux types de défauts exige des techniques complémentaires comme des mesures CV et de pompage de charges CP (technique exposée plus loin) afin de déterminer leur influence respective.

Dans le cas d'une dégradation localisée, le problème de la distinction devient plus complexe du fait que la mesure de V_{Th} devient beaucoup moins affectée par les dommages qui sont fortement localisés au drain (structure LDD).

On obtient typiquement les courbes des Fig. II.6 et II.5. Dans la première on observe sur transistor NMOS dégradé sous contrainte DC au maximum de courant substrat, un décalage des courbes non parallèle, ce qui indique la prédominance de la génération d'états d'interface localisés. La seconde figure présente un décalage parallèle au cours d'un stress uniforme effectué à la tension $V_G = 2.5V$.

II.1.3 Variation de la transconductance

Le premier chapitre (I.3.3) a mis en lumière la forte dépendance de la transconductance vis à vis de la mobilité, mais aussi des résistances d'accès aux Drain/Source. Rappelons que :

$$g_m = \frac{\mu_0 C_{ox} W/L}{(1 + \theta(V_{GS} - V_{Th}))^2} \quad (\text{II.8})$$

où $\theta = \theta_0 + Gm_0 R_{SD}$. Dans l'exemple donné dans la Fig. II.7, on observe sur un dispositif à canal de type N soumis à une contrainte correspondant au maximum du courant de substrat, une réduction importante de la mobilité, principalement due à un nombre important de pièges (de type accepteur) générés à l'interface près du Drain et le long du canal. Un changement de transconductance est effectivement le signe d'une diminution de la mobilité des porteurs dans le canal, conséquence d'interactions coulombiennes avec des charges piégées à l'interface Si-SiO₂. De même la réduction de G_m est aussi imputable à un accroissement de la résistance série R_{SD} . Si la dégradation de la mobilité (et le courant de Drain) induit un retard pendant le fonctionnement AC, la dégradation de la transconductance entraîne une augmentation des temps

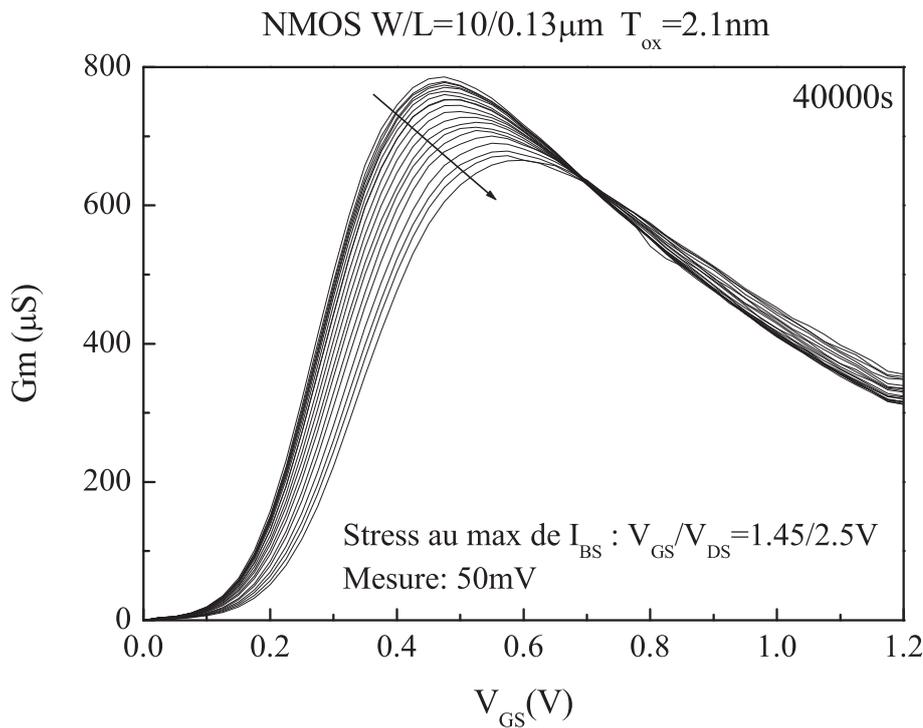


FIG. II.7 – Détérioration de la transconductance lors d'un stress de 40000s aux conditions du maximum de courant substrat.

de transition et donc de la réponse du circuit. Il est donc important de connaître les variations de G_m , afin d'évaluer les effets induits par les défauts générés pendant la contrainte. Puisque G_m détermine la réponse d'un dispositif à une augmentation soudaine du potentiel de Grille, toute dégradation de ce paramètre augmentera le temps de montée ou le temps de descente. Quand la réduction de G_m se produit pour de faibles tensions de grille, c'est que le dispositif est soumis à une forte réduction de mobilité des porteurs, ce qui se traduit globalement par un retard de la réponse du circuit. En revanche s'il y a augmentation de la résistance série, due à une dégradation localisée au Drain, la réduction de G_m sera observée pour des tensions V_{GS} élevées, et engendre des temps de montée/descente augmentés.

II.1.4 Variation de la pente sous seuil

Bien que ce paramètre ne soit pas un critère de fonctionnement essentiel pour les circuits digitaux, il peut permettre de distinguer les dégradations de la structure. L'obtention de l'expression du courant de Drain en régime d'inversion faible [10, 11] est décrite en annexe A :

$$I_{DS} = \mu_0 \frac{W}{L} \exp\left(\beta \frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} (V_{GS} - V_{GS}^*)\right) \exp(-2\beta\phi_F) \quad (\text{II.9})$$

où $\beta = q/kT$.

On peut remarquer dans (II.9) qu'il est plus commode de représenter le courant sur une

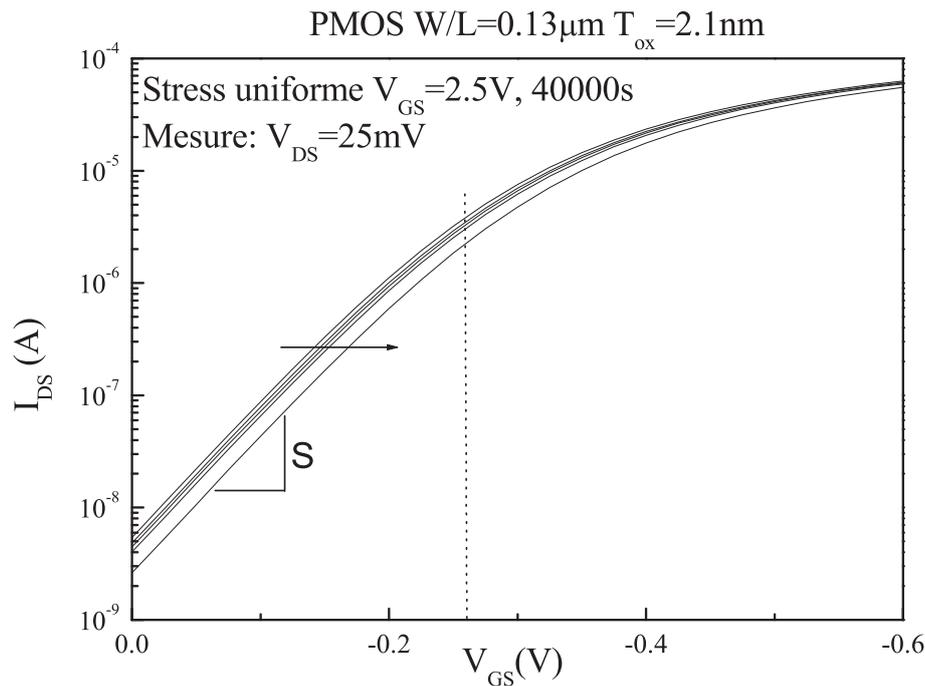


FIG. II.8 – Evolution de la caractéristique I_{DS} de V_{GS} au cours d'un stress à V_{GS} constant.

échelle logarithmique. Ceci permet d'obtenir facilement la pente de $\log(I_D)$ [11] :

$$S = \frac{d \log(I_{DS})}{dV_{GS}} = \beta \frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} \quad (\text{II.10})$$

On peut distinguer deux cas [11] :

1. Un décalage de la courbe parallèlement à elle-même : la pente ne change pas mais la caractéristique se décale, indiquant une dérive de V_{FB} et donc de V_{Th} due à la présence de charges piégées dans l'oxyde. Ce cas est typique d'un stress DC uniforme réalisé à $V_{DS} = 0V$ durant le stress, observable dans l'exemple de la Fig. II.8. Ce type de stress peut également induire un décalage parallèle (déduit par mesure CV d'un décalage ΔV_{FB}) qui peut également être associé à des charges négatives piégées dans l'oxyde de Grille.
2. les changements de pente sous le seuil de la caractéristiques I_{DS} de V_{GS} reflète la création de défauts d'interface, de type accepteur dans le cas d'une diminution du courant dans les NMOS, et donneurs dans le cas du PMOS. En effet quand un stress est réalisé au maximum du courant substrat, on peut observer la combinaison d'une diminution de S avec un décalage parallèle. Ceci est dû à la localisation d'une quantité significative de charges négatives piégées dans la région des zones de diffusion LDD/MDD, tandis que des pièges d'interface de type accepteur sont plus largement répartis le long du canal. Cependant, une charge négative localisée a le même effet sur le courant qu'une répartition plus étendue des pièges d'interface à travers le canal [12]. C'est pourquoi cette technique est préférable pour les dégradations uniformes, où le décalage parallèle permet de remon-

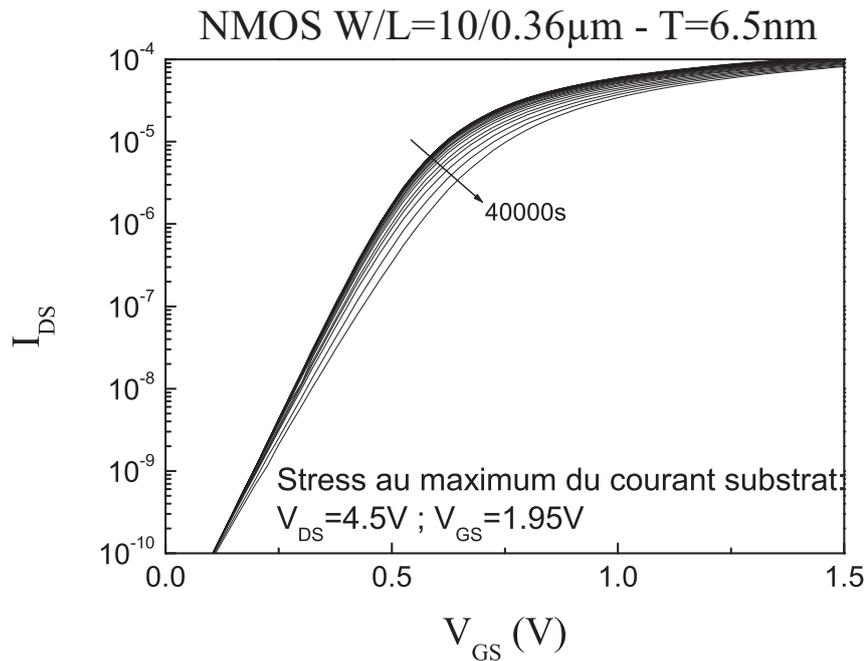


FIG. II.9 – Evolution de la caractéristique sous seuil I_{DS} de V_{GS} .

ter à la charge piégée. En effet lorsque les conditions de stress induisent une dégradation localisée, la détermination du type de défauts et de leur localisation s'avèrent délicate avec les seules informations recueillies avec cette méthode.

Par conséquent, en présence de dégradations localisées, toutes les caractérisations sont fortement sensibles à l'étalement de la région dégradée. Seule une valeur moyenne de la densité de piège d'interface est donc accessible. On l'obtient par la mesure de la pente avant et après le stress, puis la détermination des changements au cours de la période de stress. En traçant :

$$\Delta N_{it} = \frac{\beta C_{ox}}{q} \Delta S^{-1} \quad (\text{II.11})$$

pour une gamme de valeurs du potentiel de surface comprises entre ϕ_f et $2\phi_f$, on peut alors extraire la valeur moyenne du nombre de pièges produits à l'interface en régime d'inversion faible. Dans l'exemple de la Fig. II.10, la courbe de ΔN_{it} à l'aide ΔS on obtient une création d'états d'interface de $2.2 \cdot 10^{12} \text{cm}^{-2}$ pendant un stress de 40000s effectué au maximum du courant substrat ($V_D = 4.5V$ et $V_G = 1.95V$) sur un transistor NMOS ($W/L = 10/0.36\mu\text{m}$, $T_{ox} = 6.5\text{nm}$).

Un décalage parallèle des courbes implique un piégeage de charges dans l'oxyde, conséquence d'une dégradation uniforme. Si le décalage s'accompagne d'un changement de pente, on ne peut pas conclure catégoriquement quant à la présence ou non de charges piégées. Il nous faut donc faire appel à d'autres techniques plus fines pour dissocier ΔN_{it} de ΔN_{ot} . Dans la prochaine section nous allons nous intéresser au pompage de charges.

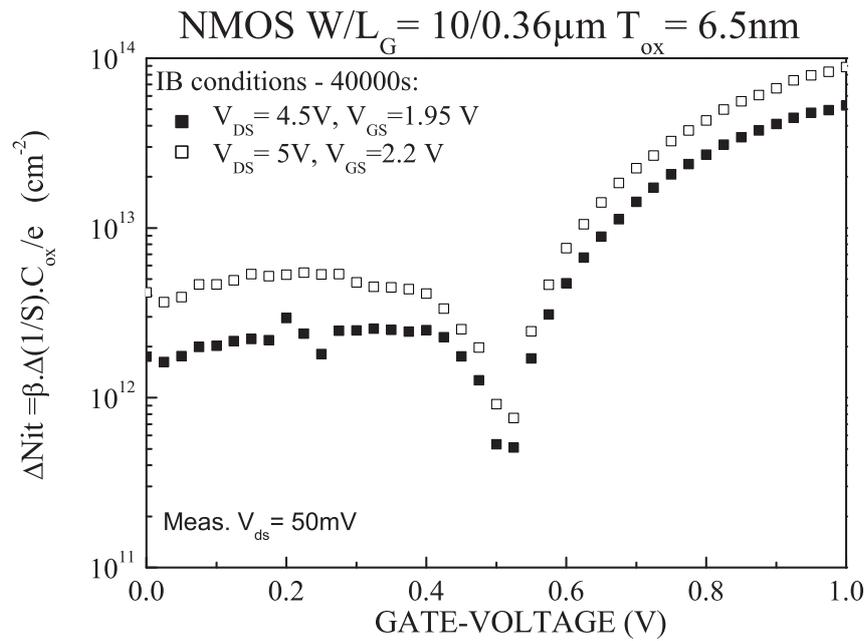


FIG. II.10 – Evaluation de la variation du nombre d'états d'interface exprimé à l'aide de ΔS en fonction de la tension de Grille.

II.2 Analyse des défauts par pompage de charges

Le pompage de charge est un puissant moyen d'investigation des défauts dans les dispositifs MOS. La mesure par pompage a été développée par Brugler et Jesper en 1969 [13]. Plus tard, en 1973, Simmons et Wei [14][15] proposeront une première explication des cinétiques de vidage et de remplissage des pièges d'interface, finalement améliorée par Groeseneken et al [16] en 1984. Cette méthodologie, mise en place par Groeseneken est toujours d'actualité et autorise la dissociation des mécanismes responsables des dommages causés à la structure testée. Des travaux plus récents ont apporté une sophistication à la technique permettant d'obtenir $D_{it}(E)$ la densité énergétique [17][18]. Actuellement, l'expérimentateur est confronté aux problèmes posés par les faibles dimensions des dispositifs : la finesse des isolants de Grille est responsable d'un courant tunnel devenu non négligeable devant le courant pompé. Pour pallier à problème, Masson et Autran [19] ont proposé une méthode pour séparer les deux composantes par le calcul du courant de grille en fonction du signal de Grille sinusoïdal, que nous avons appliquée aux signaux triangulaires [20].

D'une manière générale, le pompage de charge est basé sur la mesure du courant substrat généré par la recombinaison de la charge piégée sur les états d'interface avec la charge présente dans le canal. On peut distinguer deux façons de procéder, en utilisant des signaux à deux niveaux, ou à trois niveaux. Si la première, la plus utilisée, renseigne sur la densité surfacique des défauts, la seconde permet d'obtenir la répartition énergétique de ceux-ci dans la bande interdite. Toutes deux présentent des limitations expérimentales pour les dispositifs à oxydes ultra-minces.

II.2.1 Pompage de charges standard à deux niveaux

a) Principe

On applique sur la Grille un signal à deux niveaux dont un ou les deux balayent un intervalle de tensions comprenant V_t et V_{fb} . Le Drain et la Source sont mis à la masse, le courant pompé est mesuré au substrat. Ces signaux sont schématisés sur la Fig. II.11. Lors des variations du signal de Grille et donc du potentiel de surface, les états d'interface émettent et capturent des électrons pour rester en équilibre avec les bandes d'énergie et le niveau de Fermi. Dans le cas d'une structure MOS à substrat P en équilibre en accumulation, un signal de grille fait basculer rapidement le dispositif en inversion, les électrons sont fournis par la source et le Drain (qui constituent des réservoirs d'électrons). Les états d'interface capturent les électrons du canal, afin de maintenir l'équilibre. Lorsque le signal de grille impose à nouveau l'accumulation, les trous remontent vers l'interface et les pièges émettent les électrons dans le canal vers la Source et le Drain. La formation de la couche d'accumulation est plus rapide que l'émission des électrons, ceux-ci se recombinent donc avec les trous du substrat donnant naissance à un flux de trous dans le substrat. Ce flux est à l'origine d'un courant Substrat, dit courant pompé noté I_{CP} .

b) Théorie de la statistique de génération-recombinaison

La statistique de remplissage et de vidage des pièges de charges dans la bande interdite fut entièrement décrite par Shockley, Read et Hall [21]. Nous allons ici en exposer les bases. On peut définir les pièges par des centres de génération ou de recombinaison de charges (centre GR).

Vis-à-vis des électrons, il faut distinguer :

- les pièges accepteurs, chargés si occupés et neutres sinon.
- les pièges donneurs, neutres si occupés et chargés sinon.

Outre cette caractéristique, on utilise également la section efficace de capture $\sigma(m^2)$ pour définir les centres GR. On peut ainsi définir le coefficient de capture (m^3s^{-1}) d'une charge par le piège, respectivement pour les trous et les électrons :

$$c_p = \sigma_p v_T \quad (\text{II.12})$$

$$c_n = \sigma_n v_T \quad (\text{II.13})$$

où $v_T(m s^{-1})$ est la vitesse des porteurs due à leur agitation thermique. Le coefficient de capture augmente donc avec la température du semi-conducteur. De même on peut définir des coefficients d'émission $e_p(m^3s^{-1})$ et $e_n(m^3s^{-1})$ pour les trous et les électrons et $e_n(m^3s^{-1})$. $c_{p/n}$ et $e_{p/n}$ représentent respectivement les probabilités de capture et d'émission.

Nous avons introduit plus haut N_{it} et f_s la densité de défauts et la probabilité d'occupation

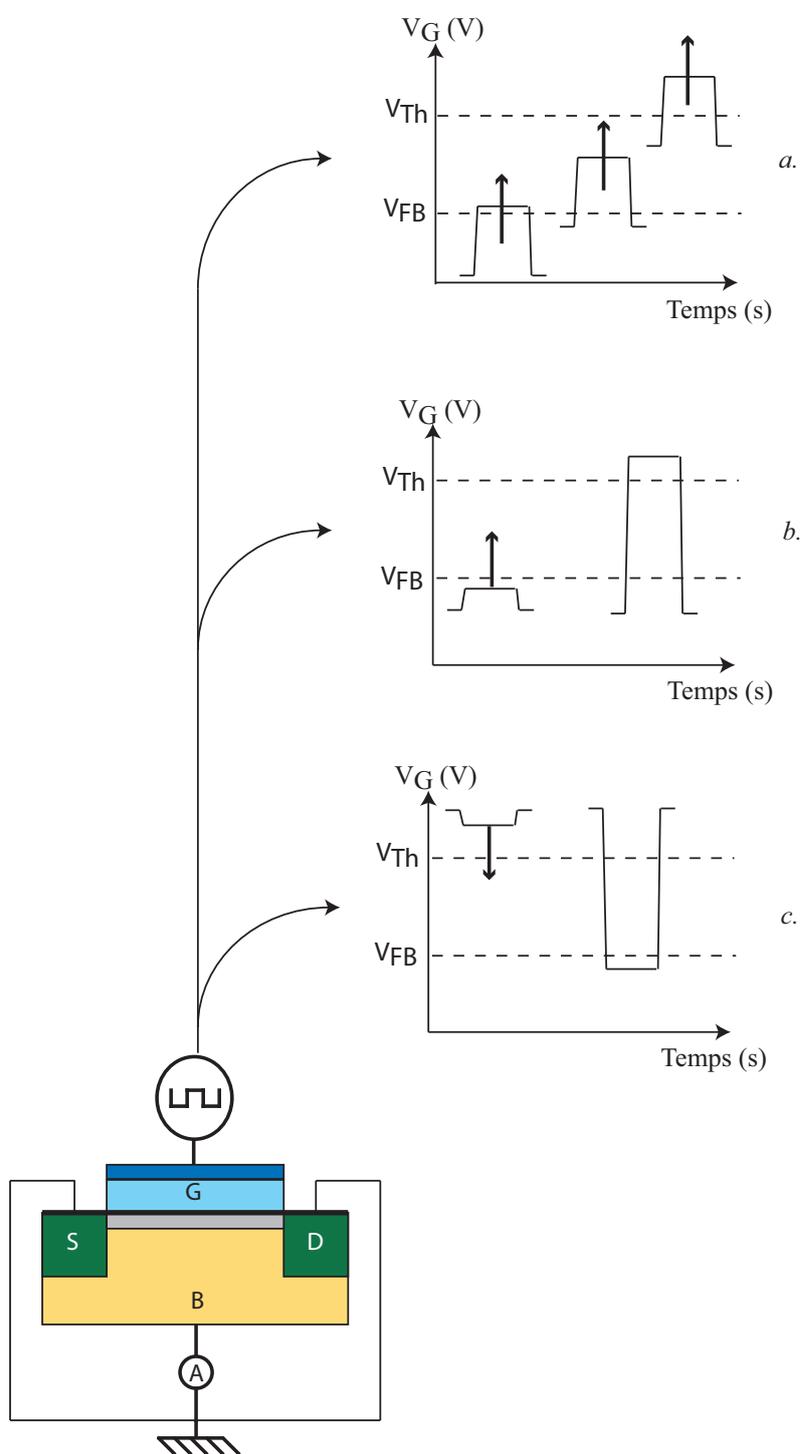


FIG. II.11 – Forme des signaux employés pour le pompage de charge à deux niveaux. a : la forme du signal est translatée, b : V_G^{haut} est variable, c : V_G^{bas} varie.

d'un centre par un électron. La proportion de centres occupés est donc $N_{it}f_s$, et la proportion de centres inoccupés est $N_{it}(1 - f_s)$. Ces quantités sont utiles pour définir les taux de capture des trous et des électrons $\Gamma_{c,p}$ et $\Gamma_{c,n}$ ainsi que les taux d'émission $\Gamma_{e,p}$ et $\Gamma_{e,n}$. La probabilité de capture d'un électron qui "descend" de la bande de conduction vers la bande valence est le produit du nombre de pièges libres par le coefficient de capture, soit :

$$\Gamma_{c,n} = c_n N_{it} (1 - f_s) \quad (\text{II.14})$$

De même pour l'émission des électrons, on multiplie le coefficient d'émission par le nombre de pièges occupés :

$$\Gamma_{e,n} = e_n N_{it} f_s \quad (\text{II.15})$$

Avec un raisonnement identique pour les trous (l'émission d'un trou correspond à la capture d'un électron, et inversement), on obtient :

$$\Gamma_{e,p} = e_p N_{it} (1 - f_s) \quad (\text{II.16})$$

$$\Gamma_{c,p} = c_p N_{it} f_s \quad (\text{II.17})$$

c) Expression du courant pompé I_{CP}

Brugler et Jesper [13] observèrent les premiers ce courant et proposèrent l'expression :

$$I_{CP} = f_{CP} S_{eff} Q_{it} + \beta f_{CP} S_{eff} C_{OX} (V_{GB} - V_t) \quad (\text{II.18})$$

dans laquelle S_{eff} est la surface effective de la Grille, α représente la quantité de porteurs minoritaires de la couche d'inversion qui se recombine à l'interface avec les trous du substrat lors du passage du régime d'inversion au régime d'accumulation.

Le premier terme de (II.18) est le courant dû à la charge Q_{it} pompée sur les états d'interface [16][18] :

$$Q_{it} = q \int_{E_1}^{E_2} D_{it}(E) dE = q^2 D_{it} \Delta\psi_S \quad (\text{II.19})$$

où $D_{it}(E)$ est la densité d'états d'interface au niveau d'énergie E , D_{it} est ainsi la densité moyenne sur le domaine $\Delta E = E_2 - E_1$ correspondant à une variation du potentiel de surface $\Delta\psi_S$. Le courant pompé sera donc nul lorsque $\frac{\partial\psi_S}{\partial t} = 0$, et son calcul se ramène à l'étude du vidage et du remplissage des pièges suivant les variations du potentiel de surface. On appelle E_S la fonction de remplissage des pièges. La figure II.12 montre sur la même échelle de temps les variations du signal de grille, et les variations des niveaux d'énergie des bandes de valence et de conduction, ainsi que les cinétiques de remplissage des pièges. A partir de la cinétique du remplissage des pièges d'interface [14][15], on décrit cette cinétique de vidage-remplissage en sept étapes [16] à l'aide de la Fig II.12 :

1. On suppose que la structure est en accumulation à l'équilibre thermodynamique (temps t_i), c'est à dire pour $E_S = E_{FI}$, le niveau de Fermi correspondant à V_{GBas} . Lorsque le signal de grille commence à augmenter, les pièges libèrent leurs trous vers la bande de valence (BV) afin de maintenir l'équilibre et E_S suit les variations de E_F . Ceci se produit en déplétion jusqu'à l'inversion faible, $V_{GS} = V_{IF}$ et $E_S = E_{IF}$.
2. Lorsque l'émission des trous n'est plus suffisante pour maintenir l'équilibre, celle-ci se fait alors hors équilibre, vers le substrat. La probabilité de capture des électrons de la bande de conduction (BC) devient supérieure à celle d'émission des trous vers la bande de valence.
3. Lorsque la couche d'inversion est formée (inversion forte), pour $V_{GS} = V_t$, $E_S = E_{em,h}$, les électrons provenant de la Source et du Drain sont capturés par les états d'interface.
4. Lorsque l'équilibre avec E_F est atteint (temps t_h), V_{GS} est constant et égal à V_{Ghaut} , tous les pièges sont remplis.
5. Quand V_{GS} commence à diminuer, les pièges émettent les électrons vers la BC, dans le canal de manière à maintenir l'équilibre avec le niveau de Fermi.
6. Quand l'émission des électrons vers la BC ne peut plus maintenir l'équilibre, il y a capture de trous en provenance de la BV, hors équilibre et $E_S = E_{em,e}$.
7. Pour $V_{GS} < V_{fb}$ l'équilibre est maintenu par la capture des trous de la BV.

Remarquons dès à présent les notations V_t et V_{fb} . Si d'un point de vue conceptuel, leur signification s'apparente aux paramètres transistors V_{Th} et V_{FB} , ils correspondent au remplissage des niveaux d'énergie par les pièges.

Le courant pompé peut alors se décomposer en quatre composantes I_1 , I_2 , I_3 et I_4 telles que [16] :

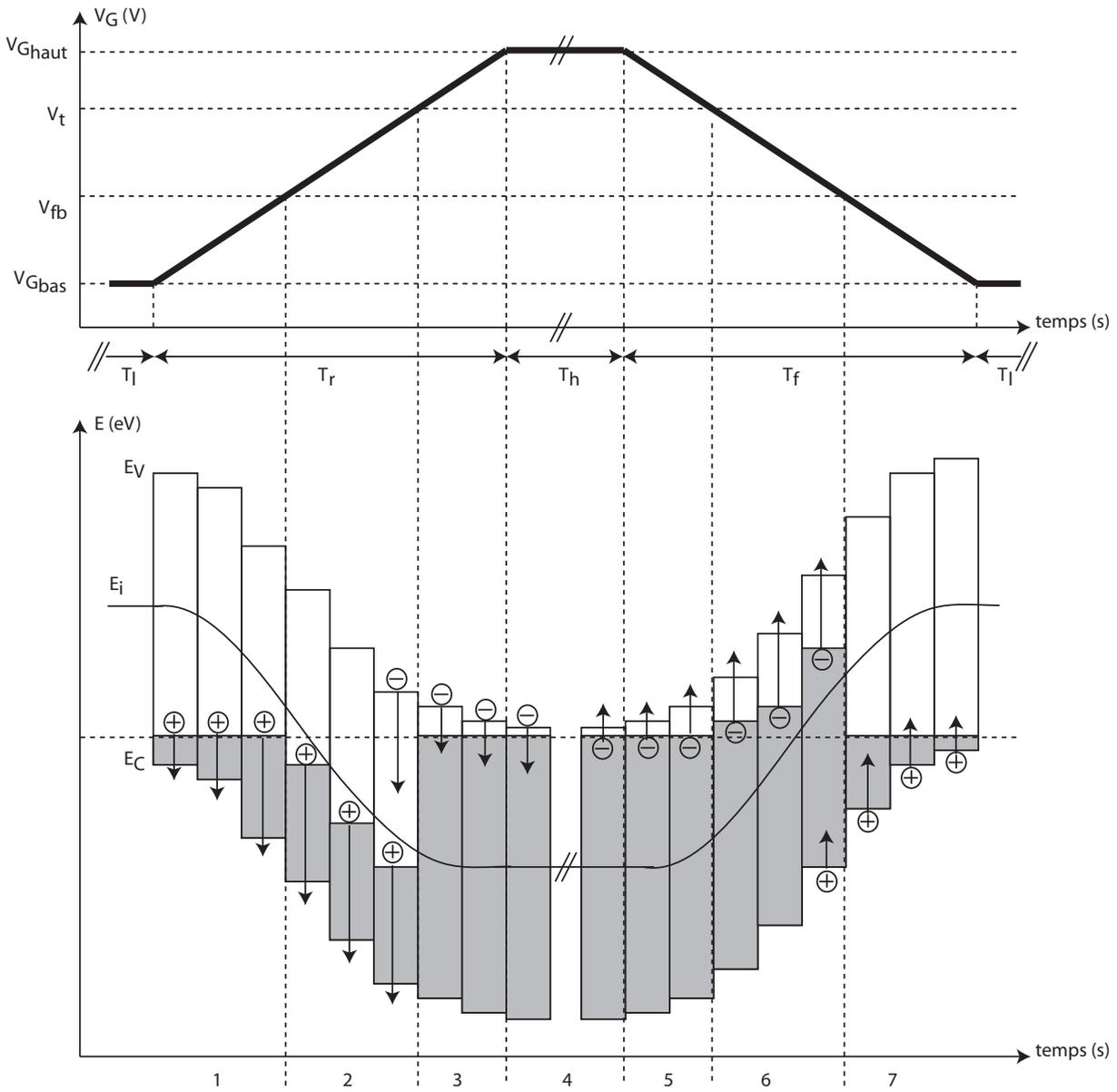


FIG. II.12 – Processus d'émission et de capture entrant en jeu pendant un cycle du signal de grille pendant le pompage de charge à deux niveaux [16]

$$I_{D/S} = q^2 f_{CP} S_{eff} \left[\underbrace{-D_{it} \Delta \psi_S^e}_{I_1} + \underbrace{D_{it} \Delta \psi_S^{ee}}_{I_2} \right] \quad (\text{II.20})$$

$$I_B = q^2 f_{CP} S_{eff} \left[\underbrace{-D_{it} \Delta \psi_S^h}_{I_3} + \underbrace{D_{it} \Delta \psi_S^{he}}_{I_4} \right] \quad (\text{II.21})$$

Les courants I_1 et I_2 correspondent aux étapes 3, 5 et 6, de même I_3 et I_4 sont relatifs aux étapes 1, 2 et 7. Dans le tableau II.1 on peut trouver les significations des différents $\Delta \psi$.

Variations de ψ	Mouvements de charges associés
$\Delta \psi_S^e$	capture des électrons de la BC en provenance des Source et Drain
$\Delta \psi_S^{ee}$	émission des électrons vers la BC collectés à la Source et au Drain
$\Delta \psi_S^h$	capture des trous vers la BV collectés au substrat
$\Delta \psi_S^{he}$	émission des trous de la BV en provenance du substrat

TAB. II.1 – Signification des $\Delta \psi$ de (II.20) et (II.21)

Le calcul de I_{CP} revient à calculer $(-\Delta \psi_S^e + \Delta \psi_S^{ee})$ ou $(-\Delta \psi_S^h + \Delta \psi_S^{he})$. Dans le cas du courant pompé mesuré au substrat, on a :

$$I_{CP} = q f_{CP} S_{eff} \left[- \int_{E_{IF}}^{E_{em,e}} D_{it}(E) dE + \int_{E_{IF}}^{E_{em,h}} D_{it}(E) dE \right] \quad (\text{II.22})$$

$$= q f_{CP} S_{eff} D_{it}(E_{em,h} - E_{em,e}) \quad (\text{II.23})$$

Basé sur les travaux de Simmons [14][15], le calcul de $E_{em,e}$ et $E_{em,h}$ a été simplifié par un développement limité effectué par Groeseneken :

$$E_{em,e} = E_i - kT \ln \left(v_T \sigma_{e,n} n_i t_{em,e} + \exp \left(\frac{E_i - E_{F,inv}}{kT} \right) \right) \quad (\text{II.24})$$

$$E_{em,h} = E_i + kT \ln \left(v_T \sigma_{h,n} n_i t_{em,h} + \exp \left(- \frac{E_i - E_{F,acc}}{kT} \right) \right) \quad (\text{II.25})$$

Notons ici que ces expressions sont obtenues par développement limité, mais que Masson [22] a obtenu dans le cadre d'une étude dynamique des expressions similaires avec un facteur $\frac{1}{\ln(2)}$ devant v_T . En notant que $E_i - E_{F,inv}$ et $E_i - E_{F,acc}$ s'expriment en fonction de ϕ_F et que les

variations de $E_{em,e}$ et $E_{em,h}$ sont pratiquement indépendantes du dopage, on peut écrire :

$$E_{em,e} = E_i - kT \ln (v_T \sigma_{e,n} n_i t_{em,e}) \quad (\text{II.26})$$

$$E_{em,h} = E_i + kT \ln (v_T \sigma_{e,h} n_i t_{em,h}) \quad (\text{II.27})$$

Nous avons déjà défini la vitesse thermique v_T , les sections efficaces de capture $\sigma_{e,n/h}$ des porteurs et $t_{em,e}$ et $t_{em,h}$ sont les temps durant lesquels, les électrons/trous sont émis hors équilibre, c'est à dire le temps passé entre V_t et V_{fb} [23] :

$$t_{em,e} = \frac{|V_t - V_{fb}|}{\Delta V_G} t_f \quad (\text{II.28})$$

$$t_{em,h} = \frac{|V_t - V_{fb}|}{\Delta V_G} t_r \quad (\text{II.29})$$

En réarrangeant les termes sous le logarithme on obtient finalement l'expression du courant pompé [14, 15] :

$$I_{CP} = 2qkT D_{it} f_{CP} S_{eff} \left[\ln (v_T n_i \sqrt{\sigma_p \sigma_n}) + \ln \left(\frac{|V_t - V_{fb}|}{\Delta V_G} \sqrt{t_r t_f} \right) \right] \quad (\text{II.30})$$

Dans le cas d'un signal triangulaire, $t_{em,e}$ et $t_{em,h}$ s'expriment comme [16] :

$$t_{em,e} = \frac{|V_t - V_{fb}|}{\Delta V_G} \frac{\delta}{f_{cp}} \quad (\text{II.31})$$

$$t_{em,h} = \frac{|V_t - V_{fb}|}{\Delta V_G} \frac{(1 - \delta)}{f_{cp}} \quad (\text{II.32})$$

ou δ est le rapport cyclique du signal employé ($0 < \delta < 1$). Le courant pompé devient :

$$I_{CP} = 2qkT D_{it} f_{CP} S_{eff} \left[\ln (v_T n_i \sqrt{\sigma_p \sigma_n}) + \ln \left(\frac{|V_t - V_{fb}|}{\Delta V_G} \frac{\sqrt{\delta(1 - \delta)}}{f_{cp}} \right) \right] \quad (\text{II.33})$$

Dans le cas d'un signal trapézoïdal, employé pour les mesure à V_{Ghaut} , V_{Gbas} ou V_{Gbase} fixe, les temps de capture sont donnés par :

$$t_{em,e} = \frac{|V_{fb} - V_t|}{\Delta V_G} t_m \quad (\text{II.34})$$

$$t_{em,h} = \frac{|V_{fb} - V_t|}{\Delta V_G} t_d \quad (\text{II.35})$$

Ces expressions permettent d'obtenir le courant pompé lorsque le signal de Grille a une forme trapézoïdale avec des temps de montée/descente t_m/t_d :

$$I_{CP} = 2qkTD_{it}f_{CP}S_{eff} \ln \left[v_T n_i t \sqrt{\sigma_p \sigma_n} \frac{|V_{fb} - V_t|}{\Delta V_G} \sqrt{t_d t_m} \right] \quad (\text{II.36})$$

Enfin pour un signal sinusoïdal la forme des temps de montée/descente est plus complexe [18] :

$$t_{em,e} = t_{em,h} = \frac{1}{2\pi f_{CP}} \underbrace{\left[\sin^{-1} \left(\frac{2|V_{fb} - V_{off}|}{\Delta V_G} \right) + \sin^{-1} \left(\frac{2|V_{off} - V_t|}{\Delta V_G} \right) \right]}_{t^*} \quad (\text{II.37})$$

Où V_{off} est la tension de décalage du signal (représentant la valeur moyenne du signal). Ainsi on arrive au courant pompé :

$$I_{CP} = 2qkTD_{it}f_{CP}S_{eff} \ln \left[\frac{v_T n_i t^* \sqrt{\sigma_p \sigma_n}}{2\pi f_{CP}} \right] \quad (\text{II.38})$$

L'emploi de signaux sinusoïdaux présente deux avantages. Le premier est de maximiser les échanges de charges entre les états d'interface et le canal, en effet le signal "passe" plus de temps entre V_{fb} et V_t . Le second intérêt est de minimiser le courant tunnel qui peut traverser l'oxyde si celui est ultra-mince.

Avant d'aborder l'aspect expérimental du pompage de charges à deux niveaux, il nous faut revenir à l'équation (II.18). Le second terme de l'équation est appelé composante géométrique. Ce courant est dû à la recombinaison des porteurs minoritaires qui n'ont pas le temps d'atteindre les Sources et Drain, avec les porteurs majoritaires provenant du substrat lors du basculement de l'inversion forte vers le régime d'accumulation. Pour réduire cette erreur systématique, plusieurs solutions ont été proposées :

1. Elliot [24] a suggéré d'appliquer à la Source et au Drain une polarisation inverse V_R . Ceci a pour effet de réduire la longueur du canal par extension des zones désertées autour des

jonctions S-B et D-B. Cette technique risque toutefois de masquer une partie des états d'interface proches des Source et Drain.

2. Une autre méthode, introduite par Brugler[13] consiste en l'augmentation des temps de montée et de descente du signal afin de ralentir le basculement inversion-accumulation, laissant ainsi plus de temps aux porteurs minoritaires pour regagner la Source et le Drain.
3. Une dernière façon de procéder consiste à respecter la condition géométrique :

$$W/L \gg 1$$

c'est à dire des transistors de plus petites longueurs $W \geq 20\mu m$ et $L \leq 1\mu m$. Il faut prendre garde à la surface active totale du canal : si celle-ci est trop petite, on peut être dans l'incapacité de mesurer le courant pompé, trop faible. Ceci pose le problème de la sensibilité de la mesure, il faut trouver un compromis entre fréquence, forme du signal et surface du canal pour que le matériel de mesure puisse réaliser l'acquisition.

II.2.2 Mesures expérimentales et interprétations

La figure II.13 décrit schématiquement le banc de mesure utilisé au laboratoire. Les mesures sont contrôlées par l'ordinateur de commande grâce au programme CMOS écrit en HT basic, évolution 32bit du logiciel HP basic. L'utilisation de la boîte de commutation Hp 41501B autorise de nombreuses configurations et permet grâce à l'interface HPIB, d'enchaîner les phases de stress et de mesure, y compris CP, sans interventions manuelles sur les connexions (toujours délicates vis à vis des décharges ESD). Dans le cas particulier du pompage de charge, nous avons utilisé deux méthodes pour l'acquisition sur l'électromètre Hp4156c :

- pour les mesures à ΔV_A fixé, on utilise le mode *sampling*, et le générateur Hp 8112B est contrôlé via HPIB, pour chaque point (position du pulse) on réalise une moyenne de I_{CP} sur 10s.
- les mesures à un niveau fixé sont réalisées en mode *sweet*, le niveau variable du pulse est contrôlé ("trigé") par le Hp4156c, et l'acquisition est instantanée pour chaque point.

Ce deuxième type d'acquisition est le plus pratique car le plus rapide, et donc facilite l'enchaînement des mesures et des stress sur des temps longs.

Que l'on procède à une mesure par pompage de charge avec le niveau haut ou bas variable, ou bien avec l'amplitude constante, il est impératif de balayer V_{fb} et V_t , en prenant soin que l'amplitude maximale du signal ne soit pas trop importante pour ne pas dégrader la structure. Pour chaque dispositif il faut donc calibrer la mesure. Dans tous les cas, l'expression II.30 reste valable et permet de remonter aux informations essentielles.

a) Mesures à fréquence variable

V_{fb} et V_t étant connus, on doit déterminer la sensibilité en fréquence. En effet pour les faibles fréquences, la recombinaison de la charge piégée sur les états d'interface ne se fait plus avec

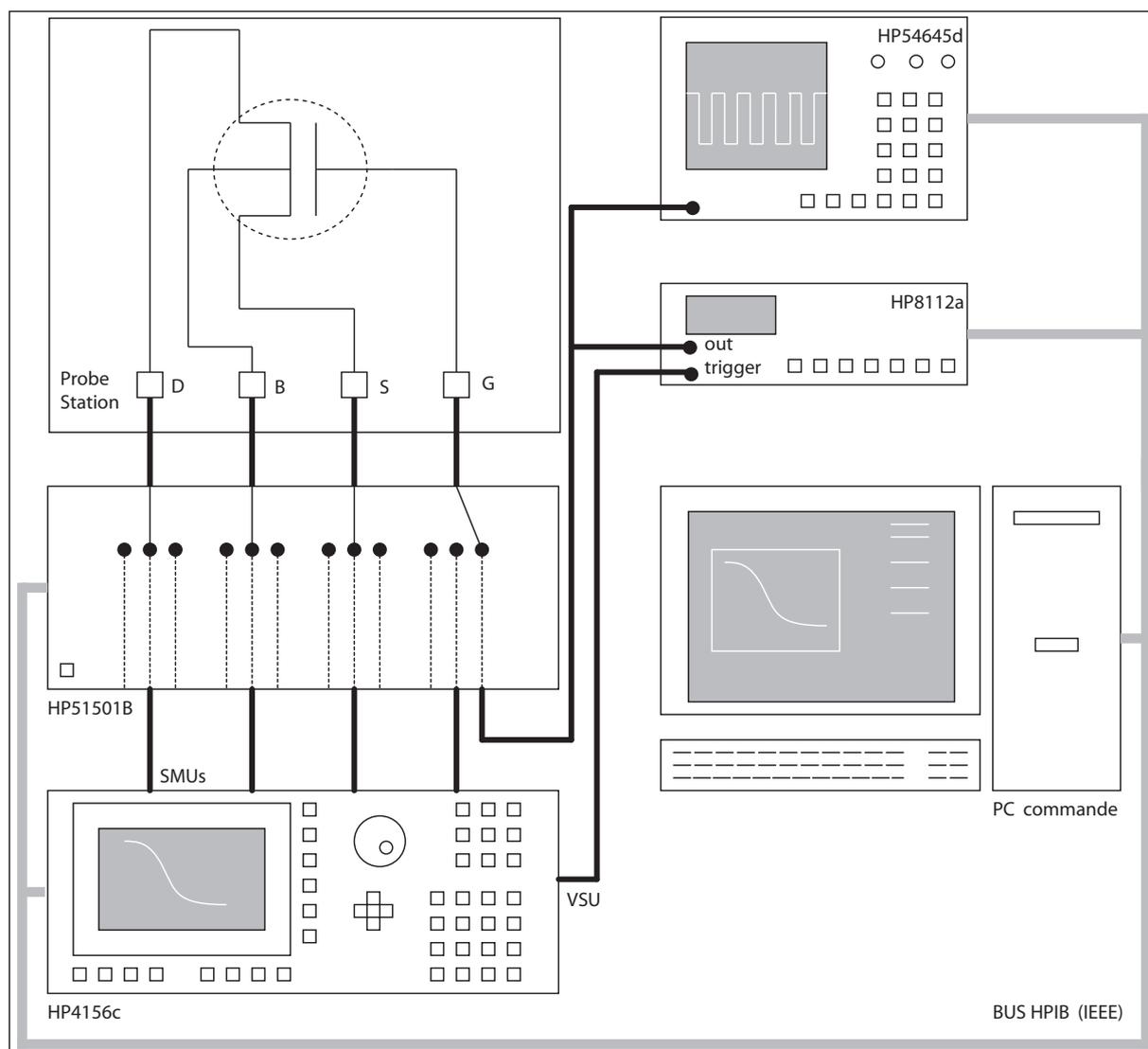


FIG. II.13 – Schéma de principe pour la mesure CP, le Drain et la Source sont au même potentiel. Les dispositifs étudiés sont sur des wafers de test prévus à cet effet avec des pads en surface pour contact par micro-pointes.

les porteurs majoritaires du substrat car ils ne sont plus fournis assez rapidement. C'est alors la composante tunnel qui prédomine ($T_{ox} < 2.5nm$). On utilise un signal triangulaire, qui balaie l'intervalle de tension comprenant la tension de bande plate et la tension de seuil, en faisant varier la fréquence. En traçant la charge pompée en fonction de la fréquence de mesure sur une

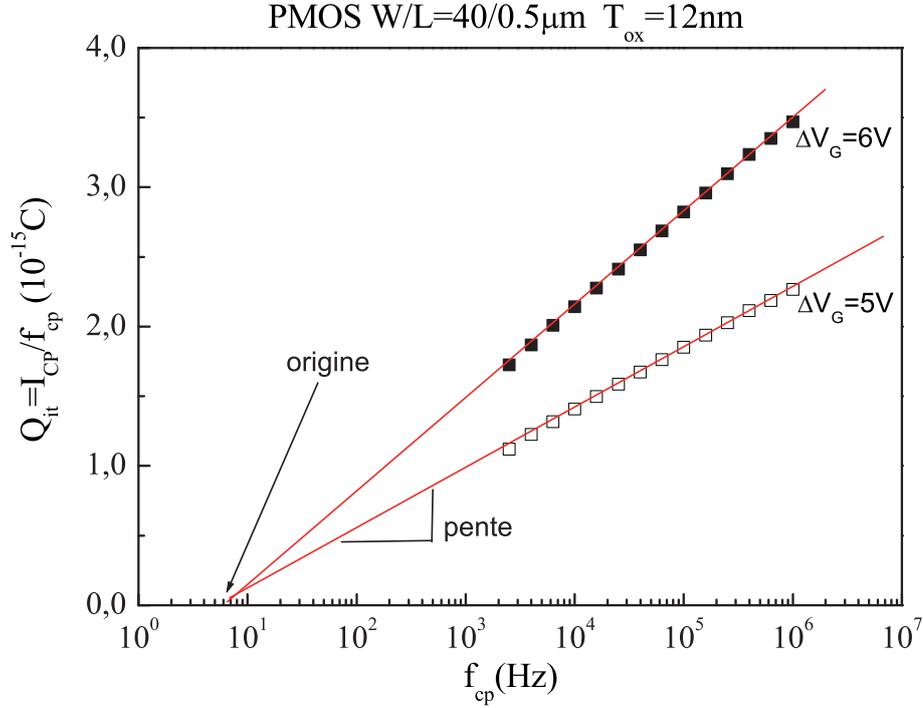


FIG. II.14 – Mesure de pompage de charges à amplitude constante avec un signal triangulaire, $V_{fb} = -1V$, $V_t = 0.7V$.

échelle logarithmique. La Fig. II.14 montre pour deux amplitudes différentes, la charge pompée Q_{it} par cycle de signal de grille. On constate d'une part l'augmentation du courant pompé avec la fréquence, et d'autre part le changement de la pente avec l'amplitude du signal. On obtient une relation linéaire dont le couple pente/origine permet de remonter à la concentration de défaut à l'interface D_{it} ainsi qu'à la moyenne géométrique de la section efficace de capture $\bar{\sigma}$. En effet, de (II.30) on déduit :

$$pente = 2 \ln(10) qkT S_{eff} D_{it} \quad (II.39)$$

$$origine = -2qkT S_{eff} f_{CP} D_{it} \ln \left(v_T n_i \sqrt{\sigma_p \sigma_n} \frac{|V_t - V_{fb}|}{\Delta V_G} \sqrt{t_r t_f} \right) \quad (II.40)$$

Sur la figure II.14, pour ΔV_G fixé à 6V, la mesure sur le transistor PMOS de technologie 0.5 μ m et d'oxyde de Grille 12nm, donne avec la pente :

$$D_{it} = 7.32 \cdot 10^{10} eV^{-1} cm^{-2}$$

De (II.40) on tire la valeur de la section de capture efficace pour les électrons et les trous

$$\sqrt{\sigma_p \sigma_n} = 9.49 \cdot 10^{-16} \text{ cm}^2$$

Remarque : On obtient une valeur approchée de v_T avec $\frac{1}{2}mv_T^2 = \frac{3}{2}kT$, soit $v_t \approx 1.1 \cdot 10^7 \text{ m.s}^{-1}$.

b) Mesures à amplitude constante

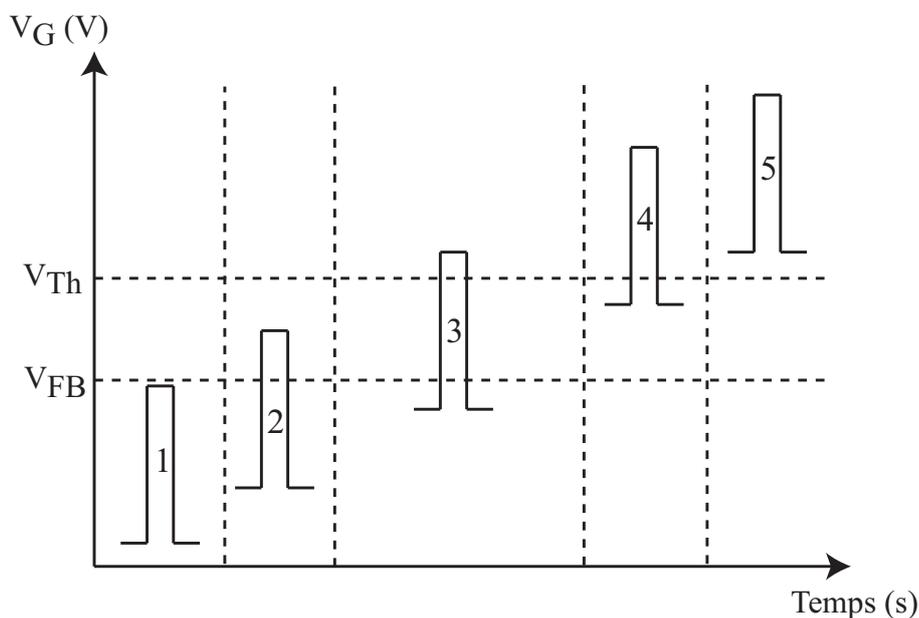


FIG. II.15 – Evolution du pulse de la mesure à amplitude constante vis à vis des tensions de seuil et de bande plate.

Pour cette mesure, l'analyseur est en mode "sampling" (temporel). L'amplitude ΔV_A du signal de grille est fixé, et c'est le niveau de la base du pulse qui sert de référence. La manipulation peut se décomposer en cinq phases (Fig. II.15) :

1. tout le pulse est hors du domaine de tension $[V_{fb}; V_t]$, le courant pompé est minimal,
2. le pulse entre dans l'intervalle et le courant pompé, ce cas de figure correspond à la mesure à $V_{G_{bas}}$ fixe
3. tout le pulse est dans l'intervalle, le courant pompé est maximal et correspond à la densité moyenne d'états d'interfaces dans la bande interdite, D_{it} .
4. une partie du pulse ressort de l'intervalle et I_{CP} commence à diminuer, on est dans le cas correspondant à $V_{G_{haut}}$ fixé
5. tout le pulse est ressorti de l'intervalle est le courant substrat est revenu à sa valeur minimale initiale

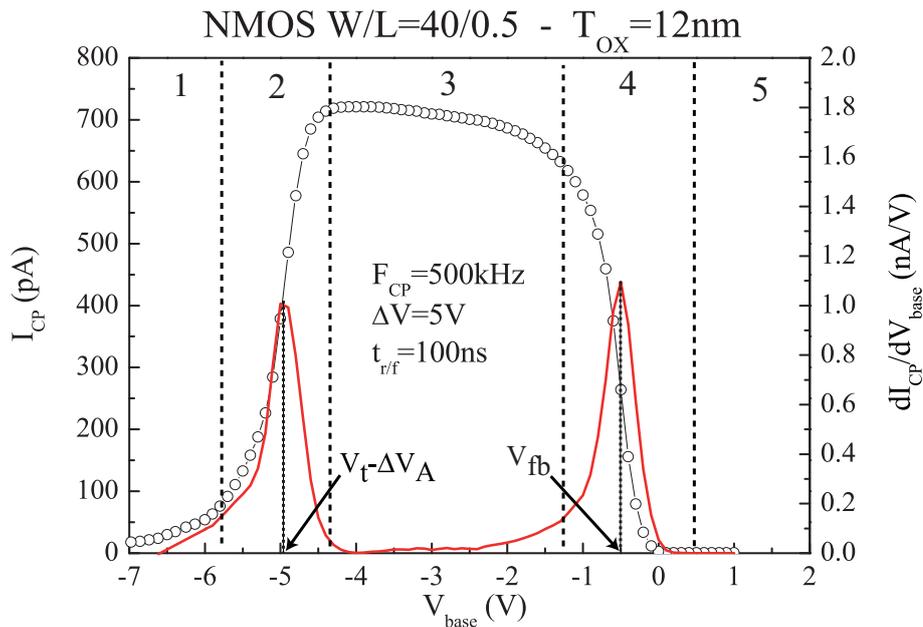


FIG. II.16 – Mesures de pompage de charges à amplitude constante sur transistor NMOS $W/L = 40/0.5\mu m$, $T_{ox} = 12nm$, les fronts du signal ont une durée de 100ns, le rapport cyclique est de 50%.

Dans l'exemple de la Fig. II.16, la valeur maximale du courant résultant du pompage de charges est $I_{CP}^{max} = 721pA$. Pour cette valeur, toute la bande interdite est sondée, et la densité d'états s'obtient avec (II.33). Le calcul donne :

$$D_{it} = 6.33 \cdot 10^{10} eV^{-1} cm^{-2}$$

On obtient une valeur cohérente avec le résultat obtenu précédemment sur un dispositif identique.

c) Mesures à V_{Gbas} ou V_{Ghaut} fixe

Pour cette mesure l'acquisition se fait en mode "sweep", c'est à dire le temps de mesure pour chaque point est réduit au temps nécessaire à l'analyseur pour l'acquisition (en fonction de la précision sélectionnée). Un des deux niveaux doit balayer l'intervalle $[V_t; V_{fb}]$, afin que l'ensemble des niveaux d'énergie susceptibles de correspondre aux pièges soit balayé. Suivant le niveau que l'on fixe, la mesure sera plus sensible à un type de défaut donné. En effet pour le NMOS (PMOS) le niveau bas fixé va permettre l'étude du piégeage de la charge positive, alors que le niveau haut fixé sera plutôt dédié à la charge piégée négative. Sur la Fig. II.17 on observe les courbes typiques de pompage de charge à un niveau fixé : le courant pompé sature lorsque toute la bande interdite a été analysée, c'est à dire quand les deux niveaux du pulse sont de part et d'autre de l'intervalle $[V_t; V_{fb}]$. Lorsque la mesure démarre du niveau haut (fixé et supérieur

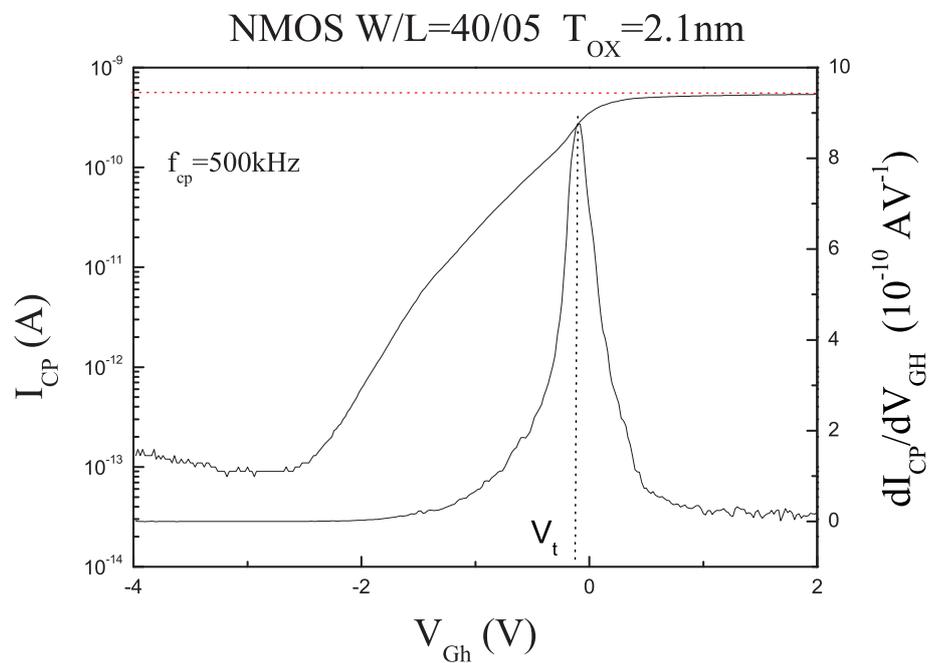
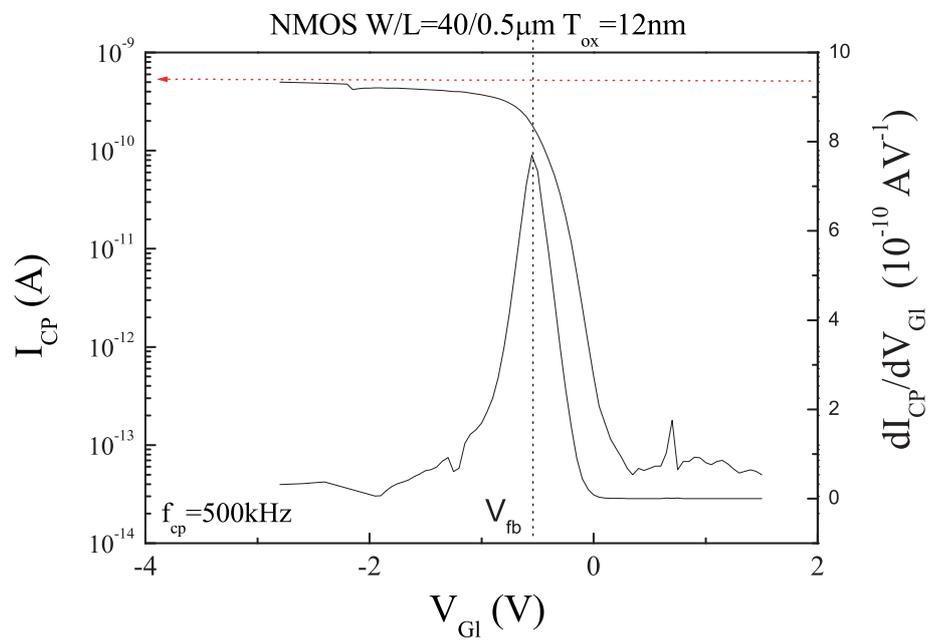


FIG. II.17 – *Pompage de charge à deux niveaux, avec un niveau fixé sur transistor NMOS $W/L = 40/0.5\mu m$, $T_{ox} = 12nm$, les fronts du signal ont une durée de 100ns, le rapport cyclique est de 50%.*

à V_t pour un NMOS) la saturation est atteinte quand V_G devient inférieur à V_{fb} . pour le niveau bas fixé (inférieur à la tension de bande plate) la saturation se produit pour $V_G > V_t$. La valeur maximale de I_{CP} permet de remonter à D_{it} si $\bar{\sigma}$ est connu. Il peut arriver que la saturation ne soit pas totale, ceci peut être attribué au rapport courant tunnel/courant pompé non négligeable ; Declercq et Jespers [25] ont attribué ce phénomène à des états lents situés près de l'interface. La couche d'inversion augmentant avec la tension V_{Ghaut} , ces pièges lents ont plus de chance de capturer un électron et de contribuer ainsi au phénomène de pompage [20]. Dans notre exemple on a $I_{CPmax} = 5.41 \cdot 10^{10} A$, et en utilisant la valeur de sigma du paragraphe précédent, on trouve avec (II.36) :

$$D_{it} = 4.48 \cdot 10^{10} eV^{-1} cm^{-2}$$

L'avantage de ces deux méthodes, est le temps de mesure réduit d'une part par la rapidité d'acquisition et d'autre part par le nombre de points nécessaires plus faible. En effet, la mesure à V_{Gbas} fixé correspond aux étapes 1, 2 et 3 de la Fig. II.16 : le niveau inférieur est maintenu au dessous de V_{fb} alors que le niveau haut parcourt $[V_{fb}; V_t]$. Quand à la mesure à V_{Ghaut} fixé, elle correspond aux étapes 3, 4 et de la Fig. II.16. Une seule des deux courbes suffit à remonter à la valeur moyenne de N_{it} , obtenue par la valeur maximale de I_{CP} .

II.2.3 Evolution au cours d'un stress

a) Analyse de défauts CP standard

De part sa structure, notre banc de mesure permet de commuter automatiquement entre les configurations de stress, de mesure I-V et de mesure CP. L'évolution des courbes CP au cours du temps renseigne sur l'évolution de la quantité de défauts [27]. L'augmentation de la hauteur du plateau indique une variation de la valeur moyenne de D_{it} . Un décalage des courbes vers les tensions négatives indique une diminution du seuil CP ($V_t - \Delta V_A$) et donc de la tension de seuil du transistor indiquant l'apparition de charge positive dans l'oxyde. Inversement, les décalages vers les tensions positives des courbes CP révèlent une charge piégée négative dans l'isolant de Grille. Ces observations sont plus simples dans le cas des dégradations uniformes (cas a de la Fig. II.18). En effet, la dégradation non-uniforme (cas c et d), génèrent une déformation de la courbe CP typique en cloche dû aux effets cumulés de N_{it} et N_{ot} localisés. On peut s'en rendre compte sur les schémas de la figure II.18. Sur le cas c. on observe sur un NMOS la présence de charges positives piégées dans l'oxyde (décalage du front dans les tensions négatives) associé à des états d'interface près du Drain (augmentation de I_{CP}^{max}). Enfin le troisième schéma illustre ces effets sur un PMOS, le décalage des fronts vers les tensions positives est imputable à une charge négative piégée dans l'oxyde, et les états d'interface sont décrit par l'augmentation du maximum I_{CP} .

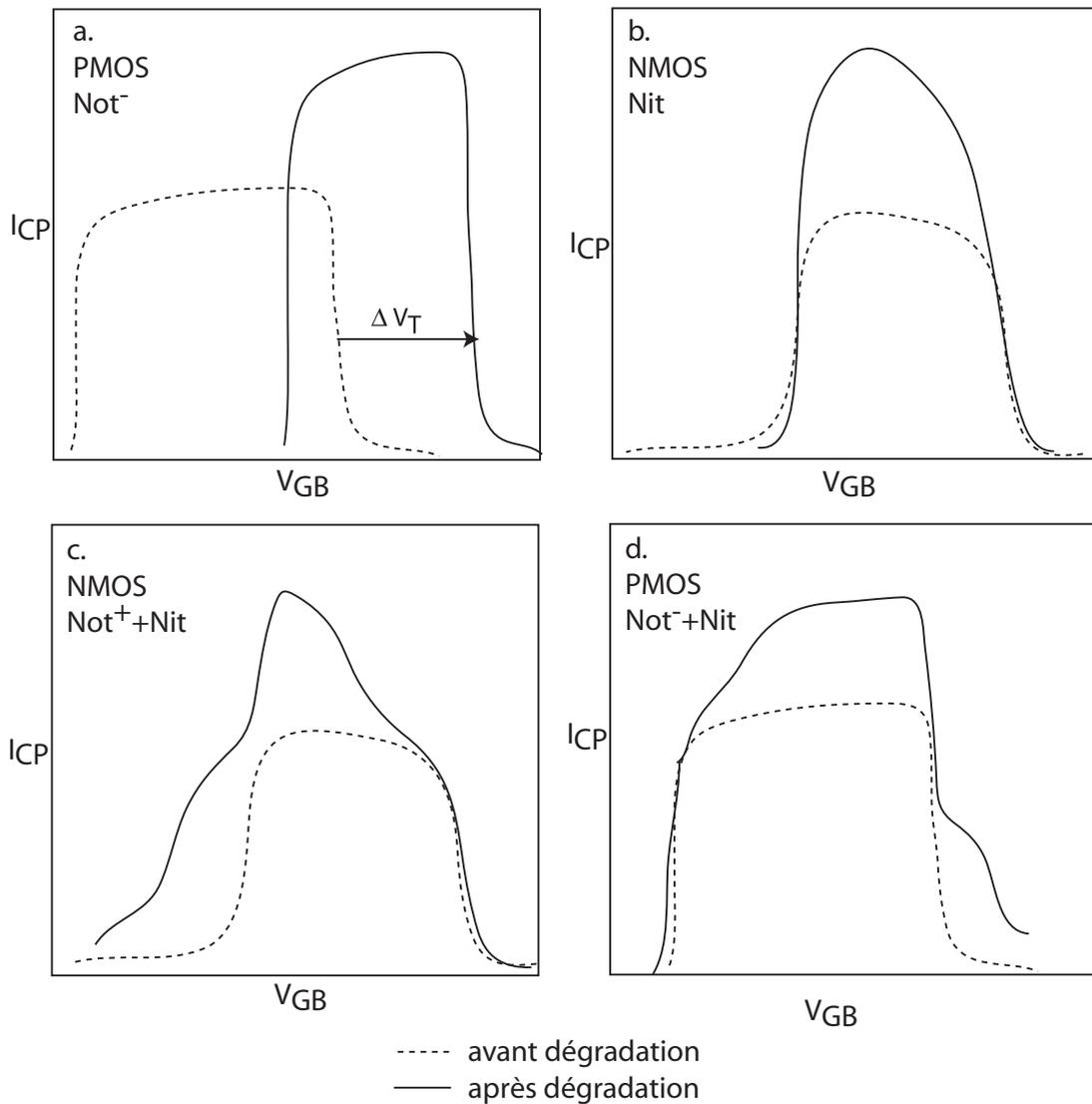


FIG. II.18 – Représentation schématique des cas standards de dégradation. a : charge négative importante piégée dans l'oxyde sur PMOS, b : création d'états d'interface dominante sur NMOS, c : charge positive associée à la création d'états d'interface sur NMOS, d : charge négative associée à la création d'états d'interface sur PMOS [26].

b) Technique CP Différentielle

Il est plus pratique et plus rapide de réaliser un pompage de charge à un niveau fixé pendant les enchaînements de séquences stress-mesure. On peut ainsi suivre l'évolution du courant pompé au cours du temps comme sur l'exemple de la Fig. II.19, effectué en alternance avec des injections d'électrons chauds (phénomène qui sera décrit dans le chapitre 3) sur un PMOS. La première information essentielle est l'augmentation du plateau CP qui traduit directement l'augmentation de la valeur moyenne de la densité d'états d'interface D_{it} . En revanche sur un tel tracé il est peu commode d'obtenir des informations sur la charge piégée dans l'oxyde, puisque l'augmentation du plateau "masque" les décalages du front des courbes.

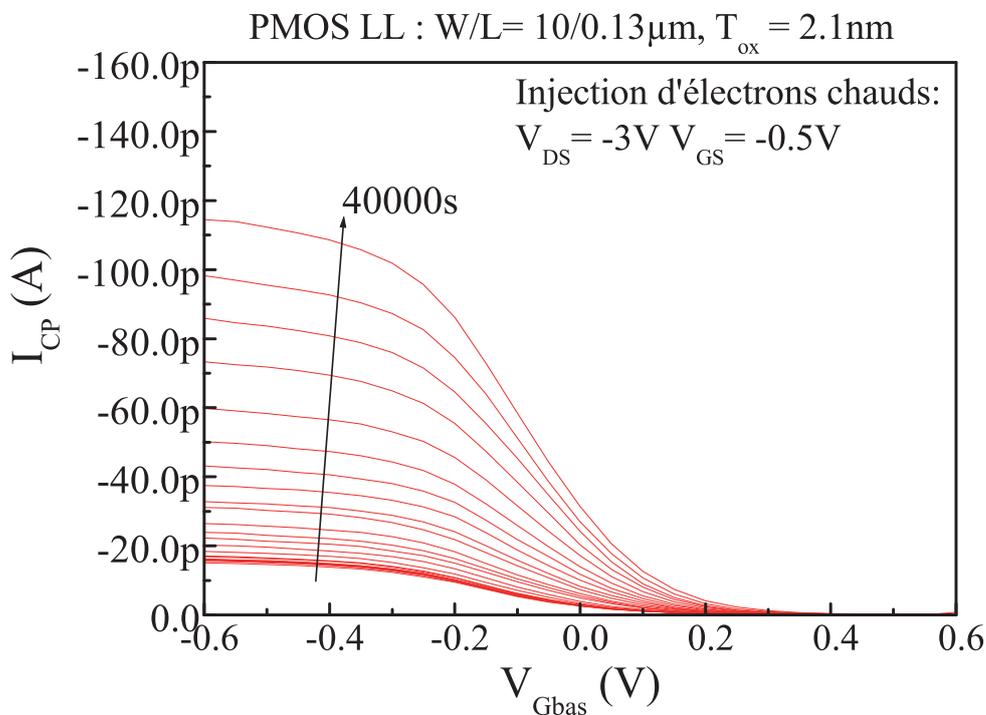


FIG. II.19 – Mesures de pompage de charge successives au cours d'injections d'électrons chauds, le niveau haut est fixé à $V_{Gh} = 0.6V$.

Pour dissocier les deux aspects, on utilise la technique CP différentielle proposée par Chen [28]. Pour cela, on trace la dérivée par rapport à V_{Gl} . Les déplacements des pics obtenus sur la Fig. II.20 donnent directement l'information sur les déplacements latéraux des courbes CP, dans le cas d'un transistor PMOS dégradé à la condition d'injection d'électrons. La distinction peut s'effectuer par rapport à la courbe à $t = 0s$ de $\left(\frac{dI_{CP}}{dV_{Gl}}\right)$ (labels ronds) et en suivant les différence de $\Delta\left(\frac{dI_{CP}}{dV_{Gl}}\right)$ en fonction du temps. Dans cet exemple, on observe la croissance de la charge négative piégée dans l'oxyde ($\Delta V_{CP}^+ = 0.105V$) et cela dès les temps courts de stress. On remarquera dans cette technologie à oxyde de Grille de $2.1nm$, le piégeage plus important aux temps courts et la diminution du piégeage comme première conséquence du dépiégeage.

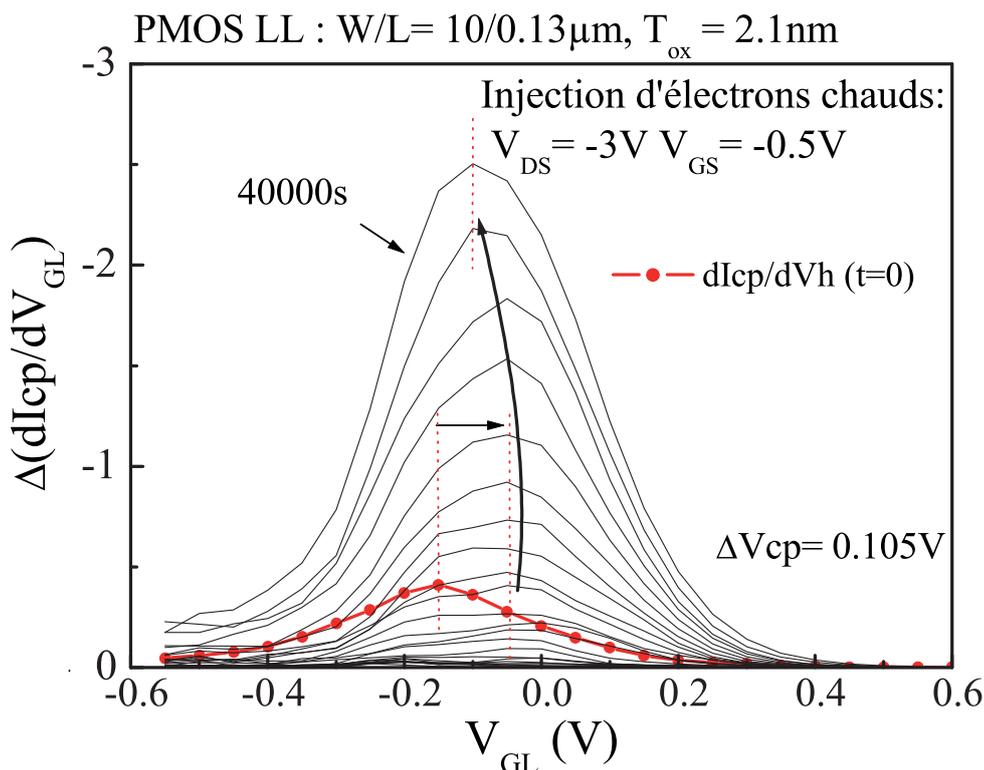


FIG. II.20 – Dérivée par rapport au niveau bas variable du pulse des courbes CP pendant les injections d'électrons chauds pendant $t_s = 40000$ [29].

Ces aspects seront abordés au Chapitre III.

II.2.4 Contraintes et limitations

La technique de pompage de charges s'avère plus délicate à mettre en pratique que les techniques I-V présentées plus haut. En effet les paramètres des pulses à utiliser sont différents pour chaque technologie analysée. La mesure ne doit pas être génératrice de dégradation dans la structure, le signal de Grille doit donc se situer dans l'intervalle $[-V_{DD}; V_{DD}]$, de manière à maintenir le champ dans l'oxyde dans des proportions raisonnables. Ceci revient à respecter les conditions :

$$V_{fb} - \Delta V_A \geq -V_{DD} \quad (\text{II.41})$$

$$\Delta V_A + V_t \leq V_{DD} \quad (\text{II.42})$$

$$|V_{fb} - V_t| \leq \Delta V_A \quad (\text{II.43})$$

Dans le cas des mesures à un niveau fixé, les conditions (II.41) et (II.42) doivent être réalisées à la fin de la mesure, c'est à dire quand ΔV_A est maximal.

Si ces conditions sont relativement faciles à réaliser, les technologies successives ont introduit

une limitation qui parait de plus en plus présente au fur et à mesure que l'oxyde de Grille s'amincit. De part la configuration de mesure, lorsque l'on mesure I_{GB} , on mesure la somme du courant de Grille et du courant pompé. Pour les oxydes épais, ceci ne pose bien évidemment pas de problème particulier. En revanche, si le dioxyde de silicium est le siège de courant tunnel, celui ci peut être du même ordre de grandeur que le courant pompé. Masson [19] a proposé une manière simple de s'affranchir de ce problème en soustrayant la composante tunnel au courant pompé. Cette composante tunnel se calcule en intégrant la composante DC du courant de Grille suivant la valeur du potentiel appliqué :

$$I_{Tunnel} = \frac{1}{T_p} \int_0^{T_p} I_{GB} [V_G(t)] dt \quad (\text{II.44})$$

$I_{GB} [V_G(t)]$ est la soustraction du courant de Grille et du courant substrat obtenus dans les conditions statiques. La Fig. II.21 représente les composantes tunnel entre $V_{GS} = \pm V_{DD}$ pour une technologie $2.1nm$. On remarque suivant le sens de la mesure et l'amplitude choisie, le courant tunnel peut être important en inversion faible (PMOS). D'autre part, on peut noter que le courant substrat reste très faible, à la limite du banc ($10fA$) et augmente à partir de V_{FB} en accumulation, ce qui nous permet d'approximer $I_{GB} \approx I_G$.

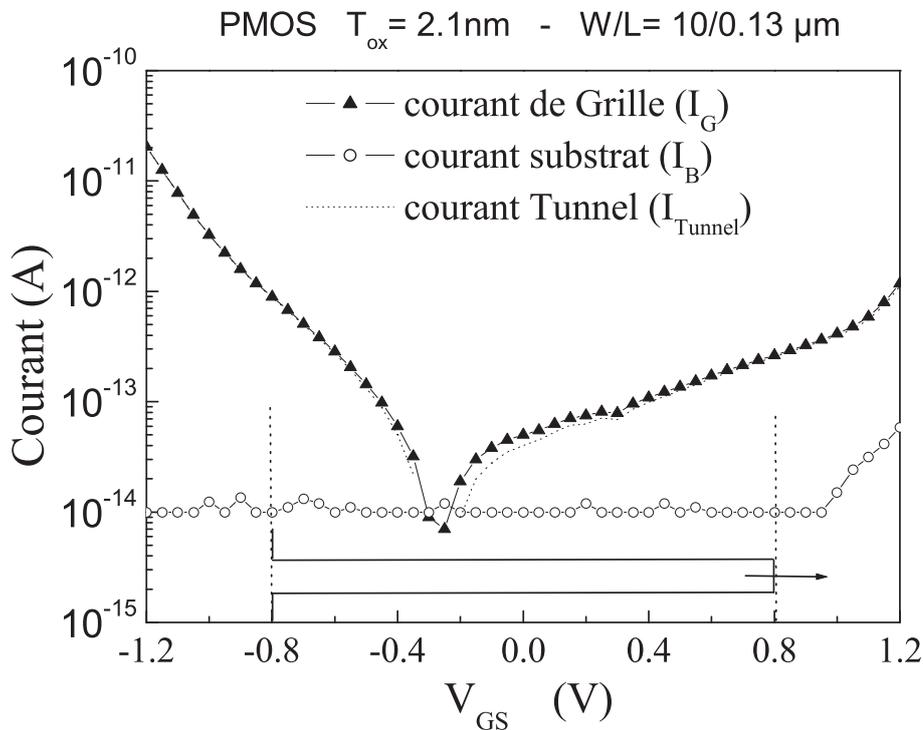


FIG. II.21 – Courant de Grille et de substrat nécessaires au calcul de $I_{GB} = I_G - I_B$

Dans le cas de signaux sinusoïdaux, le signal de Grille s'écrit comme la fonction du temps [19]:

$$V_G(t) = V_{off} + \frac{\Delta V_A}{2} \sin(2\pi f_{CP}) \quad (\text{II.45})$$

où V_{off} est la valeur moyenne du signal et $\Delta V_A = V_{gh} - V_{Gl}$ sa moyenne. En reportant (II.45) dans (II.44) on obtient l'expression donnant la composante DC du courant [19] :

$$I_{Tunnel} = \frac{2}{\pi \Delta V_A} \int_{V_{Gl}}^{V_{Gh}} \frac{I_{GB}(V)}{\sqrt{1 - \left(\frac{V - \Delta V_{off}}{\Delta V_A/2}\right)^2}} dV \quad (\text{II.46})$$

En faisant la soustraction de I_{Tunnel} à I_{CP} on élimine le décrochage à basse fréquence observé sur la Fig. II.23.

Remarque : La Fig. II.14 ne présente pas ce phénomène du fait de l'épaisseur de l'oxyde, qui ne permet le passage d'un courant tunnel direct.

Dans le cadre de notre étude, nous privilégions le pompage de charge à V_{Gh} ou V_{Gl} fixé pour sa facilité de mise en oeuvre. Nous employons des signaux trapézoïdaux de rapport cyclique 1/2, pour lesquels on négligera les temps de montée/descente, le calcul de la composante DC pendant une période se ramène simplement à la moyenne (pour un rapport cyclique de 50%) des niveaux haut et bas au cours de la période CP [20]:

$$I_{Tunnel}(t) = \frac{1}{2} [I_{GB}(V_G(t)) + I_{GB}(V_{G_{fixe}})] \quad (\text{II.47})$$

Les Fig. II.22 et II.23 présentent un exemple de ce que cette technique peut apporter dans le cas où les courbes CP ne saturent pas nettement, empêchant le calcul de D_{it} de manière claire et sure. Pour ce dispositif, le courant tunnel est important à cause de la surface du dispositif ($10\mu m \times 0.4\mu m$). La prise en compte du courant tunnel permet de corriger le plateau de saturation d'environ 4%.

D'autres techniques existent pour pallier l'augmentation croissante des courants de fuites tunnel à travers l'oxyde mais également à travers les "spacer" (au bord de l'oxyde, au dessus des zones de recouvrements). On peut citer notamment les travaux de Chung [30, 31] basés sur une étude en fréquence de la mesure à un des deux niveaux fixé. Comme nous l'avons établi dans la section précédente (II.2.2.a), le courant pompé augmente avec la fréquence du signal de Grille. On effectue la mesure en diminuant la fréquence jusqu'à obtenir la fréquence où seule la composante DC reste mesurable vis-à-vis de la sensibilité du banc de mesure.

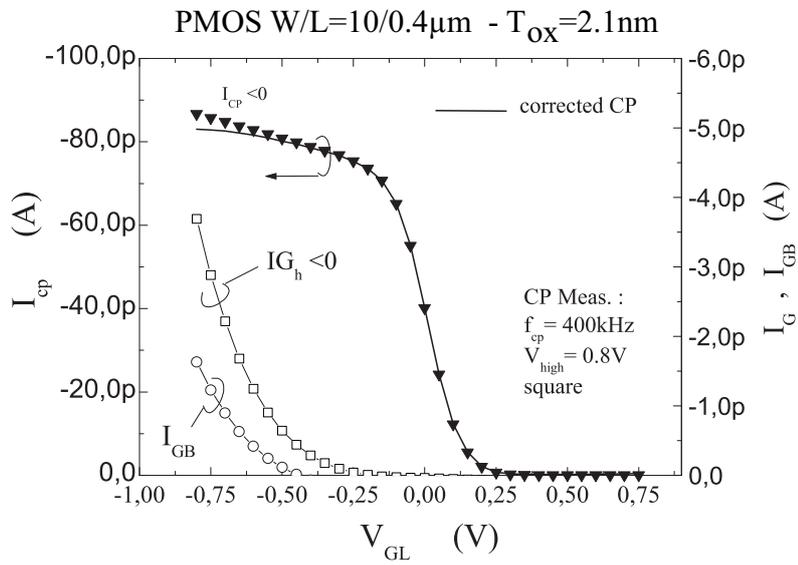


FIG. II.22 – Comparaison des courbes CP – V_{GL} avec les courbes de courant de Grille (I_G) et le courant tunnel (I_{GB}) calculé avec (II.47) [20].

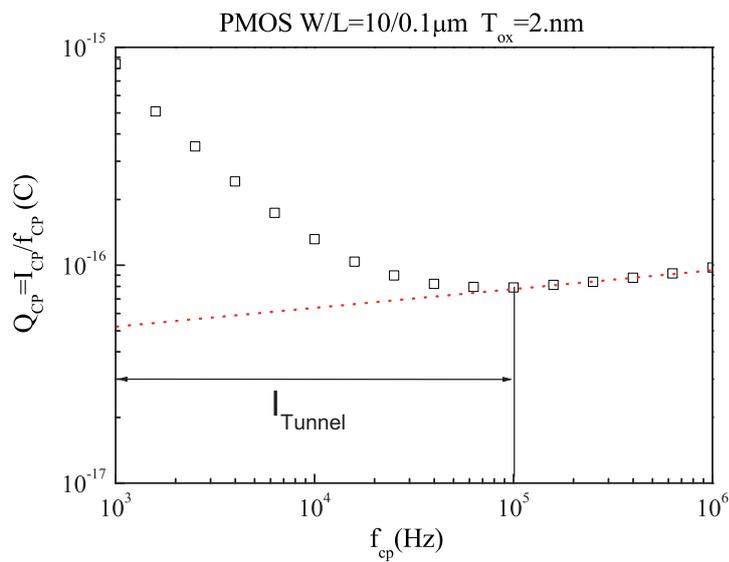


FIG. II.23 – Décrochage typique dû à une forte composante tunnel dans un mesure CP (Signal sinusoïdal, amplitude $\Delta V_A = 2.4V$).

II.3 Autres méthodes d'analyse

Pour compléter ce chapitre sur les méthodes d'analyse, nous allons énoncer quelques techniques supplémentaires proposées dans la littérature. Nous allons succinctement exposer le pompage de charges à trois niveaux, le profilage latéral et la méthode de la grille flottante.

II.3.1 Pompage de charge à trois niveaux

Commençons par le pompage de charge à trois niveaux qui s'inspire directement du pompage de charge standard, mais utilise un signal de Grille plus complexe [17], en intercalant un troisième plateau entre V_{Gh} et V_{Gl} (Fig. II.24). Ce troisième plateau dans le signal de grille

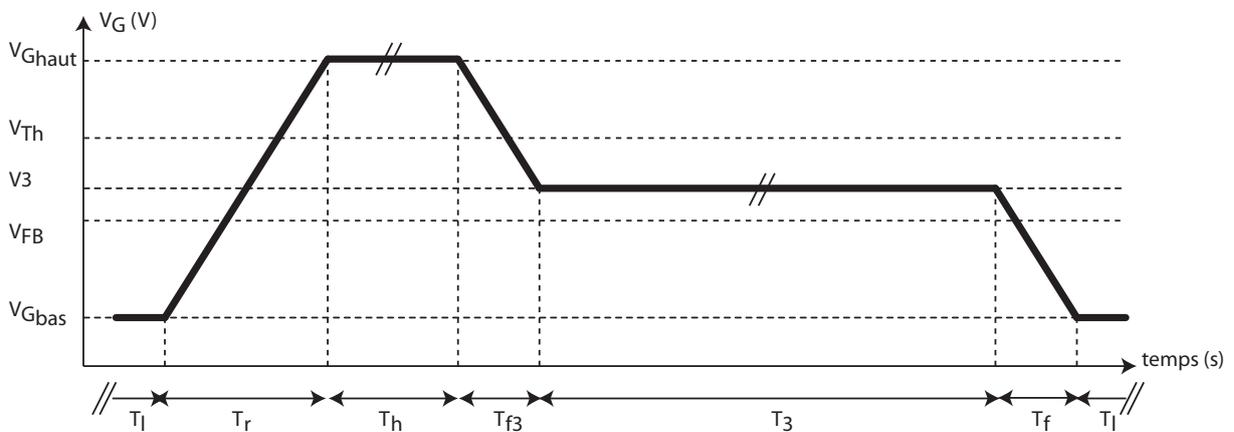


FIG. II.24 – Synopsis du signal de Grille pour le pompage de charge à 3 niveaux sur une période.

établit un équilibre thermodynamique au niveau E_3 qui correspond au potentiels de grille V_3 et de surface ψ_3 , et permet ainsi un pompage sélectif sur le niveau E_3 . Si le temps t_3 du plateau de V_3 est suffisamment long pour atteindre l'équilibre, les niveaux haut et bas étant fixés par ailleurs, il est ainsi possible de sonder les demi-gap inférieur et supérieur. Si la période du signal débute par le niveau haut (bas), les tensions V_3 comprises dans le demi gap supérieur permettent de suivre la capture d'électrons et l'émission d'électrons (capture de trous et émission d'électrons), alors que si V_3 est dans le demi gap inférieur, on observe la capture d'électrons et l'émission des trous (capture de trous et émission de trous) [32]. De même si l'on fixe la valeur V_3 , on peut déterminer les temps de capture ou d'émission en faisant varier la durée du troisième plateau, T_3 . En se rappelant que $E_3 = q(\psi_3 - \phi_F) + E_i$, le courant pompé s'exprime :

$$D_{it}(\psi_3) = (qS_{eff}f_{CP})^{-1} \left| \frac{dI_{CP}}{d\psi_3} \right| \quad (\text{II.48})$$

A l'aide de la relation $\psi - V_G$ (obtenu par mesure CV) on remonte à la densité d'états par niveau d'énergie. De part les faibles valeurs de courant mesurées au substrat, cette technique demande

des dispositifs de grande surface, avec des oxydes relativement épais pour diminuer les courants tunnels à travers l'isolant. Par ailleurs la mesure CP à trois niveaux ne permet pas la distinction entre les charges d'oxydes et la charge piégée sur les états d'interface.

II.3.2 Profilage de la zone de défauts

La technique du profilage peut être effectuée à l'aide des paramètres IV ou CP. Cette dernière permet de dissocier l'étendue latérale des défauts N_{it} et N_{ot} . On peut citer la technique de profilage de Chen et Ma [28] ou d'Ancona et al. [33], basée sur le principe du pompage de charges. Elle permet de donner la localisation des charges stockées dans l'oxyde et sur les états d'interface (Fig. II.25). Considérons un transistor MOSFET à canal n . Dans le cas d'une distribution de défauts non uniforme le long du canal, les paramètres V_t et V_{fb} présente une dépendance spatiale [34] comme représenté sur la Fig. II.25. Ceci se traduit par un courant pompé pouvant s'exprimer comme [35]:

$$I_{CP} = qW_{eff}f_{cp} \int_0^{L_{eff}} N_{it}(x)g[V_{Gh} - V_t(x)]dx \quad (\text{II.49})$$

où $N_{it}(x)$ est la répartition spatiale de la densité d'états d'interface le long du canal et g est donné par :

$$g(y) = \begin{cases} 0, & \text{if } y > 0, \\ 1, & \text{if } y < 0. \end{cases}$$

En dérivant (II.49) par rapport à V_{Gh} , on obtient l'équation différentielle :

$$\frac{dI_{CP}}{dV_{Gh}} = qW_{eff}f_{cp} \int_0^{L_{eff}} N_{it}(x)\delta[V_{Gh} - V_t(x)]dx \quad (\text{II.50})$$

Un signal en opposition de phase au signal de grille est appliqué au Drain ou/et à la Source, il s'en suit la formation de la ZCE dont la longueur ΔL dépend de $V_D = V_{rev}$. Le canal à alors pour longueur $L - \Delta L$, les équation (II.49) et (II.50) s'intègrent sur l'intervalle $[L; L - \Delta L]$, et pour deux valeurs successives de V_{rev} , les variations associées de I_{CP} et ΔI_{CP} sont données par :

$$\Delta I_{CP} = qW_{eff}f_{cp}N_{it}(L')g[V_{Gh} - V_t(L')]\Delta L \quad (\text{II.51})$$

$$\Delta \left(\frac{dI_{CP}}{dV_{Gh}} \right) = qW_{eff}f_{cp}N_{it}(L')\delta[V_{Gh} - V_t(L')]\Delta L \quad (\text{II.52})$$

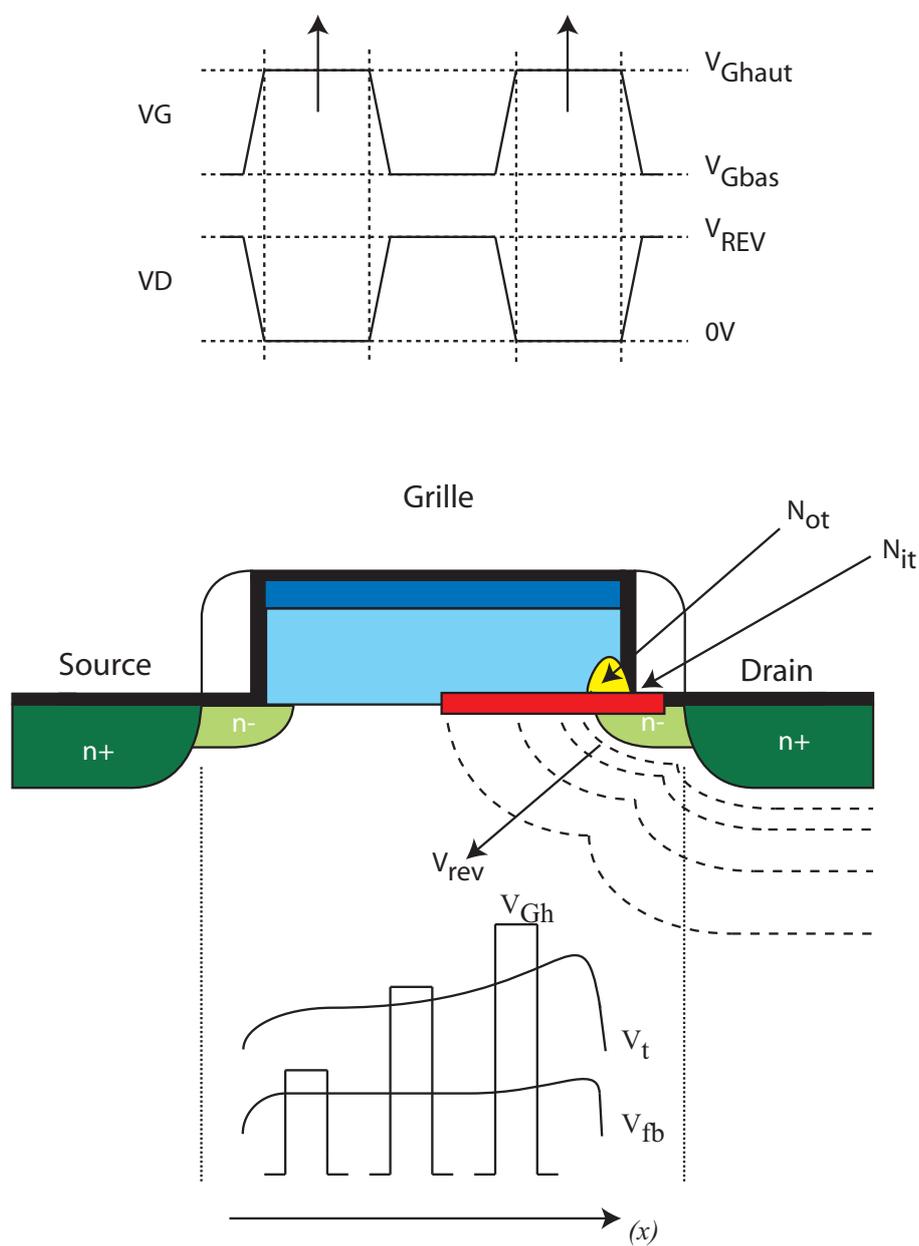


FIG. II.25 – Influence de la tension V_{ref} sur la ZCE au Drain.

où $L' = L - \Delta L/2$, avec ΔL correspondant à la diminution de la longueur effective du canal, imposée par une variation de la tension de Drain ΔV_{rev} . Dans la pratique, la mesure CP donne un faisceau de courbes $I_{CP} - V_{Gh}$ pour chaque V_{rev} , desquelles on déduit les tracés $(dI_{CP}/dV_{Gh}) - V_{Gh}$ pour chaque V_{rev} . On construit finalement $\Delta(dI_{CP}/dV_{Gh}) - V_{Gh}$ en les soustrayant une à une. Les pics de ces courbes indiquent les points pour lesquels $V_t(x) = V_t(L')$. La valeur de $V_t(L')$ est déduite des courbes $(dI_{CP}/dV_{Gh}) - V_{Gh}$, alors que L' peut être obtenu à l'aide de modèles tels que SPICE ou BSIM. Quand aux valeurs de V_{fb} , elles s'obtiennent avec le même procédé sur les courbes $I_{CP} - V_{Gl}$. Lorsque les profils des paramètres CP sont connus, on peut calculer $N_{ot}(x)$, lié à $V_t(x)$ et $N_{it}(x)$ lié à $I_{CPmax}(x)$.

Une limitation de la technique est la relation qui lie ΔL à V_{rev} , pour l'établir, il faut en effet connaître exactement le profil du dopage dans le canal dans le sens latéral $N_a(x)$. Sur la Fig. II.26 on a dessiné le modus operandi de la méthode, notre banc permet de commuter du pompage standard au profilage latéral de manière entièrement automatisée.

II.3.3 Grille flottante

La technique de la Grille flottante [36, 37, 38, 12] permet de caractériser les dommages - pièges de charges dans l'oxyde et états d'interface SiO_2 - générés par des injections de porteurs chauds. Cette méthode d'analyse n'est utilisable que pour des oxydes tels que $T_{ox} > 7nm$. On procède en trois étapes :

1. on charge la Grille en polarisant la structure avec les tension $V_{GS} = V_{ft}$ et V_{DS} (Fig. II.27),
2. la grille est déconnectée pour être flottante ($t = 0s$),
3. on mesure $I_{DS}(t)$.

Le potentiel de la Grille isolée est modifié par les porteurs injectés entraînant un champ dans l'oxyde. La variation de courant de Drain induite ΔI_D pendant de l'intervalle temps Δt engendre le courant de Grille donné par:

$$I_{G,e} = \frac{C_g \Delta I_D}{g_m \Delta t} \quad (II.53)$$

où le g_m est la transconductance et C_g est la capacité de la Grille (y compris la capacité constante du pad destiné à la pointe de mesure). La tension de Grille est directement donnée par la relation:

$$V_G = \frac{I_D}{g_m} + V_t \quad (II.54)$$

Le transconductance g_m et la capacité de Grille sont facilement mesurées à toutes les tensions de Grille et de Drain appliquées sur le dispositif pendant la mesure FG, la tension de seuil étant mesurée par une courbe linéaire $I_D - V_G$. Les mesures sont effectuées sur un transistor MOSFET

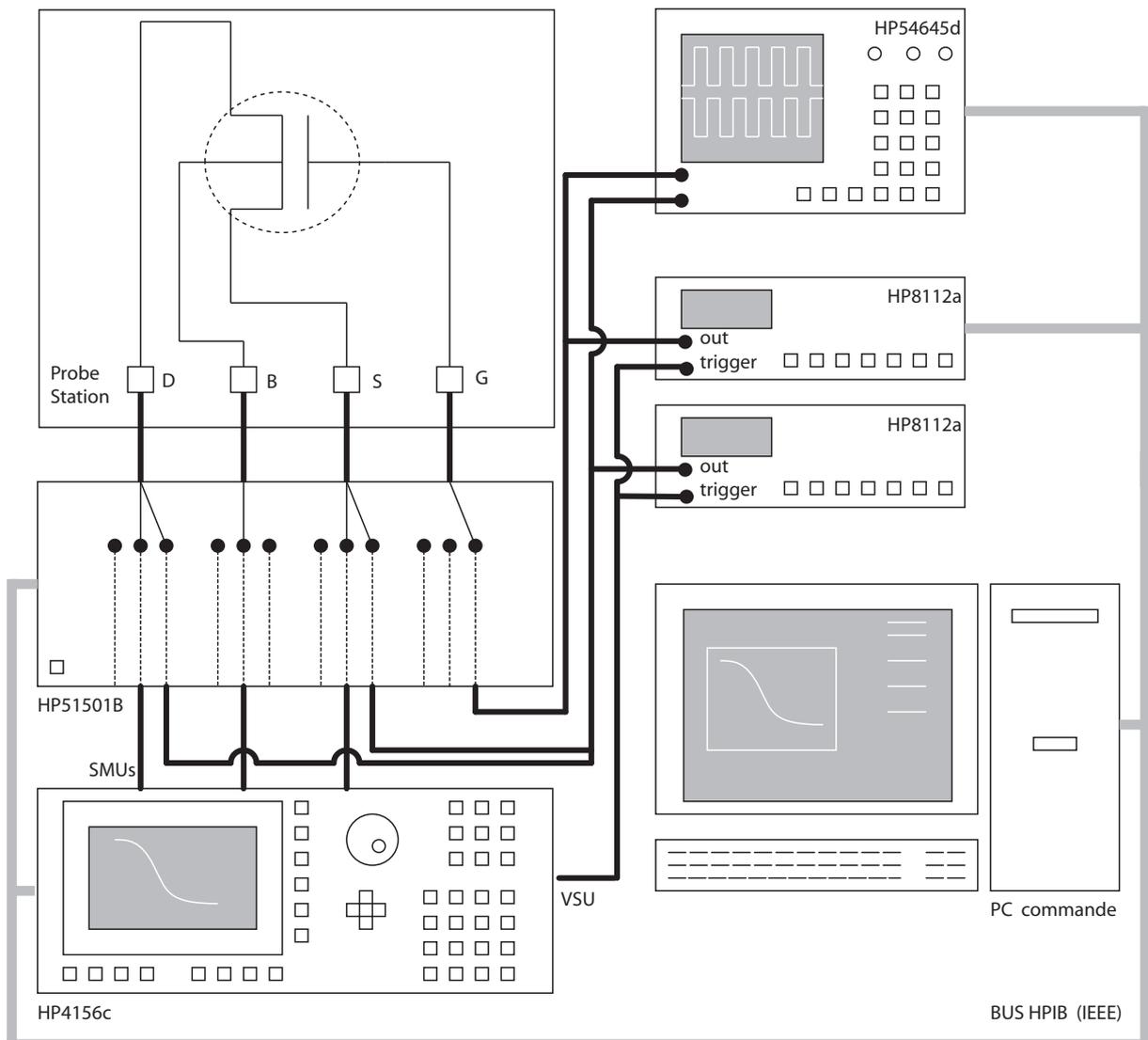


FIG. II.26 – Synopsis du banc pour le profilage, les connexions supplémentaires apportées par rapport à la Fig. II.13 sont en fait implantées de manière permanente sur le banc et sont utilisées à la demande grâce à la configuration de la boîte de commutation.

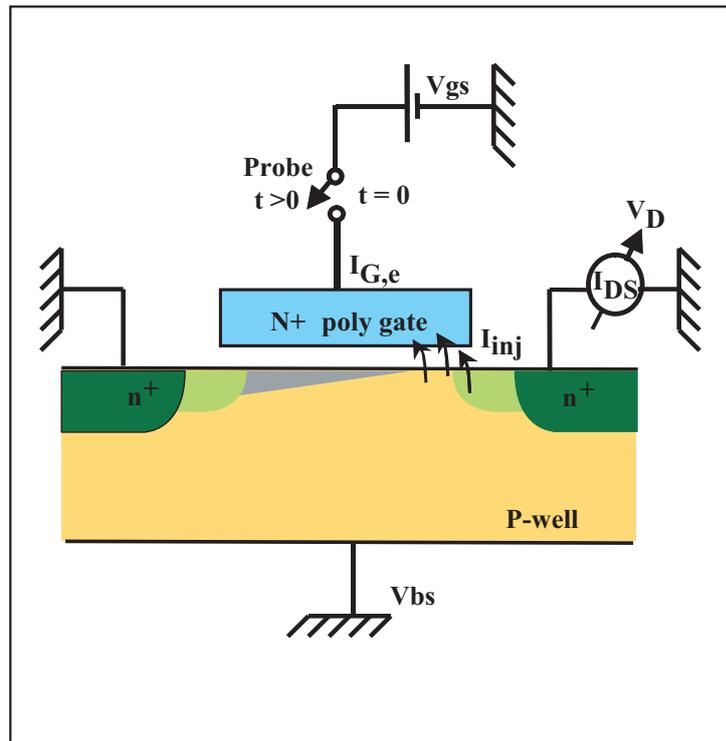


FIG. II.27 – Principe de la mesure par Grille Flottante.

isolé. Considérons un transistor à canal n (l'étude pour le canal p se fait avec les changements habituels de notations et de polarité). Avant la mesure FG elle-même, nous mesurons $I_D - V_G$ en régime linéaire ($V_D = 50mV$) et en régime saturé ($V_D = V_{Dfg}$) pour la tension de drain utilisée pendant la mesure de FG. La tension de Grille est ensuite augmentée pendant un temps fixe (par exemple 30 s), puis la pointe de Grille est relevée (il est préférable de disposer de pointes automatisées). La variation temporelle du courant de drain est mesurée jusqu'à ce qu'elle atteigne une valeur correspondante à une tension de Grille choisie plus basse d'après les caractéristiques saturées $I_D - V_G$. Comme la Grille a été positivement chargée, chaque électron injecté (du fait la tension de Grille élevée) atteignant la Grille diminue la charge stockée dans l'isolant, ce qui correspond à une diminution de la tension de Grille et à une diminution du courant de Drain. À partir de la caractéristique $I_D - t$, on calcule le courant de Grille électronique $I_{G,e}$ à l'aide de la dérivée de la courbe selon II.53 et nous en déduisons la tension de Grille correspondante à partir de $I_D - V_G$ selon II.54.

La technique FG est employée pour suivre l'évolution de $I_{G,e}$ induite par les injections HE. Comme la technique de FG utilise la "source" de dégradation (c'est à dire le mécanisme d'injection de porteurs chauds lui-même) pour sonder les caractéristiques $I_{G,e} - V_G$, cette technique est, en fait, une sonde locale pour mesurer les effets des charges d'oxyde localisés sur le comportement de $I_{G,e}$. Ainsi la technique de FG est une méthode directe pour sonder les défauts localisés autour du point d'injection HE près du drain ou dans la région de recouvrement Drain-

Grille du transistor MOSFET. La mise en évidence de la création de pièges d'oxyde de type accepteur se déduit de l'étude de l'évolution de $I_{G,e}$ pendant le vieillissement mesuré par la technique de FG [38]. D'après le modèle de l'électron chanceux [1], les paramètres qui peuvent notablement modifier $I_{G,e}$ sont :

- le champ électrique latéral maximum ξ_m
- la hauteur de barrière effective que les électrons doivent surmonter pour être injectés dans l'oxyde et pour être collectés à la Grille.

Dans ce modèle, une expression simple du courant de Grille est donnée par la relation suivante :

$$I_{G,e} = 0.5 \frac{I_{DS} T_{ox}}{\lambda_r} \left(\frac{\lambda \xi_m}{\Phi_b} \right)^2 P(F_{OX})_L \exp^{-\Phi_b / \xi_m \lambda} \quad (\text{II.55})$$

où λ_r et λ sont, respectivement, les libres parcours moyens dans le cas de collisions élastiques et inélastiques, Φ_b est la hauteur de la barrière et $P(F_{OX})_L$ est la probabilité pour un électron d'atteindre la Grille sans subir de collision. Cette probabilité dépend du champ dans l'oxyde F_{OX} mesuré au Drain. Dans cette équation, $I_{G,e}$ diminue avec l'augmentation de Φ_b ou la diminution de ξ_m . Cette augmentation est due à une génération de charges négatives dans l'oxyde, qui induit une augmentation du potentiel de barrière le long du canal [12]. Cette augmentation de ξ_m devrait induire une augmentation de $I_{G,e}$. Dans le cas contraire, on peut conclure que la diminution observée de $I_{G,e}$ est dominée par une augmentation de la hauteur de barrière effective induite par ces charges négatives d'oxyde.

Cette méthode a notamment permis de mettre en évidence le rôle et la nature des pièges d'oxyde dans les structures LDD, au niveau des "spacer". Il a ainsi été montré que les pièges d'oxyde dans cette génération LDD de n-MOSFET sont de type accepteurs et deviennent électriquement actifs après injections d'électrons. [12]. Cette méthode, combinée aux mesures CP donne une caractérisation fine de la dégradation des structures CMOS.

Conclusion

Dans ce chapitre nous avons étudié les principales techniques qui permettent d'analyser le niveau de dégradation atteint par les dispositifs MOSFETs après une contrainte électrique. Ces méthodes ont été abordées indépendamment des mécanismes de vieillissement mis en jeu. Ainsi nous avons vu un ensemble de méthodes dites courant-tension qui consistent à suivre au cours du stress les variations des paramètres électriques statiques du composant, afin de déduire le type de dommage occasionné. Dans la deuxième partie de ce chapitre, la physique du pompage de charge a été décrite. Enfin les techniques d'extraction des paramètres (D_{it}, N_{it}, N_{ot}) relatifs à la quantification de la dégradation ont été exposées. L'ensemble de ces outils étant maintenant connu nous allons pouvoir nous intéresser aux mécanismes de dégradation eux-mêmes.

Références du Chapitre II

- [1] P. J. Tsang, S. Ogura, W. W. Walker, J. F. Shepard, and D. L. Crithlow, “Fabrication of high-performance LDDFET’s with oxide sidewall-spacer technology,” *IEEE Trans. Electron Devices*, vol. ED-29, p. 4. **II**
- [2] T. Hori, J. Hirase, Y. Odake, and T. Yasui, “Deep submicrometer large angle tilt implanted drain (latid) technology,” *IEEE Trans. Electron Devices*, vol. 39, no. 10, p. 2312, 1992. **II**
- [3] T. Liou, C. S. Teng, and R. B. Merrill, “Hot-electron-induced degradation of conventional, minimum overlap, ldd and ddd n-channel mosfets,” *IEEE Circuits and devices Mag.*, pp. 9–15, March 1988. **II**
- [4] H. Wakabayashi, M. Ueki, M. Narihiro, T. Fukai, N. Ikezawa, T. Matsuda, K. Yushida, K. Takeuchi, Y. Ochiai, T. Mogami, and T. Kunio, “Sub-50-nm physical gate length CMOS technology and beyond using steep halo,” *IEEE Trans. Electron Devices*, vol. 49, no. 1, pp. 89–94, 2002. **II**
- [5] C. H. Wann, K. Noda, T. Tanaka, M. YOSHIDA, and C. Hu, “A comparative study of advanced MOSFET concept,” *IEEE Trans. Electron Devices*, vol. 43, no. 10, p. 1742, 1996. **II**
- [6] J.Chung, M.-C. Jeng, G. May, P. Ko, and C.Hu, “Hot-electron currents in deep-submicrometer MOSFET’s,” in *IEDM’ 88 Tech. Dig.*, pp. 200, 1988. **II**
- [7] E. Takeda, H. Kume, T. Toyabe, and S. Asa, “Submicrometer MOSFET structure for minimizing hot-carrier degradation,” *IEEE J. Solid-State Circuits*, vol. 17, no. 2, p. 241, 1982. **II**
- [8] B. Doyle, M. Bourcerie, J.-C. Marchetaux, and A.Boudou, “Interface state creation and charge trapping in the medium-to-high gate voltage range ($vd/2 \leq vg \leq vd$) during hot-carrier stressing of n-MOS transistors,” *IEEE Trans. Electron Devices*, vol. 37, no. 3, p. 744, 1990. **1, b), III.3, III.3.1**
- [9] A. Bravaix, D. Goguenheim, N. Revil, and E. Vincent, “Hot-carrier damage in deep-submicrometer CMOS technologies,” in *International Reliability Workshop, 1999*. **1, 2**
- [10] M. Barron, “Low level currents in insulated gate field effect transistor,” *Solid State Electronics*, vol. 15, no. 2, p. 293, 1972. **II.1.4**

- [11] R. Overstraeten, G. J. Declerck, and P. Muls, "Theory of the MOS transistor in weak inversion new method to determine the number of surface states," *IEEE Trans. Electron Devices*, vol. 22, no. 5, p. 282, 1975. [II.1.4](#), [II.1.4](#), [II.1.4](#)
- [12] D. Vuillaume, J. Marchetaux, P.-E. Lippens, A. Bravaix, and A. Boudou, "A coupled study by floating-gate and charge-pumping techniques of hot-carrier-induced defects in submicrometer LDD n-MOSFET's," *IEEE Trans. Electron Devices*, vol. 40, no. 4, p. 773, 1993. [2](#), [II.3.3](#), [II.3.3](#)
- [13] S. Brugler and P. Jaspers, "Charge pumping in MOS devices," *IEEE Trans. Electron Devices*, vol. 16, no. 3, pp. 297–302, 1969. [II.2, c \)](#), [2](#)
- [14] J. Simmons and L. Wei, "Theory of dynamic charge and capacitance characteristics in mis systems containing discrete surface traps (a)," *Solid-State Electronics*, vol. 16, pp. 43–52, 1973. [II.2, c \)](#), [c \)](#), [c \)](#)
- [15] ———, "Theory of dynamic charge and capacitance characteristics in mis systems containing discrete surface traps (b)," *Solid-State Electronics*, vol. 16, p. 52, 1973. [II.2, c \)](#), [c \)](#), [c \)](#)
- [16] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. D. Keersmaecker, "A reliable approach to charge-pumping measurements in MOS transistors," *IEEE Trans. Electron Devices*, vol. 31, no. 1, p. 42, 1984. [a \)](#), [II.2, c \)](#), [c \)](#), [c \)](#), [II.12, c \)](#)
- [17] N. S. Saks and M. G. Ancona, "Determination of trap capture cross section using three level charge pumping," *IEEE Electron Device Lett.*, vol. 11, no. 8, 1990. [II.2](#), [II.3.1](#)
- [18] J. L. Autran and C. Chabrerie, "Use of the charge pumping technique with a sinusoidal gate waveform," *Solid-State Electronics*, vol. 39, pp. 1394–1395, 1996. [II.2, c \)](#), [c \)](#)
- [19] P. Masson, J. Autran, and J. Brini, "On the tunneling component of charge pumping current in ultrathin gate oxide MOSFET's," *IEEE Electron Device Lett.*, vol. 20, no. 2, February 1999. [II.2](#), [II.2.4](#), [II.2.4](#), [II.2.4](#)
- [20] A. Bravaix, D. Goguenheim, N. Revil, and E. Vincent, "Comparison of low leakage and high speed deep submicron pmosfet's submitted to hole injection," in *International Reliability Workshop '02*, 2002, 2002. [II.2, c \)](#), [II.2.4](#), [II.2.2](#)
- [21] W. Shockley and W. Read, "Statistics of the recombination of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 62–69, 1952. [b \)](#)
- [22] P. Masson, "Etude par pompage de charge et par mesures de bruit basse fréquence de transistors mos a oxynitruure de grille ultra-mince," *Thèse de doctorat, Lyon*, 13 janvier 1999. [c \)](#)
- [23] P. Heremans, J. Witters, G. Groeseneken, and H. Maes, "Analysis of the charge pumping technique and its application for the evaluation of MOSFET degradation," *IEEE Trans. Electron Devices*, vol. 36, no. 7, p. 1318, 1989. [c \)](#)
- [24] A. Elliot, "The use of charge pumping currents to measure surface state densities in MOS transistors," *Solid-State Electronics*, vol. 19, pp. 241–247, 1976. [1](#)

-
- [25] M. Declercq and P. Jespers, in *Analysis of interface properties in MOS transistors by means of charge pumping measurements*, vol. 9, pp. 244–253, 1974. [c](#))
- [26] C. T. Wang, *Hot Carrier Design Considerations for MOS Devices and Circuits*, V. N. Reinhold, Ed., New York, 1992. [II.18](#)
- [27] H. M. G. Groeseneken, “Basics and applications of charge pumping in submicron MOS-FETs,” *Microelectronic Reliability*, no. 38, pp. 1379–1389, 1998. [a](#))
- [28] W. Chen, A. Balasinski, and T.-P. Ma, “Lateral profiling of oxide charge and interface traps near MOSFET junctions,” *IEEE Trans. Electron Devices*, vol. 40, p. 187, 1993. [b](#)), [II.3.2](#)
- [29] A. Bravaix, L. Gauthé, and D. Goguenheim, “Comparison of the carrier injection mechanisms in low leakage and high speed HC9L7 PMOSFET’s,” *Confidential STMicroelectronics Crolles report*, 2002. [II.20](#)
- [30] S. S. Chung and al, “A novel direct determination of the interface traps in sub 100nm CMOS devices with direct tunneling regime 12-16Å gate oxide,” in *Symposium on VLSI Technology*, pp. 74, 2002. [I.4.5](#), [II.2.4](#)
- [31] S. Chung, “Low leakage reliability characterisation methodology for advanced CMOS with gate oxide in 1nm range (invited),” in *IEEE IEDM ’ 04*, pp. 477, 2004. [II.2.4](#)
- [32] M. G. Ancona and S. Saks, “Numerical simulation of 3-level charge pumping,” *J. Appl. phys*, vol. 71, no. 9, pp. 4415–4421, 1992. [II.3.1](#)
- [33] M. G. Ancona, N. Saks, and D. M. Carthy, “Lateral distribution of hot-carrier-induced interface traps in MOSFET’s,” *IEEE Trans. Electron Devices*, vol. 35, no. 12, p. 2221, 1988. [II.3.2](#)
- [34] M. Tsuchiaki, H. Hara, , T. Morimoto, and H. Iwai, “A new charge pumping method for determining the spatial distribution of hot carrier induced fixed charge in p-MOSFET’s,” *IEEE Trans. Electron Devices*, vol. 40, no. 10, pp. 1768–1779, October 1993. [II.3.2](#)
- [35] W. Chen and T.-P. Ma, “A new technique for measuring lateral distribution of charge oxide charge and interface traps near MOSFET junctions,” *IEEE Trans. Electron Devices*, vol. 12, no. 7, pp. 393–395, 1991. [II.3.2](#)
- [36] N.Saks, P. Heremans, L. Hove, H. Maes, R. D. Keersmaecker, and G. Declerck, “Observation of hot-hole injection in NMOS transistors using a modified floating-gate technique,” *IEEE Trans. Electron Devices*, vol. 33, no. 10, p. 1529, 1986. [II.3.3](#)
- [37] F. H. Gaensslen and M. Aitke, “Sensitive technique for measuring small MOS gate currents,” *IEEE Trans. Electron Devices*, vol. EDL-1, p. 231, 1980. [II.3.3](#)
- [38] J. Marchetaux, M. Bourcerie, A. Boudou, and D. Vuillaume, “Application of the floating-gate technique to the study of the n-MOSFET gate-current evolution due to hot-carrier aging,” *IEEE Electron Device Lett.*, vol. 11, no. 9, p. 406, 1990. [II.3.3](#), [II.3.3](#)
-

- [39] S. Tam, P. K. Ko, and C. Hu, "Lucky electron model of hot electron injections in MOS-FET's," *IEEE Trans. Electron Devices*, vol. ED-31, p. 1116, 1994. [II.3.3](#), [III.1.1](#), [III.1.1](#), [III.1.1](#)

CHAPITRE III

Mécanismes de vieillissement des transistors MOSFETs soumis aux injections de porteurs chauds

Introduction

Dans les deux précédents chapitres nous avons étudié le principe de fonctionnement des structures MOS, les méthodes de caractérisations électriques des transistors basées sur cette technologie, ainsi que les effets parasites liés à l'évolution de leur procédé de fabrication. Cette évolution, dictée par la recherche du rendement et des performances maximales ne devant pas se faire au détriment de la fiabilité, les études du vieillissement des technologies CMOS restent indispensables et complémentaires à leur caractérisation électrique. Dans ce chapitre nous décrirons dans un premier temps les mécanismes de génération et d'injection de porteurs chauds dans les transistors NMOS et PMOS. Seront ensuite abordés les différents mécanismes et leur impact suivant les générations CMOS en terme de réduction d'épaisseur d'oxyde de Grille. Enfin des expérimentations de stress alternés permettront de discuter des rôles respectifs des injections d'électrons et de trous dans la génération des défauts impliqués dans le vieillissement des transistors MOSFETs, et cela dans la gamme d'épaisseur d'oxyde de Grille $12nm - 2.1nm$.

III.1 Le régime porteurs chauds : généralités

On appelle communément porteurs chauds, tout porteur de charge ayant acquis une énergie cinétique importante, c'est à dire ayant été fortement accéléré sous l'effet d'un champ électrique intense. Dans un cristal, ces porteurs énergétiques entraînent la génération de paires électron-trou lors de collisions inélastiques (avec transfert d'énergie) avec les atomes du réseau. Suivant la nature des porteurs (trous ou électrons), les conséquences ne seront pas les mêmes suivant la distribution des champs électriques. Ce mécanisme d'injection est utilisé dans certains dispositifs comme les mémoires non-volatiles (Flash EEPROM). Pour les transistors MOSFETs, ce phénomène est indésirable et peut entraîner des défaillances de fonctionnement des circuits.

III.1.1 Injection d'électrons chauds

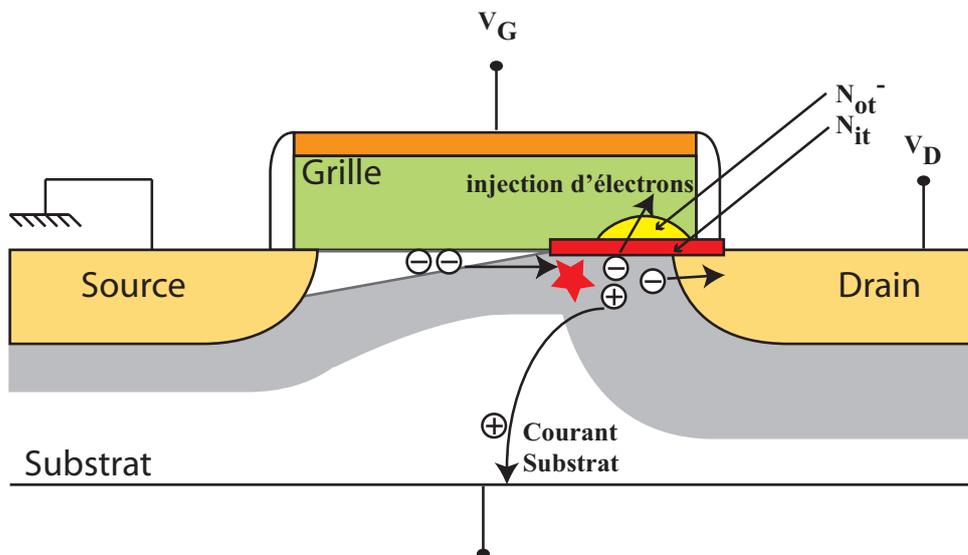


FIG. III.1 – Génération et injection d'électrons chauds dans un transistor MOSFET à canal de type n .

Le mécanisme d'injection de porteurs chauds dans un transistor MOSFET à canal n (Fig. III.1) est typiquement décrit par le modèle de l'électron chanceux ("lucky electron") [1, 2]. Quand le dispositif est polarisé en mode de saturation, des électrons sont injectés dans la région de désertion de la jonction Drain-Substrat. Si le champ électrique est suffisamment grand dans cette région, quelques électrons peuvent acquérir assez d'énergie pour causer l'ionisation par impact lors de chocs avec les atomes de silicium. Il y a alors génération de paires d'électron-trou. Les électrons responsables de ce mécanisme sont appelés électrons chauds (CHE) [3]. Le maximum du champ électrique est situé entre le point de pincement du canal et la jonction métallurgique Drain-Substrat. Certains électrons issus des paires électrons/trous produites dans la ZCE peuvent être réorientés vers l'oxyde de Grille (changement de leur moment et vecteur

d'onde). Dans un même temps, les trous sont repoussés dans le substrat générant un courant de Substrat, I_{BS} , comme illustré sur la Fig. III.1. La mesure de I_{BS} est sans doute le meilleur moyen de modéliser l'intensité de génération des porteurs chauds. En effet le courant s'exprime par [4, 1] :

$$I_{BS} = C_1 I_{DS} \exp\left(-\frac{\phi_{i,e}}{q\lambda_e\xi_m}\right) \quad (\text{III.1})$$

où $\phi_{i,e}$ est l'énergie minimale pour qu'un électron puisse provoquer l'ionisation par impact, λ_e leur libre parcours moyen et C_1 une constante qui a été évaluée proche de 2 dans [2]. C_1 dépend faiblement du champ électrique latéral ξ_m mais est fortement liée à la technologie considérée : on a vu dans (I.112) que C_1 est le rapport des coefficients d'ionisation α_i et β_i . Une valeur plus élevée de I_{BS} correspond à un taux d'ionisation par impact plus élevé. Les électrons chauds injectés ont une énergie $E > 3.2eV$, qui est la hauteur de la barrière de potentiel entre le silicium et le SiO_2 pour les électrons. Cette valeur de $\phi_{b,e}$ est obtenue sans abaissement dû à l'application des potentiels, c'est à dire pour $F_{OX} = 0V/cm$. En présence d'un champ vertical, la hauteur de barrière vue par les électrons est fortement dépendante de F_{OX} [18] :

$$\phi_{b,e} = 3.2 - \beta\sqrt{F_{OX}} - \vartheta F_{OX}^{2/3} \quad (\text{III.2})$$

où $\beta = 2.59 \times 10^{-4}(Vcm)^{1/2}$ pour du SiO_2 et ϑ doit être ajusté expérimentalement autour de $10^5 V^{1/3}.cm^{2/3}$ [1]. Cet abaissement de la hauteur de la barrière est représenté sur la Fig. III.2. Les électrons peuvent traverser la barrière et être emprisonnés dans l'oxyde ou engendrer un courant de Grille [2],[6]. Il est également possible pour les électrons chauds de créer des pièges d'interface, ou états d'interface, à l'interface $Si - SiO_2$, si leur énergie dépasse $3.5eV$. Cette énergie correspond à la somme de la hauteur de barrière $\phi_{b,e} = 3.2eV$ [1] et l'énergie de liaison $\equiv Si_3H$ [7]. Ce type de liaison représente des sites préférentiels de l'interface pour la génération des états d'interface, liée à la dépassivation des liaisons Si-H (l'espèce H est introduite lors de l'étape d'oxydation). Cette valeur de $3.5eV$ est à rapprocher des $3.7eV$ trouvés expérimentalement par Hu [2]. Le taux de génération d'interface s'exprime par :

$$\Delta N_{it} = C_2 \left[t \frac{I_{DS}}{W_{eff}} \exp\left(-\frac{\phi_{it}}{q\lambda_e\xi_m}\right) \right]^n \quad (\text{III.3})$$

Dans cette expression, C_4 et n sont des constantes. La dégradation par porteurs chauds dans des transistors MOSFETs à canal n résulte de la génération d'états d'interface et de l'apparition de charges fixes. Lors d'injections de porteurs chauds on observe [2] que:

1. la tension de seuil augmente avec le temps lorsque le dispositif est polarisé en mode linéaire, ce qui a pour conséquence une diminution de I_{on} ;
2. la mobilité des porteurs du canal diminue avec le temps, ce qui a également pour conséquence la réduction de I_{on} ;
3. la pente sous seuil varie,
4. la transconductance est modifiée de façon significative.

Basé sur le modèle proposé par Hu *et al.* [2], l'augmentation de la tension de seuil liée à la génération des pièges d'interface s'explique par la réduction de la densité et de la mobilité des porteurs dans la zone du Drain. Les décalages de tension de seuil et de transconductance sont proportionnels à la densité moyenne de pièges, qui est inversement proportionnelle à L_{eff} [2][8]. L'augmentation de la tension de Drain V_{DS} entraîne également un taux de génération supérieur puisque le champ électrique dans la zone désertée est plus fort. Considérons la Fig. III.1 où sont décrits les mouvements de charges lors du phénomène d'ionisation par impact. Si le champ électrique est suffisamment intense, certains électrons de la bande de conduction peuvent acquérir une énergie telle que leur impact sur un atome du réseau cristallin aboutisse à la rupture d'une liaison de valence. On obtient donc deux électrons dans la bande de conduction et un trou dans la bande de valence. Ce processus peut devenir cumulatif et conduire au phénomène d'avalanche suivant l'augmentation du champ latéral ($\propto V_{DS}$). Les trous générés par l'ionisation peuvent emprunter différents chemins:

- Ils peuvent être collectés par l'électrode de substrat et donner suite à un important courant Substrat
- Certains d'entre eux peuvent également migrer vers la Source et créer un abaissement de la barrière à la jonction Source-Substrat. Il se produit alors une injection d'électrons supplémentaires de la source vers le canal. Cet ensemble Source-canal-Drain travaille comme un transistor bipolaire $n-p-n$ dont la base (canal) est flottante et le collecteur (Drain) se trouve dans la condition d'avalanche. L'accroissement du nombre d'électrons injectés de la Source vers le Drain provoque une augmentation du nombre de paires électrons-trous dans la zone de déplétion à proximité du Drain. Ceci implique un abaissement de la barrière à la jonction Source-Substrat encore plus prononcé et donc encore plus d'électrons injectés à partir de la Source. Ce phénomène est d'autant plus important que la longueur de canal (=base) diminue (effet transistor).

Lorsque les électrons sont accélérés au niveau de la jonction canal-Drain, une tension moyennement positive appliquée à la grille donne la possibilité à ceux qui ont acquis une énergie potentielle suffisante de franchir la barrière de potentiel, de hauteur $\phi_{b,e}$, existant à l'interface

silicium-oxyde de silicium. L'énergie de ces électrons étant supérieure à la barrière de potentiel, la conduction est dite de type thermoïonique comme illustré sur la Fig. III.2 et le courant résultant s'exprime sous la forme de l'équation de Richardson-Schottky [9]:

$$J_{therm} = \frac{4\pi q m_{ox}^*}{h^3} (kT)^2 \exp\left(\frac{-\phi_{b,e}}{kT}\right) \exp\left(\frac{1}{kT} \sqrt{\frac{qE_{ox}}{4\pi\epsilon_{ox}}}\right) \quad (\text{III.4})$$

où h est la constante de Planck, k est la constante de Boltzmann et m_{OX}^* est la masse effective des électrons dans l'oxyde.

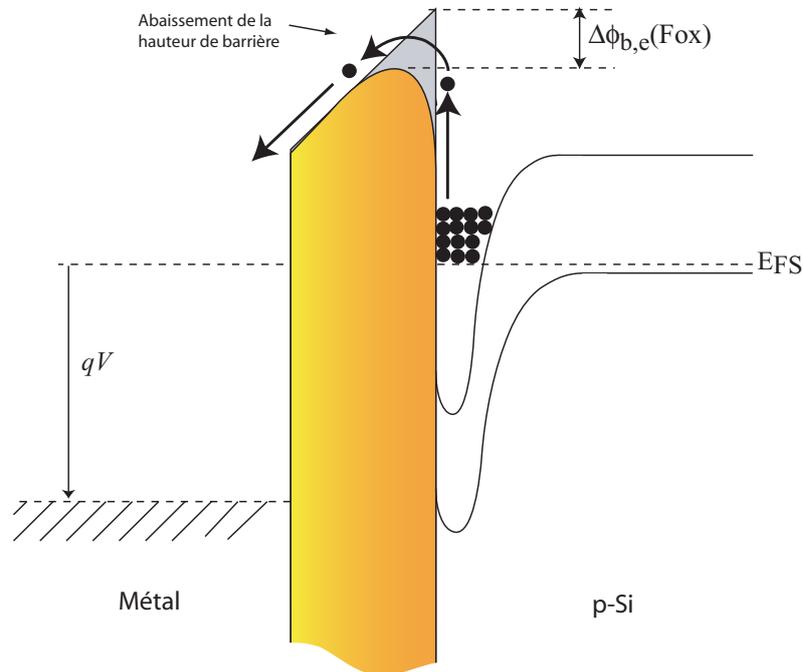


FIG. III.2 – Illustration de l'injection Richardson-Schottky et de l'abaissement de la hauteur de barrière (force image), pour les électrons dans la zone de charge d'espace proches du Drain.

Le nombre d'électrons qui atteignent la Grille peut aussi être évalué par la probabilité P_{inj} pour les électrons du canal d'acquies l'énergie nécessaire pour franchir la barrière :

$$P_{inj} \propto \exp\left(-\frac{\phi_{b,e}}{q\lambda_e\xi_m}\right) \quad (\text{III.5})$$

où $\frac{\phi_{b,e}}{q\lambda_e\xi_m}$ est une quantité représentant le fait que l'électron doit être accéléré par le champ ξ_m sur la distance λ_e pour traverser la barrière. Si l'on considère que I_{DS} représente la quantité d'électrons disponible, on peut exprimer le courant de Grille par [2] :

$$I_{GS} = C_3 I_{DS} \exp\left(-\frac{\phi_{b,e}}{q\lambda_e\xi_m}\right) \quad (\text{III.6})$$

où C_3 a été évaluée théoriquement à 2×10^{-3} [1]. En combinant les équation (III.1) et (III.6) on peut déterminer le rapport entre la hauteur de barrière et le coefficient d'ionisation, permettant ainsi d'évaluer l'abaissement de la hauteur de barrière avec F_{OX} :

$$\frac{I_{GS}}{I_{DS}} = C_3 \left[\frac{I_{BS}}{C_1 I_{DS}} \right]^{\phi_{b,e}/\phi_{i,e}} \quad (\text{III.7})$$

Nous reviendrons plus en détail sur ces expressions dans le chapitre IV, à l'occasion de l'étude des relations donnant la durée de vie des dispositifs MOSFET en fonction des courants de porteurs chauds.

III.1.2 Injection de trous chauds

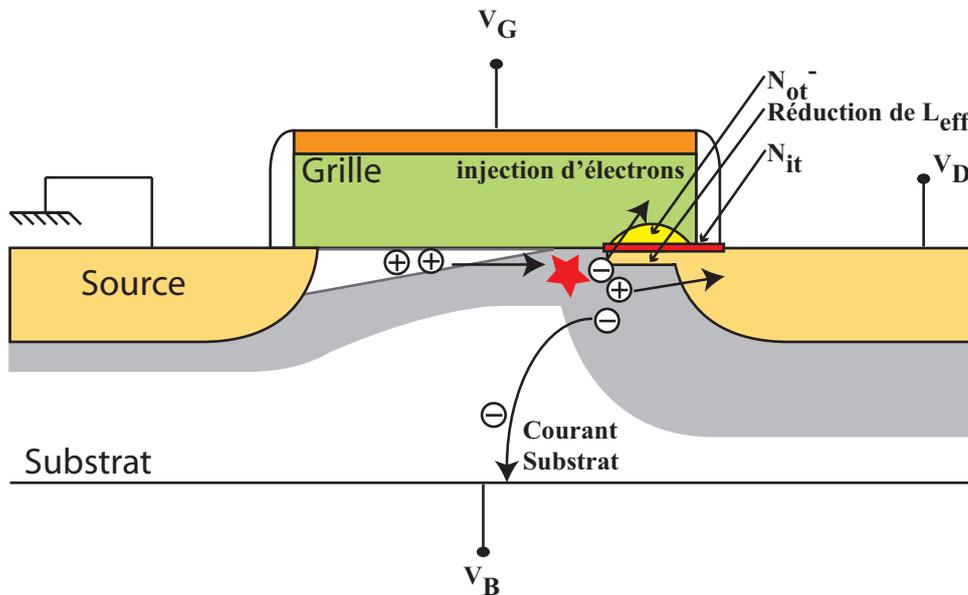


FIG. III.3 – Génération et injection de trous chauds dans un transistor MOSFET à canal p.

Ce sont Van Den bosh, Groeseneken et Maes qui ont introduit un modèle similaire au "Lucky electron" pour les trous, en développant le modèle du trou chanceu, ("lucky hole") [10]. La hauteur de la barrière entre le silicium et le dioxyde de silicium est plus faible pour l'injection

d'électrons ($3.2eV$ à $F_{OX} = 0V/cm$) que pour l'injection de trous pour lesquels $\phi_{b,h} = 4.8eV$ en l'absence de champ. L'abaissement de la hauteur de barrière induit par les lignes de champ dans l'oxyde est décrit par [10] :

$$\phi_{b,h} = 4.8 - \beta\sqrt{|F_{OX}|} - \vartheta|F_{OX}|^{2/3} \quad (\text{III.8})$$

Dans le modèle est fait l'hypothèse que β vaut $2.59 \times 10^{-4}(Vcm)^{1/2}$ tout comme pour les électrons et ϑ a été ajusté à $3.2 \times 10^{-5}V^{1/3}cm^{2/3}$. A taux d'ionisation équivalent, la dégradation par porteurs chauds est plus sévère pour les transistors NMOSFET que pour les transistors PMOSFET [11, 6]. Néanmoins, la dégradation HC dans les transistors PMOS submicroniques devient un problème délicat à étudier [12].

Quand le dispositif est polarisé en mode de saturation, des trous sont injectés dans la région désertée au Drain comme le montre la Fig. III.3. Quelques trous acquièrent assez d'énergie pour causer l'ionisation par impact, on les appelle trous chauds (HH). Les électrons produits dans la ZCE du Drain peuvent être redirigés vers l'oxyde et piégés pour les basses tensions V_{GS} . Si la densité des électrons piégés est suffisamment grande, la charge négative excessive dans l'oxyde attirera des trous à l'interface Si-SiO₂ et causera virtuellement une prolongation du Drain dans le canal. Ce phénomène a pour conséquence une réduction de L_{eff} et une diminution en valeur absolue de la tension de seuil, $|V_{Th}|$. Ceci peut être un problème sérieux pour les dispositifs à canal court, particulièrement sensibles aux modulations de L_{eff} et à l'enclenchement du phénomène de perçage qui peut conduire à la défaillance totale du PMOS. Le maximum de dégradation se produit quand le courant de Grille, I_{GS} , est maximal. Le mécanisme de piégeage d'électrons est dominant pour $|V_{GS}| < |V_{DS}|$, alors que le mécanisme d'injection de trous est dominant pour $|V_{GS}| \geq |V_{DS}|$. L'injection de trous a l'effet opposé au piégeage d'électrons : il produit une augmentation de $|V_{Th}|$.

Les effets porteurs chauds sont plus prononcés dans les dispositifs à canal court du fait qu'il n'est généralement pas possible de maintenir le même champ électrique dans les dispositifs réduits. C'est le cas quand une approche à réduction d'échelle à tension constante est mise en application. Afin de maintenir des tensions d'alimentation relativement élevées et en même temps réduire au minimum la génération HC, les technologies MOSFETs modernes mettent en application généralement une structure légèrement dopée du Drain (LDD) [3, 13]. Le but de cette région légèrement dopée entre le Drain et le canal est de décaler la position du champ électrique maximal dans la ZCE vers le Drain. L'importance du champ est également réduite [3, 13], où le pic de champ électrique montre une valeur minimum en fonction de la dose d'implant [14]. L'effet net du passage à la structure LDD est une réduction du courant substrat et un taux d'ionisation par impact inférieur, (qui a pour conséquence une génération d'états d'interface moindre) et moins d'injections d'électrons dans l'oxyde.

III.1.3 Influence de l'épaisseur de l'isolant de Grille

Pendant le fonctionnement du dispositif, le champ électrique F_{ox} à travers l'oxyde de Grille peut être assez élevé pour que les porteurs soient injectés dans l'oxyde, entraînant de ce fait des courants de fuite ou dans le pire des cas, des dommages irréversibles. Dans le cas des courants de fuite, le dispositif peut continuer à fonctionner mais à un niveau de performances inférieur; en effet, quand l'oxyde est moins isolant, le contrôle par la tension de Grille de la charge d'inversion dans le canal diminue. Dans le cas des dommages irréversibles, le dispositif cesse de fonctionner quand l'oxyde devient conducteur: la grille ne commande plus du tout la charge dans le canal.

En plus de l'injection de porteurs due à la génération HC, discutée dans la section précédente, les porteurs peuvent également être injectés dans l'oxyde de Grille par effet tunnel Fowler-Nordheim (FN) ou Direct (DT) [11] (les phénomènes tunnel sont également discutés en annexe). Le claquage de l'oxyde, ou "breakdown" (BD) se produit à des champs électriques à partir de $10MV/cm$ et se caractérise d'abord par un affaiblissement de la propriété isolante de l'oxyde dû au courant tunnel suivi de la création d'un chemin fortement conducteur entre le Substrat et le contact de Grille [11]. Le BD peut être classé dans deux catégories :

1. le claquage intrinsèque, qui se produit en l'absence de défauts, habituellement aux champs électriques élevés $F_{ox} > 10MV/cm$.
2. le claquage extrinsèque, inhérent aux défauts dans l'oxyde. Tous ces défauts, qui incluent la contamination au sodium, les contaminations métalliques, la rugosité de surface, et les amincissements localisés de l'oxyde, sont liés au processus de fabrication de l'oxyde de Grille et/ou aux processus qui précèdent ou suivent la formation du film d'oxyde [11].

Plusieurs mécanismes de claquage intrinsèque ont été proposés depuis quelques années [15, 16, 17]. Dans l'un d'entre eux, et certainement le plus représentatif, si la charge positive ou négative emprisonnée dans l'oxyde atteint une valeur critique, le courant de fuite tunnel augmente sensiblement, pouvant mener au claquage [18]. Le claquage se produit quand la densité de pièges atteint une valeur critique: un chemin conducteur local (constitué par les pièges) relie alors le contact de Grille au Substrat de silicium. On explique sa formation par le concept de percolation illustré sur la Fig. III.4. Il a été observé que la densité critique de pièges nécessaires pour claquer l'isolant diminue avec son épaisseur [18]. Pour ces modèles, la probabilité de claquage augmente avec le champ électrique et/ou avec la diminution de T_{ox} .

Puisque le claquage extrinsèque est lié aux défauts présents dans l'oxyde après fabrication, il est impératif de bien maîtriser les quantités de défauts induits par les étapes de fabrication. Ce problème est d'autant plus critique que l'isolant est mince. Les concentrations de défauts

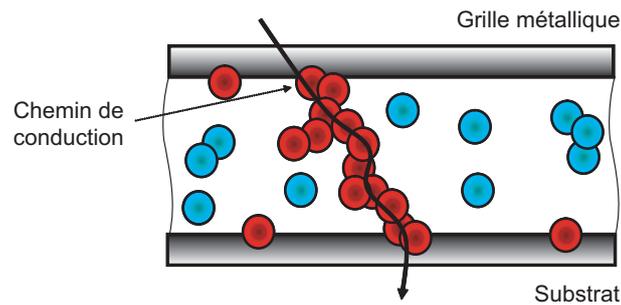


FIG. III.4 – Création d'un chemin de conduction par les pièges de charge, concept de percolation

doivent être maintenues les plus basses possibles par la maîtrise des étapes de fabrication, en particulier une préparation adéquate du substrat, mais aussi par le contrôle de la pré-oxydation [19]. Les risques de BD extrinsèques et intrinsèques des diélectriques de Grille peuvent être réduits pour les oxydes minces ($T_{ox} < 10nm$) par un processus d'empilement d'oxyde [19]. Dans ce procédé, un film d'oxyde est déposé en phase vapeur (CVD) est empilé sur l'oxyde thermique. Les avantages de ce processus vis à vis d'un processus thermique classique sont :

- la probabilité pour que les défauts soient alignés (et donc la probabilité de formation d'un chemin de conduction) est inférieure,
- l'oxyde contient peu de défauts provenant du substrat puisque moins de silicium y est consommé.

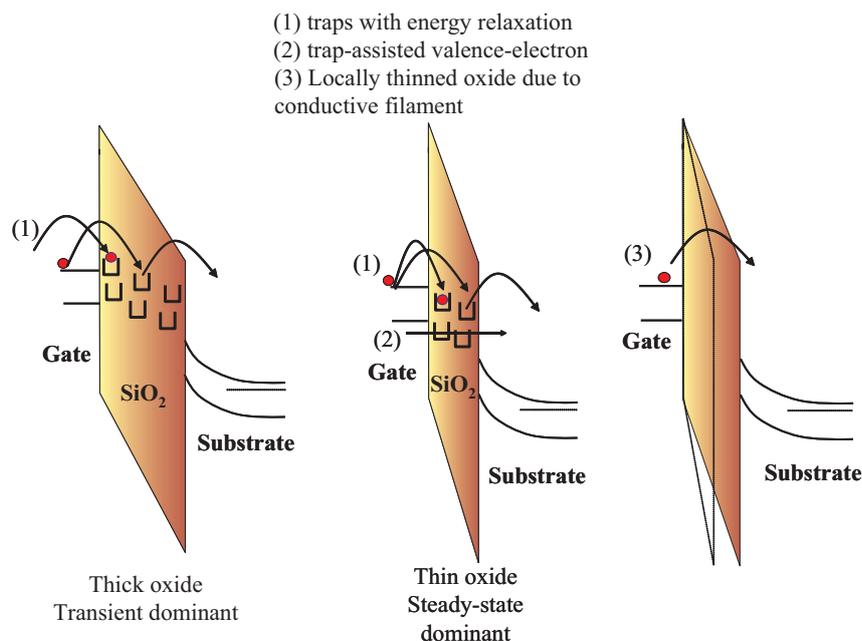


FIG. III.5 – Représentation schématique des mécanismes de conduction tunnel assistée par pièges responsables du SILC.

La fiabilité des oxydes empilés est ainsi sensiblement améliorée. Cette méthode de fabrication permet de surcroît la réalisation de films diélectriques plus minces que par les procédés thermiques conventionnels. C'est par ce principe que *STMicroelectronics* fabrique les oxydes ultra-minces de $2.1nm$ (GO1) et enchaîne sur la même plaque l'oxyde épais de $6.5nm$ (GO2). Dans les oxydes minces ($T_{ox} < 10nm$), les courants de fuite mesurés à bas champ augmentent après l'application de champs électriques élevés. Ces courants, qui sont générés en mode tunnel direct sont généralement désignés comme des courants de fuite induits par stress ou *Stress Induced Leakage current* (SILCs) [15, 20, 21]. La Fig. III.5 présente les mécanismes de conductions communément admis pour expliquer ce phénomène ; les pièges générés pendant le stress assistent la conduction (avec relaxation pour les oxyde épais, en régime permanent pour les oxydes plus fins, ou par amincissement local de l'oxyde). Le SILC peut représenter un problème de fiabilité prépondérant vis à vis de la rétention d'information dans les mémoires non-volatiles [15, 22], et peut être un signe précurseur de quasi-BD.

III.2 Expérimentations : stress statiques.

D'un point de vue pratique, la première méthode d'accélération de la dégradation est de soumettre le transistor à des contraintes DC en appliquant des tensions constantes à ses bornes. On doit choisir un V_{DS} approprié, pour être en régime HC. Ceci peut être déterminé par les courbes $I_{DS} - V_{DS}$ (V_{GS} variable) dans une gamme typique où le courant de substrat est maximum (cas référencé IB) et où le courant saturation I_{DSat} doit vérifier :

$$1.1I_{DSat}^* \leq I_{DSat} \leq 1.4I_{DSat}^* \quad (\text{III.9})$$

où I_{DSat}^* est le courant de saturation obtenu pour une condition de polarisation $V_{DS} = V_{DD}$ et $V_{GS} = V_{DD}/2$. Cette condition donne une gamme typique pour V_{DS} allant de 1.1 à 1.4 fois la tension d'alimentation V_{DD} selon la technologie examinée. Dans un second temps, le choix de V_{GS} régit le type de HC qui peuvent être majoritairement injectés dans l'oxyde. Pour un NMOSFET :

1. $V_{Th} \leq V_{GS} \leq V_{DS}/4$ pour l'injection trous chauds (HH),
2. $V_{GS} \approx V_{DS}/2$ à $V_{DS}/3$ pour la condition du maximum du courant substrat (IB),
3. $V_{GS} \approx V_{DS}$ est responsable de l'injection d'électrons chauds (HE).

Les conditions pour le PMOSFET sont obtenus en changeant les signes et en permutant les conditions 1 et 2 :

1. $|V_{Th}| \leq |V_{GS}| \leq |V_{DS/4}|$ entraîne l' injection d'électrons chauds (HE),
2. $V_{GS} \approx V_{DS}/2$ à $V_{DS}/3$ est la condition du maximum du courant substrat : injections de trous et d'électrons chauds simultanées (cas IB),
3. $V_{GS} \approx V_{DS}$ est responsable de l'injection de trous chauds (cas HH).

L'effet de la génération des porteurs chauds sur la caractéristique $I_{DS} - V_{DS}$ est montré sur la Fig. III.6. Le courant de Drain augmente et ne suit plus l'équation (I.70) du courant de Drain en régime saturé établie au premier chapitre : les électrons provenant des paires électron-trou générées s'ajoutent au courant "normal". La figure montre également le rapport du courant Substrat au courant de Drain qui suit bien le taux d'ionisation.

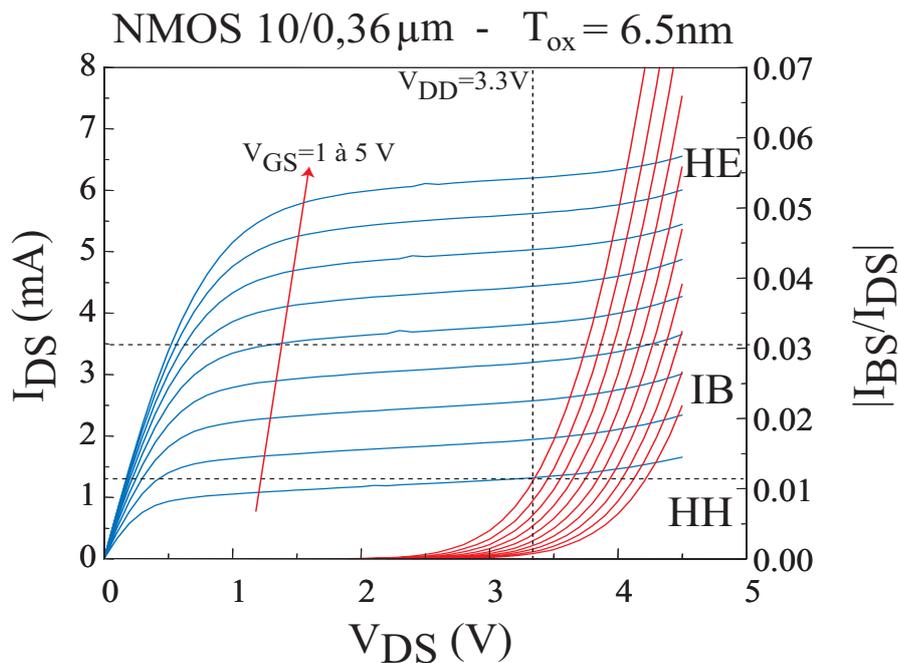


FIG. III.6 – Effet de la génération des porteurs chauds sur la caractéristique $I_{DS} - V_{DS}$ d'un transistor NMOS

III.2.1 Description des échantillons

Le suivi d'un paramètre transistor donné en fonction du temps de stress permet de déterminer la résistance d'une technologie de MOSFET au régime HC. Afin de placer l'étude des performances et de la fiabilité de la filière $0.13\mu m - 2.1nm$ dans son contexte, nous allons la comparer aux technologies ayant des épaisseurs d'oxydes plus grandes. Nous allons donc

présenter les résultats expérimentaux classés par mécanismes d'injections, et par épaisseur de l'isolant. Nous avons ainsi comparé trois épaisseurs d'oxyde: $12nm$, $6.5nm$ et $2.05nm$. Pour la première la longueur nominale de la technologie est de $0.5\mu m$. Les deux autres sont en fait issus de la même génération technologique et sont fabriqués simultanément sur les mêmes wafers. Les dispositifs avec $T_{ox} = 6.5nm$ sont destinés à gérer les entrées/sorties (I/O) des puces logiques dont les coeurs sont constitués de dispositifs d'épaisseur d'oxyde $T_{ox} = 2.05nm$. Ces derniers dispositifs à isolants de Grille ultra-minces sont divisés en deux catégories, les transistors à faibles niveaux de fuite avec une tension de seuil plus élevée, et les transistors rapides avec une tension V_{Th} faible. Les premiers sont nommés LL pour *Low Leakage* et les seconds HS pour *High Speed*. Cette variation de tension de seuil est réalisée par un ajustement de la dose de dopant à la surface du substrat dans le canal.

Ces échantillons ont été caractérisés et stressés par les protocoles décrits dans le Chap. II, sur banc de mesure sous pointes, avec connexions et mesures automatisées. Les Wafers analysés ont un diamètre de 8 pouces soit environ $20cm$. Les paramètres essentiels de nos échantillons sont consignés dans le tableau III.1.

Technologie	HCMOS5	HCMOS9-GO2	HCMOS9-GO1	
Label	T1	T2	T3 (HS)	T4 (LL)
$L(\mu m)$	0.5	0.36	0.13	
$T_{ox}(nm)$	12	6.5	2.05	
$N_{Sub}(cm^{-3})$	2×10^{16}	1.5×10^{17}	4×10^{17}	
Niveau de métallisation	5	7		
Tension nominale $V_{DD}(V)$	5	3.3	1.2	
Structure de Drain	LDD	MDD		

TAB. III.1 – Détails des dispositifs étudiés sous injections de porteurs chauds.

Les détails du procédé de fabrication pour les technologies T3/T4 sont représentés dans le Tab. III.2 et sur la Fig. III.7. Elles sont réalisées dans un double caisson : un caisson n ou p suivant le type de dispositif (canal p ou n) dans un caisson d'isolation n .

Etape	NMOS	PMOS
N-iso	P: 6×10^{13} - 1.5MeV - 7°	P: 6×10^{13} - 1.5MeV - 7°
WELL	B: 3×10^{13} - 170keV - 3°	P: 3×10^{13} - 440 - keV 3°
ANTI-PT	non	non
VT HS	In: 5×10^{12} - 160keV - 7°	As: 1×10^{12} - 120keV - 7°
VT LL	B: 6×10^{12} - 25keV - 7°	As: 5×10^{12} - 120keV - 7°
OXIDE	20.5Å	20.5Å
POLY THIC	1800Å	1800Å
LDD	As: 1.4×10^{15} - 2keV - 0°	BF ₂ : 5×10^{14} - 2keV - 0°
POCKET	As: 2.5×10^{13} - 67keV - 25°	BF ₂ : 5×10^{14} - 2keV - 25°
SPACER	L shape 80/200/500	L shape 80/200/500
Si ANNEAL	700°C, 1h	700°C, 1h
SD1	As: 2×10^{15} - 30keV - 7°	B: 2×10^{15} - 3keV - 7°
SD2	6×10^{13} - 30keV - 7°	6×10^{13} - 13keV - 7°
RTP S/D	spike 1113°C	spike 1113°C

TAB. III.2 – Paramètres de fabrication des structures GO1 : type de dopants, doses d'implants, énergie d'implantation et angle de diffusion, température des recuits.

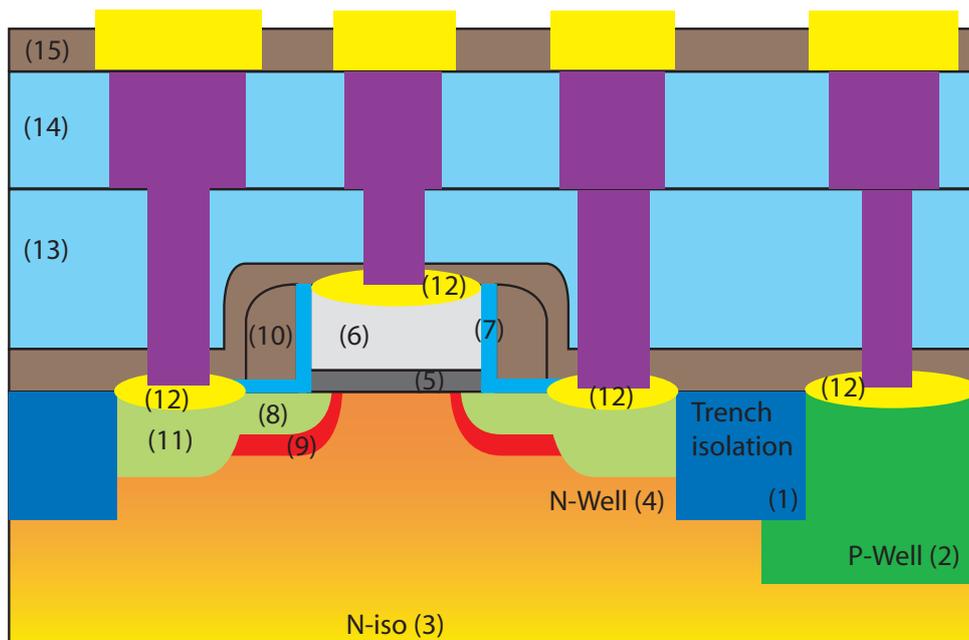


FIG. III.7 – Illustration des étapes de fabrication d'un transistor PMOS de la technologie T3 ou T4.

La fabrication débute avec la formation des tranchées d'isolation (1), on implante ensuite le caisson P-Well (Pour un PMOS) qui sert de prise Substrat (2). Le caisson isolant n-iso est ensuite réalisé (3) par diffusion de phosphore, suivie de la création du caisson N-Well (pour un PMOS) du transistor (4). L'isolant de Grille est conçu par oxydation (5) puis recouvert d'une couche de poly-Silicium de 1800\AA d'épaisseur (6), et enfin isolé latéralement par des espaceurs en forme de "L" dit "*L-shape*" (7). On enchaîne ensuite la réalisation des zones LDD (8) et des pockets (9) par diffusion de dopants (voir Tab. III.2). Les spacers sont alors réalisés (10) juste avant l'implantation des Source et Drain en deux étapes SD1 et SD2 (11). Enfin il reste à réaliser les connexions électriques avec d'abord les siliciures (CoSi_2) (12), les couches de métallisation successives (13, 14) et éventuellement, les pads extérieurs, isolés par une couche de damascène (15), pour faire des mesures électriques directement sur wafer. Nos échantillons, réalisés spécifiquement pour des mesures expérimentales, étaient pourvus d'un seul niveau de métallisation.

III.2.2 Oxydes épais : $12nm$

Cette section présente la dégradation des structures MOS avec des oxydes épais à moyens. Nous allons donc étudier les résultats des stress effectués sur la technologie *T1*. Ces résultats ont pour but de mettre en évidence les pires cas de dégradation.

a) Transistors NMOS, cas standard : maximum du courant substrat

Avant d'effectuer une série de vieillissements accélérés, il nous faut caractériser le régime porteurs chauds. Comme illustré sur les Fig. III.9 et III.8 on mesure les courants substrat et de Grille afin de connaître la polarisation donnant naissance aux pics de ces courants importants pour caractériser l'amplitude des phénomènes porteurs chauds.

$V_{DS}(V)$		5	5.5	6	6.5	7
$V_{GS}(V)$	I_{BS}^{max}	2.1	2.25	2.5	2.8	3
	$I_{G,e}^{max}$	5.9	6.5	6.9	7	-

TAB. III.3 – Polarisations du maximum des courants de Grille et de Substrat.

Les pics de courant Substrat sont obtenus pour des ratios $V_{GS}/V_{DS} \approx 0.4 - 0.5$, et les valeurs de courant obtenues pour ces polarisations sont importantes et atteignent $1.5mA$ pour une tension de Drain de $7V = 1.4V_{DD}$. Les stress IB sont les pires cas de dégradation pour les transistors NMOS à oxyde de Grille épais. Ceci est la conséquence du taux de génération de porteurs chauds très élevé. Le courant de Grille maximal est obtenu aux fortes valeurs de V_{GS} . Il s'agit d'un courant d'électrons, pour lesquels la hauteur de barrière est abaissée ($\leq 3.2eV$)

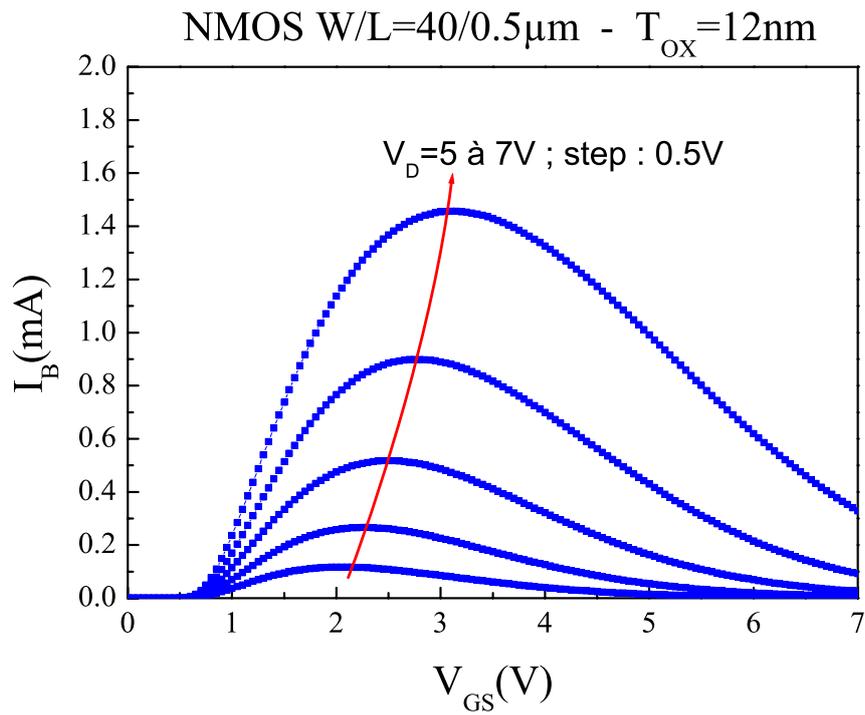


FIG. III.8 – Mesures du courant Substrat suivant V_{GS} et V_{DS} sur un transistor NMOS de la technologie T1.

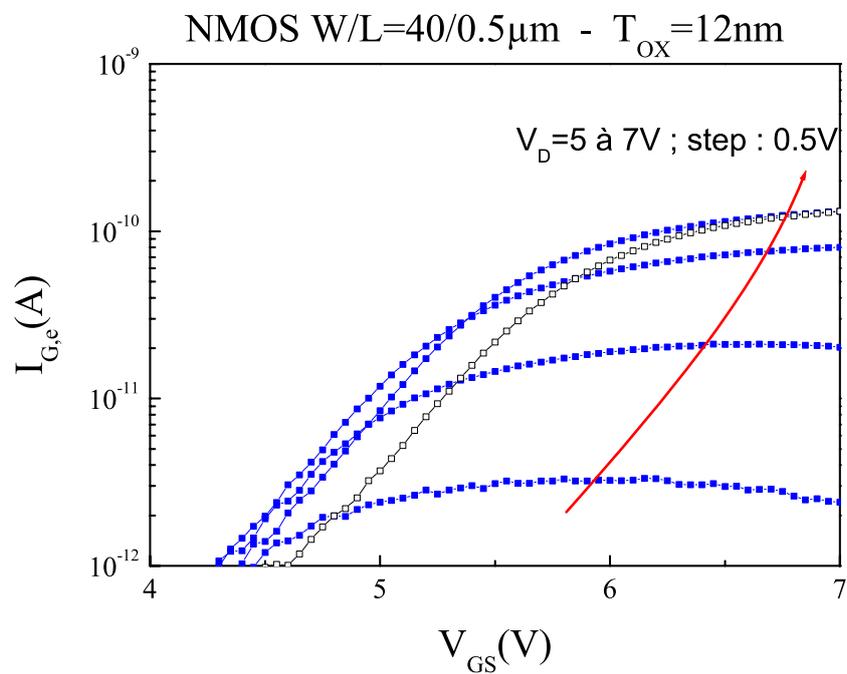


FIG. III.9 – Mesures du courant de Grille suivant V_{GS} et V_{DS} sur un transistor NMOS de la technologie T1.

du fait des fortes valeurs de F_{OX} . C'est pourquoi l'amplitude de ce courant thermoïonique $I_{G,e}$ est relativement importante (quelques centaines de pA) à fort V_{GS} . Pour caractériser le régime porteurs chauds, nous avons appliqué des tensions jusqu'à $7V$. La Fig. III.9 montre qu'en générant ces forts champs, on provoque le piégeage d'une charge négative Q_{ot}^- dans l'oxyde, dont la conséquence immédiate est de décaler la courbe de I_{GS} à cause de l'augmentation de V_{Th} (labels évidés). Enfin, l'influence de la tension de Grille sur la dégradation du courant de Drain linéaire est représenté sur la Fig. III.10, le maximum de dégradation est obtenu autour de $V_{GS} = V_{DS}/2$, qui correspond au maximum du courant Substrat.

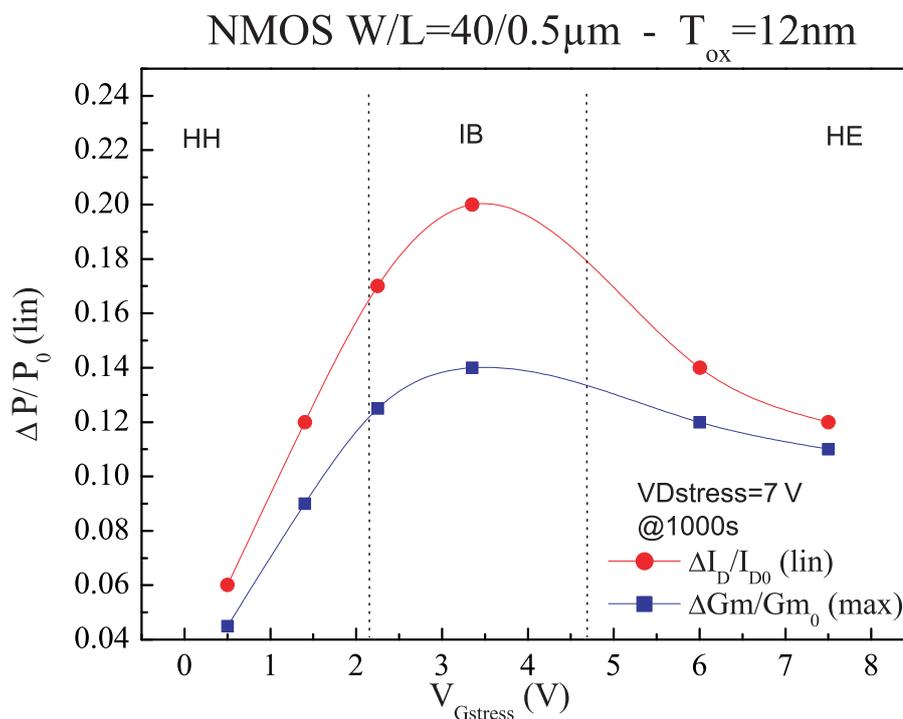


FIG. III.10 – Influence de la tension de Grille sur la dégradation du courant de Drain en régime linéaire pour $V_{DS} = 7V$ et $t = 1000s$.

Sur la Fig. III.11 sont représentées les cinétiques de génération d'états d'interfaces obtenues par mesure CP (détaillée en section II.2) pour les trois cas de stress HH, HE, IB. Conformément à ce que vous venons de voir, c'est pendant le stress IB que l'on a le plus de défauts générés avec une pente de 0.4 pour une quantité de défauts $4 \times 10^{11} cm^{-2}$ à $t = 1000s$, alors que les stress HE et HH montrent une pente plus faible (0.3 et 0.25) et des quantités de défauts à $t = 1000s$ plus faibles ($1.5 \times 10^{11} cm^{-2}$ et $9 \times 10^{10} cm^{-2}$). La partie droite de la Fig. III.11 montre la bonne corrélation entre la génération d'états d'interface $\Delta N_{it}/N_{it0}$ et la dégradation du courant de Drain $\Delta I_{DS}/I_{D0}$ (≈ 1) pour le cas IB, et une pente moins élevée pour les deux autres cas représentés HE et HH.

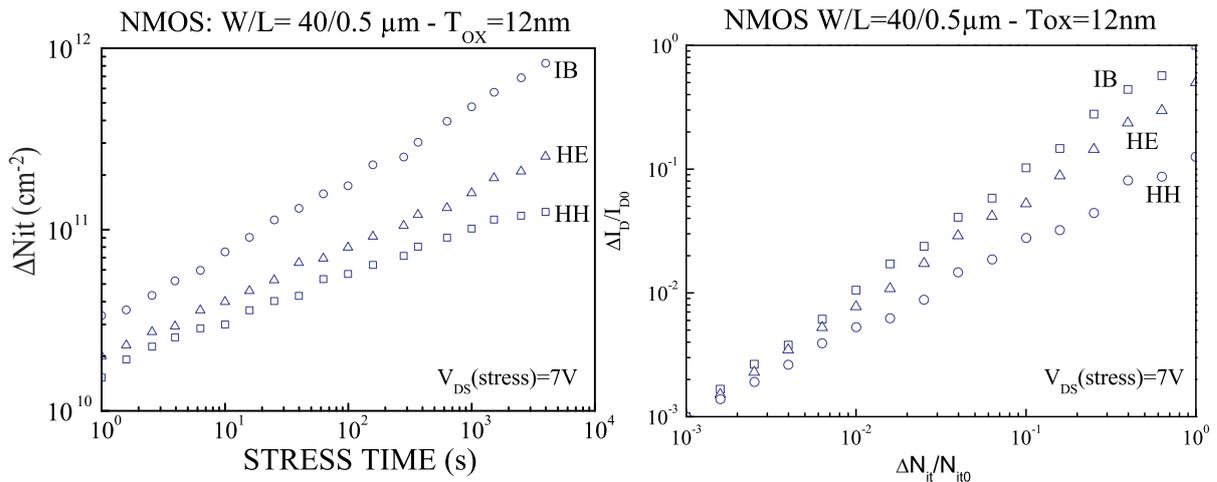


FIG. III.11 – Représentation de l'évolution du nombre d'états d'interface obtenu par mesure CP au cours des stress HH, HE et IB.

Nous avons soumis les transistors NMOS de la technologie T1 à des stress de type IB dans une gamme de tensions de Drain comprises entre 5.25V à 7.5V. Pour cette dernière tension de stress, la Fig. III.12 présente la dérive des caractéristiques $I_{DS} - V_{GS}$. Le courant de Drain et la transconductance sont fortement dégradés sur un temps relativement court ($t = 1000s$). Le pic de Gm ne se déplace pas, ce qui montre que la plus forte variation se produit pour les tensions de Grille faibles. La variation relative du courant de Drain peut être considérée comme la somme des variations de transconductance et de la tension de seuil [23, 24]:

$$\frac{\Delta I_{DS}}{I_{D0}} = \frac{\Delta Gm}{Gm_0} + \frac{\Delta V_{Th}}{V_{GS} - V_{Th0}} \quad (\text{III.10})$$

Dans cette relation $\Delta Gm/Gm_0 \approx \Delta \mu/\mu_0$. Nous avons donc représenté sur la Fig. III.13 les variations de la transconductance, du courant de Drain et la tension de seuil afin de déterminer lequel de ces paramètres est le plus sensible à l'injection de porteurs chauds au maximum du courant substrat. On observe que le courant de Drain et la transconductance exhibent à peu près (les pentes sont sensiblement différentes) la même variation, en contraste avec la dégradation de la tension de seuil qui augmente faiblement. Nous allons donc nous concentrer sur le paramètre $\Delta I_{DS}/I_{D0}$.

La Fig. III.14 montre les impacts combinés de ΔN_{it} et $\Delta \mu$ sur le courant de Drain pour la condition IB. En régime linéaire il est fortement dégradé, ce qui implique une forte réduction de la mobilité des porteurs du canal provoquée par une génération d'états d'interface importante. Pour des tensions de mesures plus élevées (5V/5V), la dégradation est plus faible. Ceci donne une première indication sur la localisation de la dégradation près de Drain. Une indica-

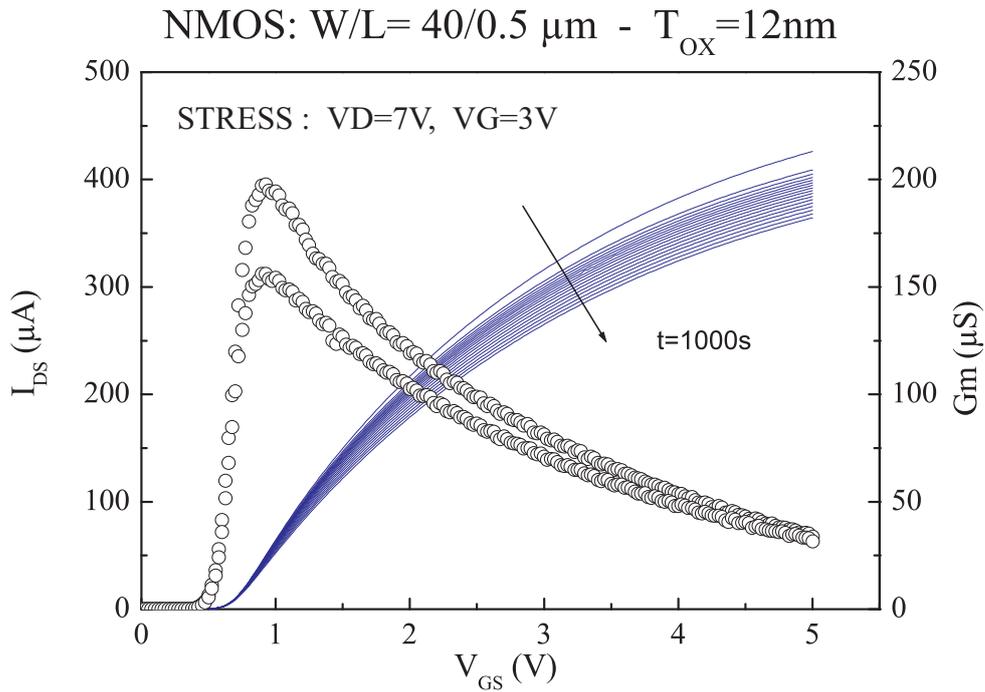


FIG. III.12 – Variation des caractéristiques $I_{DS} - V_{GS}$ et $Gm - V_{GS}$ (pour $V_{DS} = 25\text{mV}$) pendant un stress IB d'un transistor NMOS de la technologie T1.

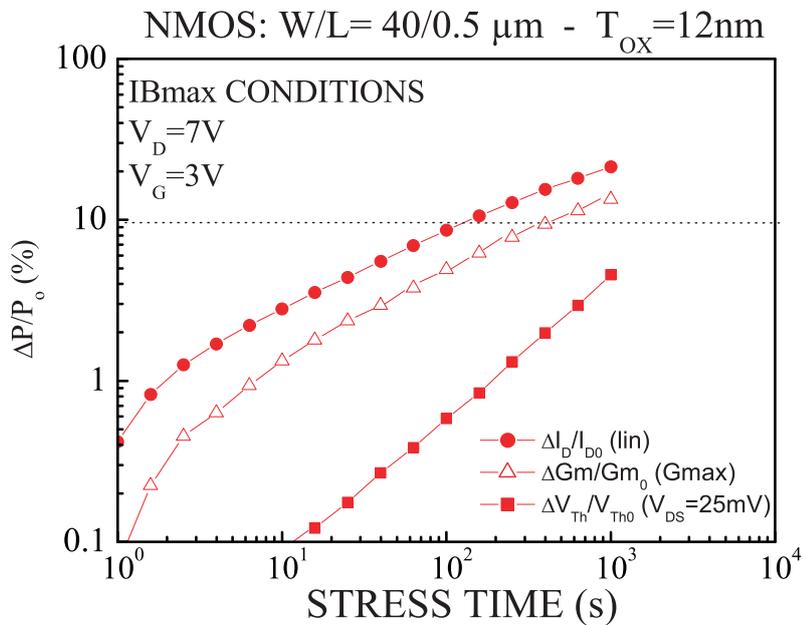


FIG. III.13 – Variation du courant de Drain, de la transconductance et de la tension de seuil pour les NMOS de la technologie T1 soumis au stress IB.

tion supplémentaire montrant la dégradation est le fait que $\Delta I_D/I_{D0}$ mesuré à V_{DD}/V_{DD} est nettement plus faible. C'est avec cette polarisation que la mesure du courant est le moins dégradé, justifiant ainsi qu'une grande partie des états d'interface sont générés près du Drain.

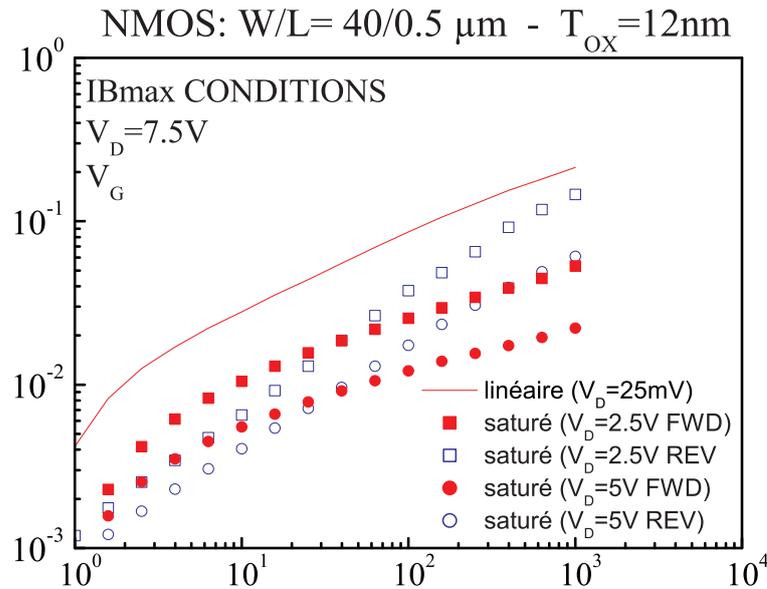


FIG. III.14 – Cinétique de la variation relative du courant de Drain d'un transistor NMOS de la technologie T1. Le Stress est réalisé au maximum du courant Substrat pour la polarisation de stress $V_{DS} = 7V$ et $V_{GS} = V_{DS}/2$.

La différence importante entre la variation du courant de Drain mesuré à V_{DD}/V_{DD} en mode FWD et REV implique que les défauts sont localisés près du Drain, puisqu'en mode FWD une partie de ces défauts sont écartés par l'extension de ZCE de la jonction Drain-Substrat, alors qu'ils influencent I_{DS} dans les zones 1, 2 et 3 (canal, zone de recouvrement et espaceur) de la Fig. II.1 dans le mode REV.

Les transistors NMOS se trouvent le plus fortement dégradés pour les conditions V_{GS}, V_{DS} correspondant au maximum de courant Substrat. Ce résultat Fig. III.11 est en accord avec les études de fiabilité porteurs chauds de la littérature pour cette génération CMOS [25]. Le Tab. III.4 donne le niveau de dégradation atteint dans les autres polarisations de stress. Les stress HH sont responsables d'un niveau de dégradation relativement faible par rapport aux conditions HE et IB.

Type de stress	tension (V_{DS}/V_{GS})	ΔV_{Th}	ΔN_{it}	$\Delta Gm/Gm_0$
VG+	7V/-	0.014V	$2.5 \times 10^{10} cm^{-2}$	8%
HH	7V/0.4V	0.02V	$3.6 \times 10^{10} cm^{-2}$	12%
HE	7V/7V	0.08V	$9 \times 10^{10} cm^{-2}$	16.5%
IB	7V/3.15V	0.05V	$7.5 \times 10^{11} cm^{-2}$	21%

TAB. III.4 – Niveaux de dégradation atteints lors des stress VG+, HH, HE et IB sur des transistors NMOS de la technologie T1, pour un temps de 1000s.

b) Transistors PMOS, cas standard : maximum du courant de Grille électronique

Le transistor MOSFET à canal p avec un isolant épais, voit ses performances fortement dégradées dans le régime porteurs chauds correspondant aux pics de courant de Grille électronique, qui représente le pire cas de dégradation. La Fig. III.15 donne le niveau de dégradation du courant de Drain en régime linéaire à 1000s en fonction de la tension de Grille pour un champ latéral suffisamment élevé ($V_{DS} = -7V$). Le pic de dégradation apparaît pour de faibles valeurs de la tension de Grille, au maximum du courant de Grille électronique. En effet pour les polarisation $V_{GS} \approx V_{DS}/4$, dans la région du Drain le champ latéral est maximal, et le champ dans l'oxyde est favorable à l'injection des électrons dans l'oxyde.

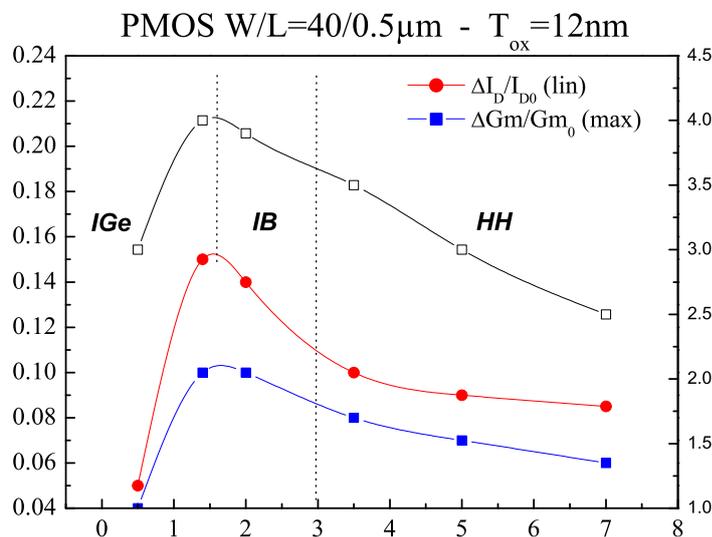


FIG. III.15 – Influence de la tension de Grille sur la dégradation du courant de Drain et la transconductance.

Lorsque l'on augmente $|V_{GS}|$, $|V_{Dsat}|$ augmente aussi, entraînant la diminution de ξ_{max} , réduisant directement le nombre de porteurs chauds dans la région de polarisation où le champ

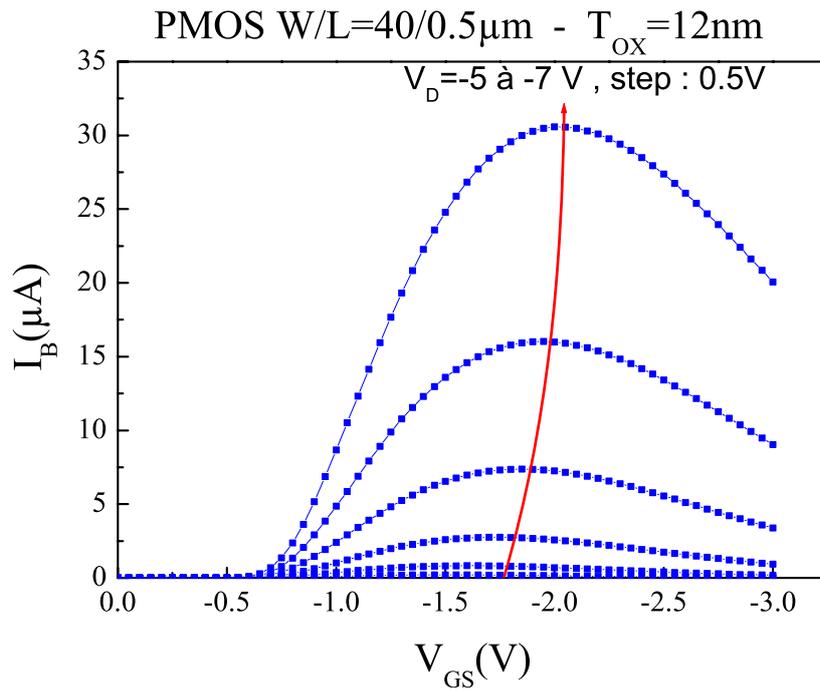


FIG. III.16 – Mesures du courant Substrat pour les transistors PMOS de la technologie T1.

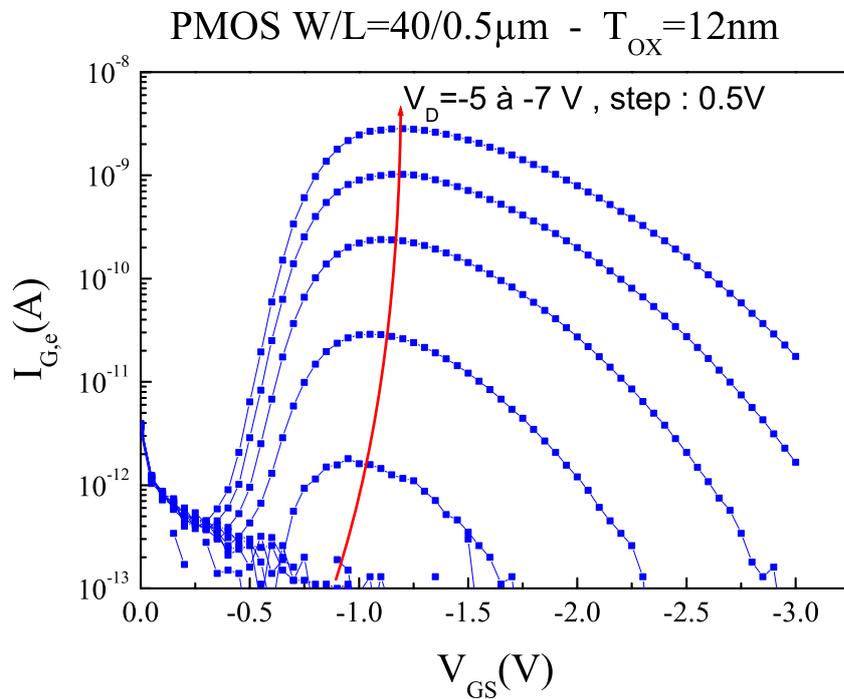


FIG. III.17 – Mesures du courant de Grille pour les transistors PMOS de la technologie T1.

vertical devient plus favorable à leur injection. De ce fait, les variations de $I_{G,e}$ suivent les variations du courant Substrat I_{BS} . Les Fig. III.17 et III.16 montrent la mesure de ces courants

en fonction de V_{GS} pour des tensions V_{DS} comprises entre $-5V$ et $-7V$. Le pic de courant électronique se déplace vers les tensions V_{GS} négatives à mesure que $|V_{DS}|$ augmente (Tab. III.5).

$V_{DS}(V)$		-5	-5.5	-6	-6.5	-7
$V_{GS}(V)$	I_{BS}^{max}	-1.6	-1.7	-1.8	-1.95	-2.05
	$I_{G,e}^{max}$	-0.95	-1.05	-1.1	-1.15	-1.2

TAB. III.5 – *Polarisations du maximum des courants de Grille et de Substrat.*

On peut remarquer que :

1. le courant Substrat reste faible par rapport au NMOS de géométrie identique, mais le courant de Grille électronique est plus conséquent ; c'est une des raisons pour laquelle le mode IGe est le plus dégradant dans le PMOS, comme nous le verrons plus tard.
2. malgré des conditions idéales pour l'injection de trous à $V_{GS} = V_{DS}$, on ne mesure pas de courant de trous à travers l'oxyde en raison d'un taux d'ionisation faible et d'une hauteur de barrière plus importante pour les trous : $\phi_{b,h} = 4.8eV$.

Intéressons nous aux cinétiques de stress IGe. La Fig. III.18 montre les variations relatives du courant de Drain, mesurées pour différentes polarisations: en régime linéaire ($V_{GS}/V_{DS} = -0.6V/25mV$), et saturé ($V_{GS} = V_{DS} = V_{DD}/2$ et $V_{GS} = V_{DS} = V_{DD}$) en mode FWD et REV. La dépendance temporelle du piégeage de charges négatives dans le PMOS suit une loi logarithmique [26], contrairement à la création d'états d'interface en loi de puissance.

On constate que c'est à $V_{DD}/2, V_{DD}/2$ que le courant est le plus réduit en valeur absolue. En augmentant le V_{DS} de mesure, le courant est moins dégradé. Si on augmente encore la tension de Drain jusqu'à V_{DD}/V_{DD} , la dégradation devient moins importante. La dégradation résulte du piégeage d'électrons, ce qui se traduit par l'extension du Drain conduisant à la réduction de la longueur électrique du canal. Cet effet est un phénomène généralement observé pour les dispositifs avec une structure de Drain de type LDD [27, 28]. Lorsqu'on augmente encore la tension de Drain, la dégradation diminue en raison de la localisation des défauts. C'est le phénomène d'écrantage qui rentre en jeu et masque la charge piégée et les états d'interface générés dans la zone de Drain au cours de la dégradation, comme nous l'avons observé dans le transistor NMOS.

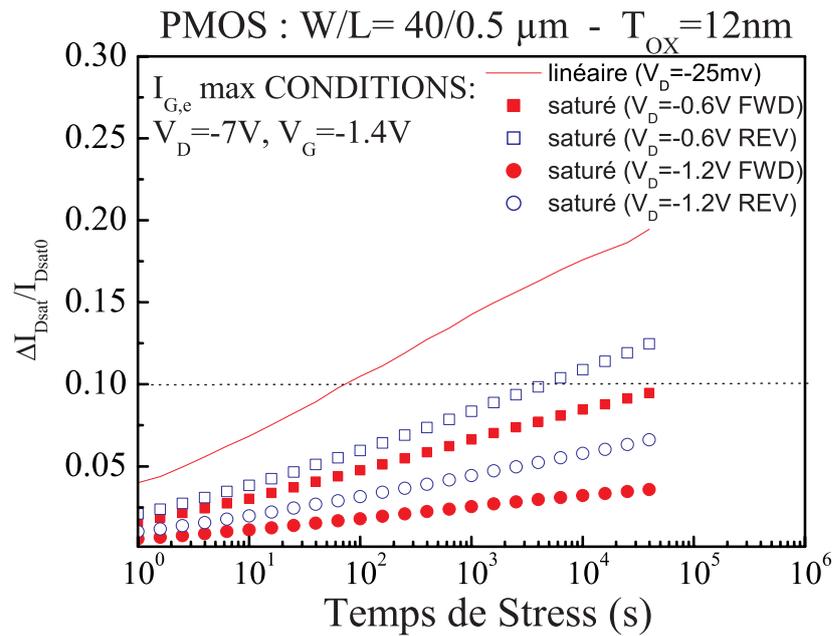


FIG. III.18 – Evolution relative du courant de Drain dans les conditions du maximum de courant de Grille électronique ($V_{DS} = -7\text{V}$ et $V_{GS} = -1.4\text{V}$) mesurée pour différentes polarisations en régime linéaire et saturé.

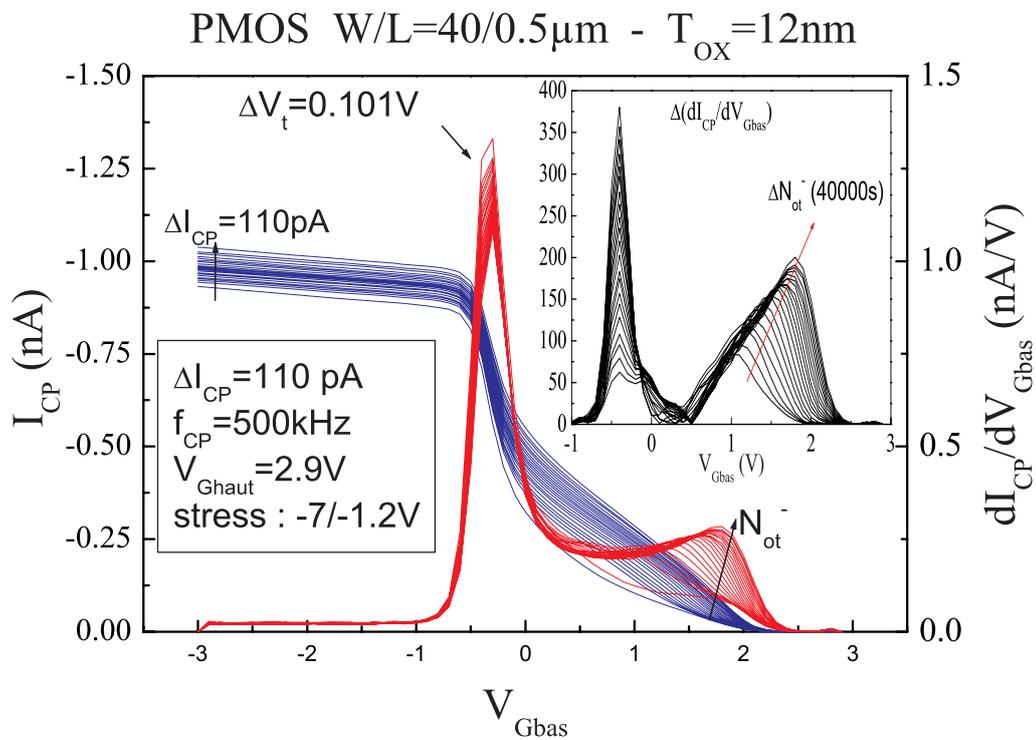


FIG. III.19 – Mesure de pompage de charge avec le niveau bas variable au cours du stress au maximum de courant de Grille électronique.

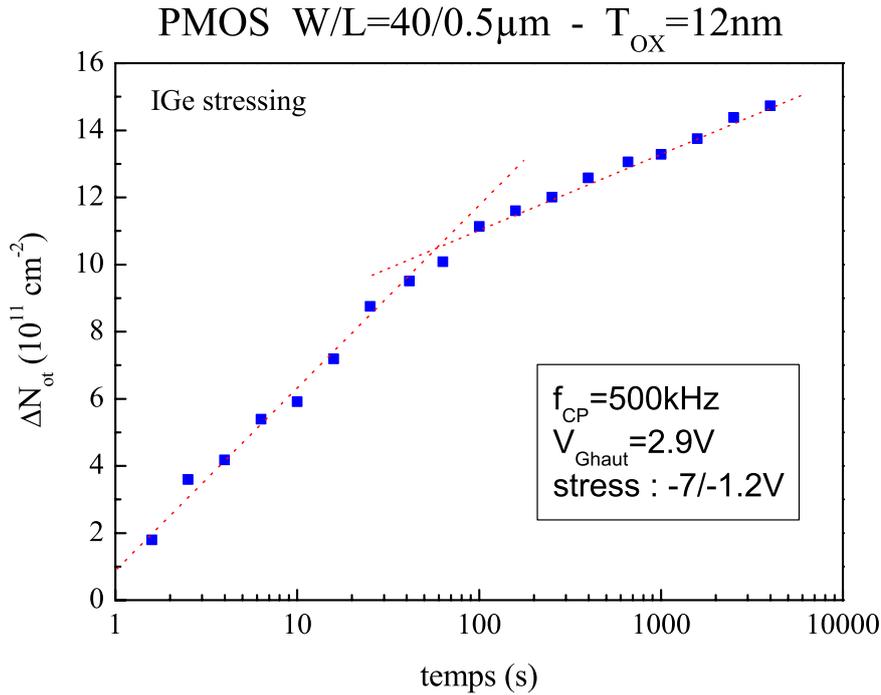


FIG. III.20 – Variation de la charge piégée dans l’oxyde calculée à partir des courbes CP pendant un stress au maximum du courant de Grille électronique.

Pour une analyse plus quantitative de la dégradation on peut s’intéresser à la mesure par pompage de charge effectuée au cours du stress (Fig. III.19). La mesure à V_{Gbas} variable présente à $t = 0s$ une valeur maximale du courant pompé de $-0.93nA$, ce qui correspond à une densité d’états $D_{it} = 5.8 \times 10^{10} eV^{-1} cm^{-2}$. Au cours du stress on observe une augmentation du courant pompé maximal de $\Delta I_{CP} = 110pA$ équivalant à une augmentation de la densité d’états $\Delta D_{it} \approx 7 \times 10^9 eV^{-1} cm^{-2}$, ce qui reste faible. Le deuxième point important est le décalage des courbes vers les tensions positives, confirmé par le déplacement du maximum des courbes dérivées. Ce décalage se retrouve dans la variation de la tension de seuil $\Delta V_{Th} = 0.101V$ due à une charge négative piégée $Q_{ot}^- = 2.9 \times 10^{-8} C.cm^{-2}$, également confirmé par la bosse sur les dérivées des courbes CP autour de $0.8V$ (technique décrite en II.2.3.b qui permet d’être plus sensible à N_{ot}^-). Dès les premières secondes du stress, cette bosse se manifeste à $1V$ du pic principal puis se décale de $\Delta V_{CP-} = 1V$ pour une variation du courant pompé $1.5 \times 10^{-10} A$, ce qui représente la valeur élevée $\Delta N_{ot} = 1.8 \times 10^{12} cm^{-2}$. La Fig. III.20 montre l’évolution de N_{ot} qui suit une loi logarithmique du temps.

Avant de terminer l’étude des oxydes épais, examinons le cas de l’injection uniforme VG- réalisée à $V_{DS} = 0V$. L’injection se fait par effet tunnel Fowler Nordheim. La Fig. III.21 compare la variation du courant de Drain en linéaire pour les deux cas IGe et VG-. Pour cet oxyde de $12nm$ le mode FN apparaît pour $V_{GS} \geq 10V$. Le stress VG- est effectué pour une tension

de Grille de $11V$, qui correspond à un champ dans l'oxyde proche de $-9MV/cm$. Le stress IGe montre une loi logarithmique du temps, ce qui n'est pas le cas pour VG- qui est en loi de puissance.

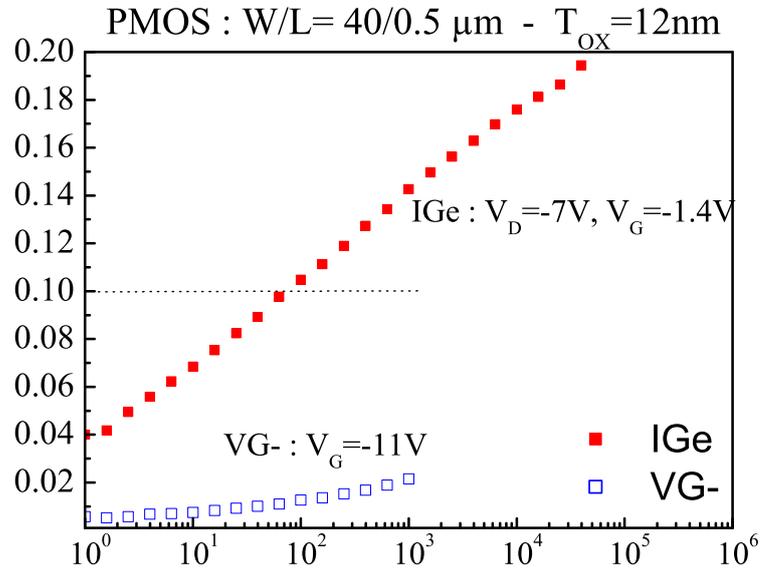


FIG. III.21 – Comparaison des dégradations du courant de Drain (linéaire) pendant les stress IGe et VG- ($-11V$).

Les mesures CP effectuées pendant le stress VG- n'ont pas permis la détection de charges piégées dans l'oxyde, mais une génération d'états d'interface caractérisée par une augmentation du courant pompé : $\Delta I_{CP} = 250pA$. Cette valeur donne une quantité d'états d'interface $N_{it} = 1.56 \times 10^{10} cm^{-2}$. Cette quantité est proche de celle obtenue dans le cas du stress IGe, pour lequel la présence des charges piégées entraîne une réduction de courant 10 fois plus importante.

III.2.3 Oxydes moyens : $6.5nm$

Pour l'étude de cette gamme d'épaisseur d'oxyde, nous avons étudié la technologie T2 où T_{ox} vaut $6.5nm$. Ces dispositifs sont conçus à partir de T3/T4 par ré-oxydation de l'oxyde de $2.1nm$ jusqu'à $6.5nm$. L'épaisseur d'oxyde de cette technologie autorise un courant tunnel plus important que pour l'oxyde $12nm$, en mode Fowler Nordheim accessible pour des tensions plus élevées. Les dégradations VG+ sont donc les plus faibles. Avant de présenter les résultats expérimentaux pour la technologie T2, précisons que les échantillons dont nous disposons avaient une diode de protection (vis à vis des décharges ESD notamment) sur la Grille, rendant impossibles les mesures par pompage de charges.

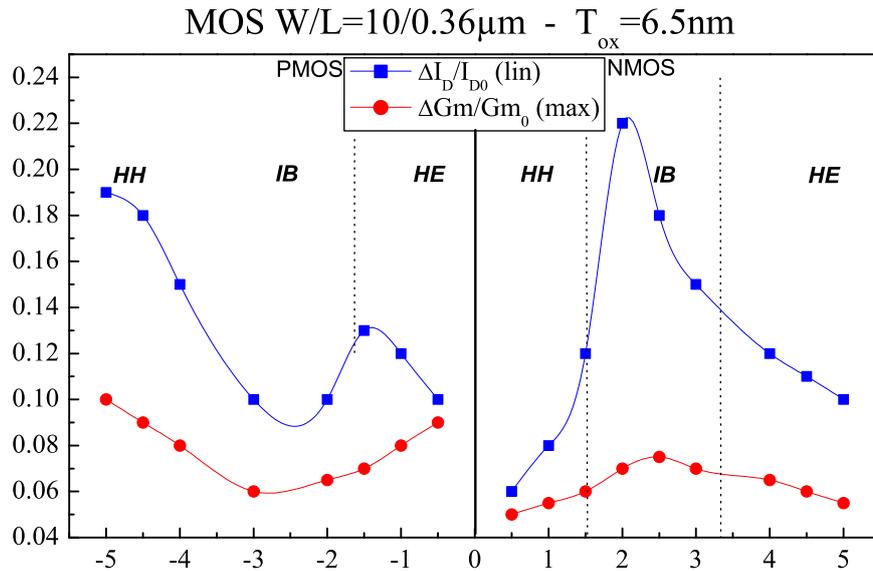


FIG. III.22 – Influence de la tension de Grille sur la dégradation du courant de Drain (régime linéaire) et de la transconductance pour la tension de Drain de stress $V_{DS} = 5V$ pour $t=1000s$.

La Fig. III.22 montre que le pire cas de dégradation pour cette technologie s’est avéré être le maximum du courant substrat pour les NMOS. Pour le PMOS, le mécanisme de dégradation le plus néfaste est l’injection HH réalisée pour $V_{DS} \approx V_{GS}$, alors que le maximum du courant substrat s’obtient à $V_{GS}/V_{DS} = 0.35$. Nous avons également étudié le cas de stress effectués dans des conditions d’injections uniformes à $V_{DS} = 0V$ ($V_{G\pm}$).

a) Cas des transistors NMOS

Par rapport aux oxydes épais, les dispositifs avec des oxydes dans cette gamme de T_{ox} ne présentent pas d’évolution majeure dans leur mode de dégradation le plus sévère, le fonctionnement au maximum du courant substrat. Examinons la caractérisation du régime porteurs chauds représentée sur les Fig. III.23 et III.24. Elle met en avant un ratio de 0.4 entre les tensions de Drain et de Grille pour obtenir le maximum de génération de porteurs chauds. Contrairement aux dispositifs avec des oxydes de $12nm$, on mesure I_{Ge} pour des tensions relativement basses (de 3 à 5.25V pour $V_{DD} = 3.3V$) et jusqu’à $2nA$ pour $W = 10\mu m$. D’autre part les pics de I_{BS} sont obtenus dans une gamme de tension de Grille où peu de porteurs (électrons et trous) atteignent la Grille, comme le montrent les mesures de courant de Grille effectuées dans les mêmes conditions (Fig. III.24). Ceci indique que la majorité des électrons est repoussée vers le Substrat à cause des conditions de champ peu favorables à leur injection dans l’oxyde pour $1V < V_{GS} < 2.7V$ et $V_{DS} = 3$ à 5.25V.

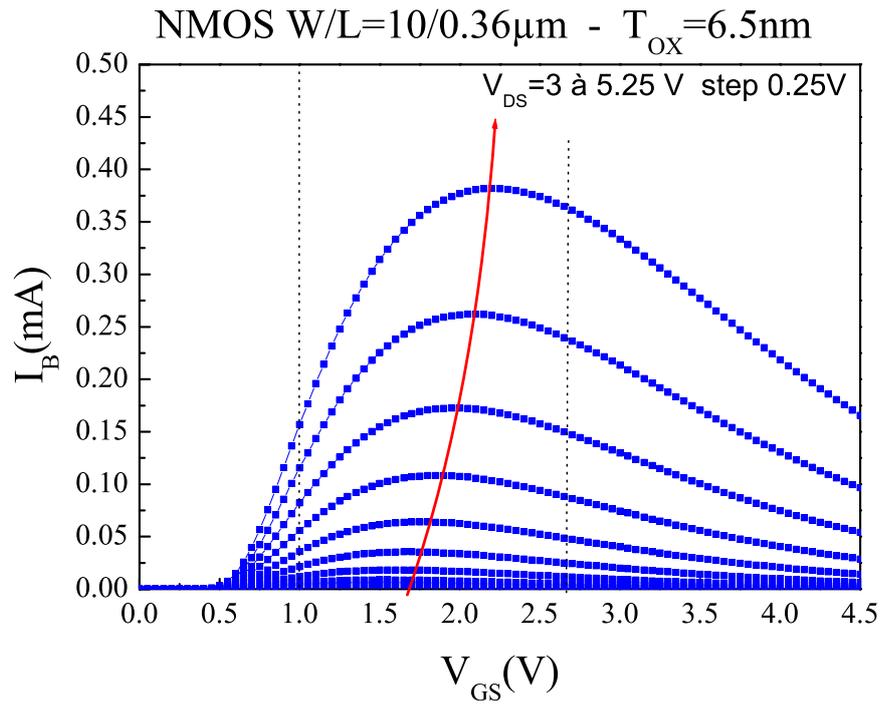


FIG. III.23 – Mesure du courant Substrat suivant V_{GS} pour différentes tensions de Drain pour un transistor NMOS de la technologie T2.

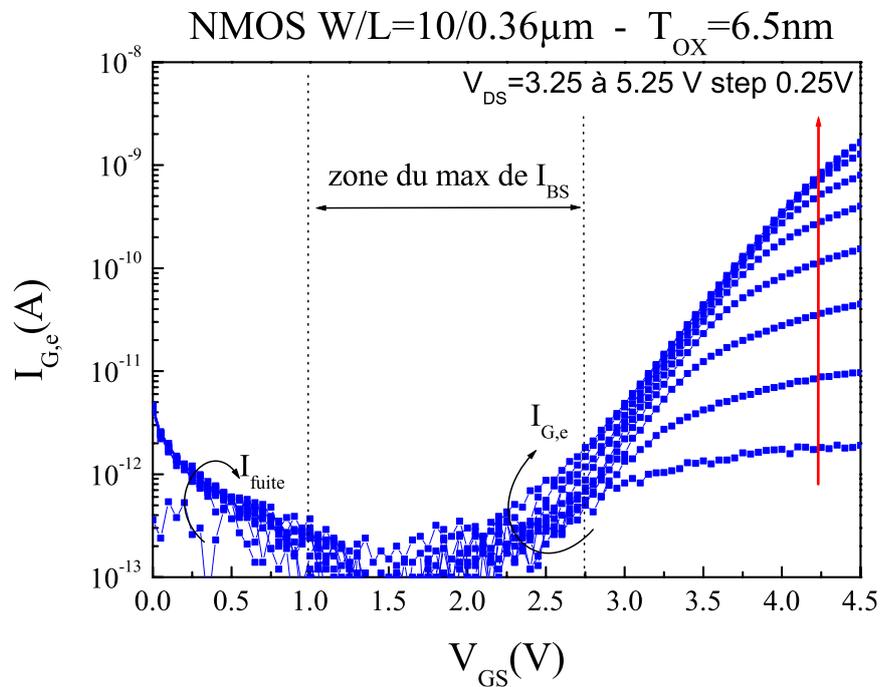
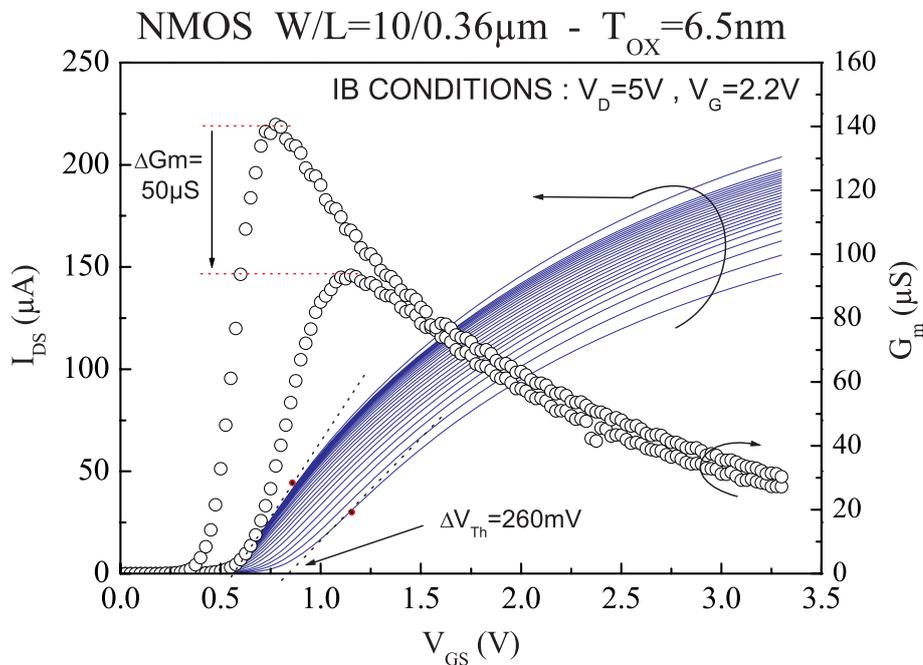


FIG. III.24 – Mesure du courant Grille suivant V_{GS} pour différentes tensions de Drain pour un transistor NMOS de la technologie T2.

$V_{DS}(V)$		3.25	3.5	3.75	4	4.25	4.5	4.75	5	5.25
$V_{GS}(V)$	I_{BS}^{max}	1.3	1.4	1.5	1.6	1.75	1.85	1.95	2	2.2
	$I_{G,e}^{max}$	4.25	4.5	4.75	5	5.25	5.5	5.75	6	6.25

TAB. III.6 – Polarisations du maximum des courants de Grille et de Substrat.

La Fig. III.25 présente l'évolution pendant le stress de la caractéristique $I_{DS} - V_{GS}$ et de la transconductance. G_m est réduit de 30% et son maximum est décalé de 0.4V. On peut exprimer la dégradation du courant de Drain comme le cumul de la variation de mobilité et de la tension de seuil (III.10) [24]. Dans ce cas nous avons donc $\Delta G_m/G_{m0} = 35\%$ et $\Delta V_{Th}/V_{Th} = 48\%$. La variation de la tension de seuil (elle augmente) est donc prépondérante sur la réduction de mobilité (à travers la variation de transconductance), sans que cette dernière ne soit négligeable pour autant. Par rapport aux dispositifs avec $T_{ox} = 12nm$, ces deux paramètres présentent ici une forte réduction : ΔG_m reflète une forte dégradation de la mobilité des porteurs dans le canal et ΔV_{Th} suggère une forte génération d'états d'interface mais également un piégeage de charges dans l'oxyde (II.2.3.a).


 FIG. III.25 – Evolution de la caractéristique $I_{DS} - V_{GS}$ au cours d'un stress au maximum de I_{BS} pour un transistor NMOS de la technologie T2.

La Fig. III.26 montre la réduction du courant de Drain $\Delta I_{DS}/\Delta I_{D0}$ à la condition du maxi-

mum du courant Substrat dans plusieurs configurations de mesures. La comparaison des modes FWD et REV nous indique qu'une partie des défauts est écrantée par l'extension de la ZCE : en mode REV la dégradation est maximisée du fait que la partie du canal est sous la zone dégradée, ce qui signifie une localisation au Drain des défauts.

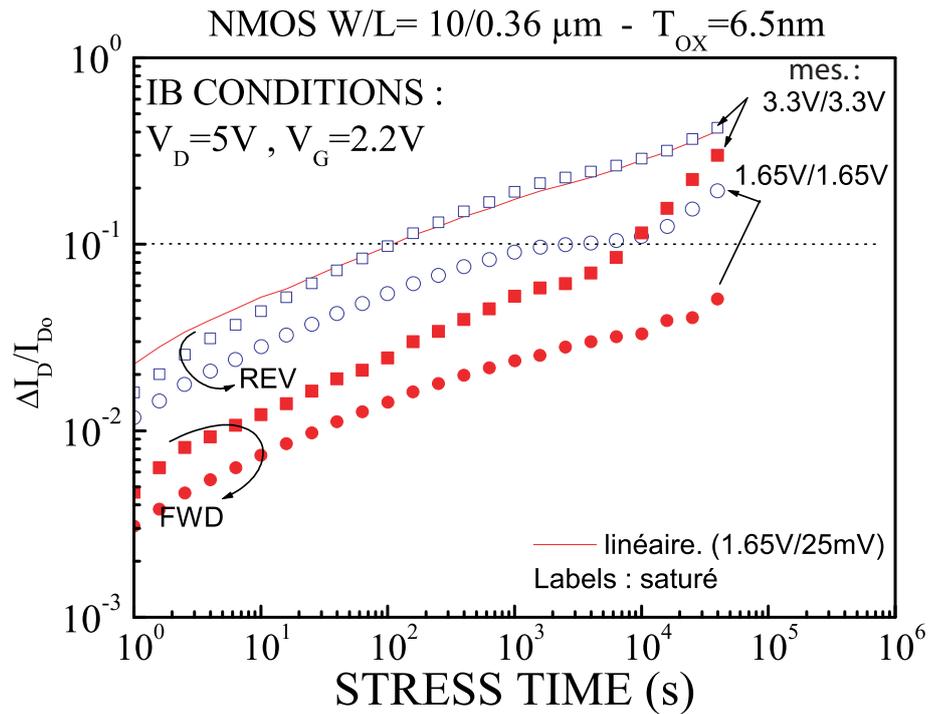


FIG. III.26 – Variation de la dégradation relative du courant de Drain au cours d'un stress au maximum du courant substrat ($V_{DS} = 5V$) pour un transistor NMOS de la technologie T2.

Concernant les injections uniformes, on note par rapport à T1, que le mécanisme d'injection uniforme VG+ a une influence plus prononcée du fait de l'augmentation de F_{OX} liée à l'isolant plus fin. Pour $V_{GS} = V_{DD}$, le champ dans l'isolant est défini par :

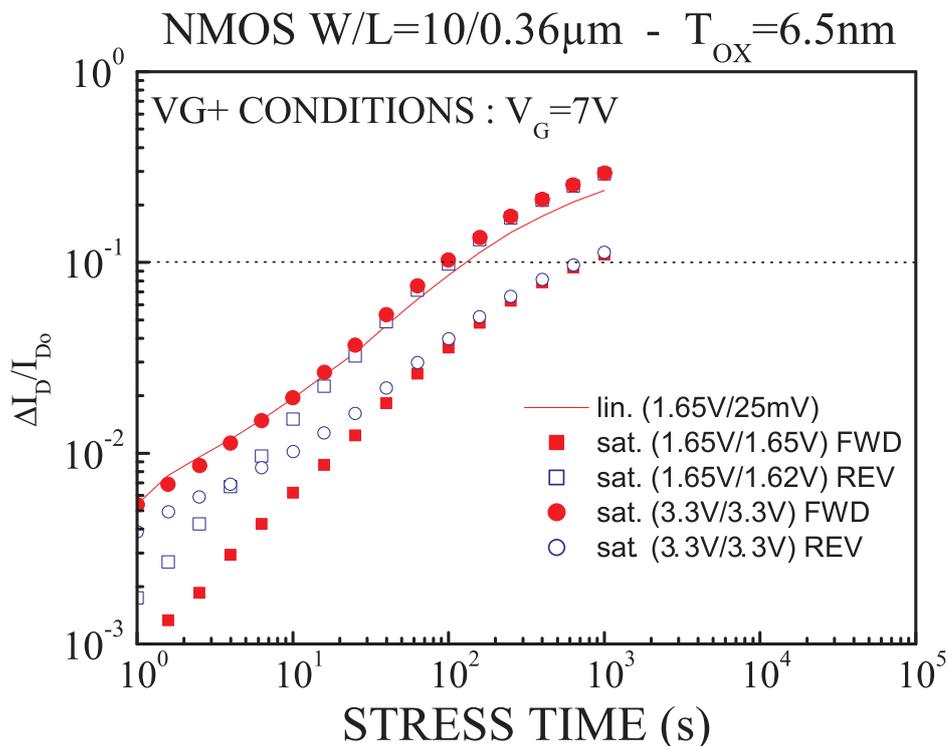
$$F_{ox} = \frac{V_{ox}}{T_{ox}} = \frac{V_{DD} - V_{FB} - 2\phi_F}{T_{ox}} \quad (\text{III.11})$$

Le tableau III.7 donne la valeur du champ dans l'oxyde pour $V_{GS} = V_{DD}$, et $V_{FB} = -E_g/2 - \phi_F - qN_{it0}^D(\phi_F)/C_{ox}$.

$T_{ox}(nm)$	12nm	6.5nm	2.1nm
$\phi_F(V)$	0.35	0.4	0.43
$V_{FB}(V)$	-0.91	-0.96	-0.99
$C_{ox}(F/cm^2)$	2.87×10^{-7}	5.31×10^{-7}	1.64×10^{-6}
$F_{ox}(MV/cm)$	4.34	5.32	6.33

TAB. III.7 – Valeur du champ électrique dans l'oxyde de Grille en fonction de son épaisseur.

Nous avons représenté sur la Fig. III.27 la variation du courant de Drain d'un transistor NMOS de la technologie T2 soumis à une injection uniforme avec une tension de Grille $V_{GS} = 7V$. Cette tension de stress est élevée vis à vis de la tension d'alimentation (3.3V), mais est révélatrice de l'effet de l'injection des électrons, les 10% de dégradation sont atteints en 100s. Comme pour le cas des stress au maximum du courant substrat, la pente de ces cinétiques est $n \approx 0.5$. Notons qu'au delà de ce temps de 100s on observe une légère saturation des courbes, qui peut être imputable à un effet de saturation de la génération de ΔN_{it} .


 FIG. III.27 – Dégradation du courant de Drain pendant une injection uniforme à $V_{GS} = 7V$ mesurée dans plusieurs conditions de polarisation (linéaire et saturé) pour un NMOS de la technologie T2.

b) Cas des transistors PMOS

Pour compléter cette étude des oxydes moyens ($6.5nm$), nous avons stressé des transistors PMOS de la technologie T2 pour identifier le pire cas de dégradation. Comme pour le cas du NMOS, l'influence du champ vertical est plus importante avec la réduction de T_{ox} . La Fig. III.28 compare les différents type d'injections possibles suivant la tension de Stress V_{GS} pour une tension de Drain de $-4.5V$ [29]. A très bas V_{GS} (au maximum du courant électronique de Grille), on observe une augmentation du courant de Drain (en valeur absolue) relative à l'effet du raccourcissement du canal dû à la charge négative. En valeur relative, cette variation reste inférieure au régime VG- et HH. Ces deux derniers cas de figure correspondent respectivement à la condition d'injections uniforme $V_{DS} = 0V$, et aux injections plus localisées de trous chauds pour $V_{GS} \approx V_{DS}$. Un premier résultat nouveau, est le fait que la condition pire-cas PMOS est devenue $V_{GS} = V_{DS}$ à fort champ latéral, montrant une réduction nette du piégeage d'électrons. D'autre part, le comportement à la polarisation I_{BSmax} , se démarque du résultat standard HE en piégeage d'électrons. Le dernier a nécessité une attention particulière que nous aborderons plus tard.

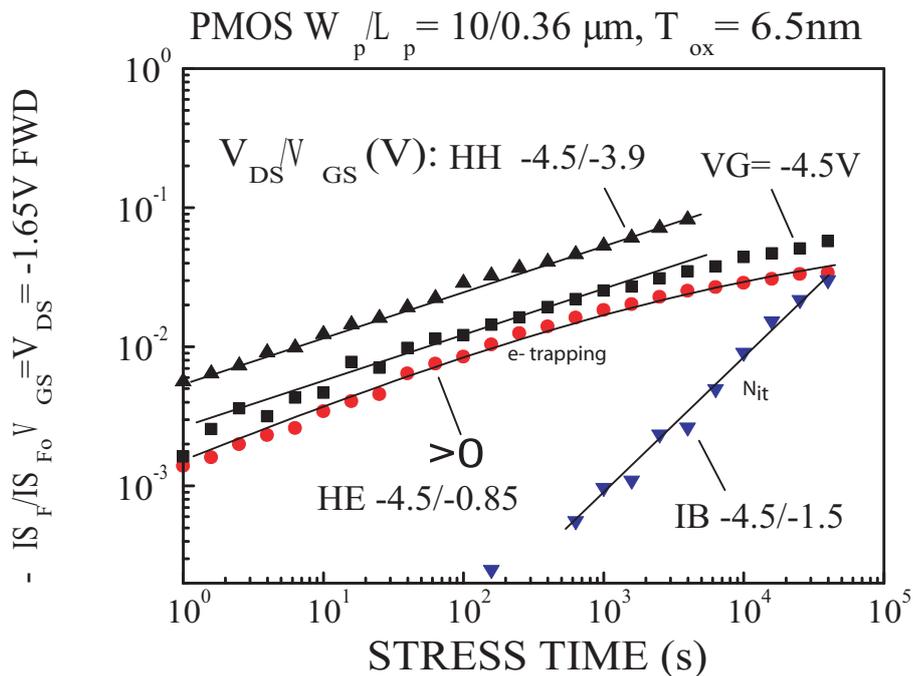


FIG. III.28 – Comparaison des mécanismes d'injection HH, IB et VG- dans les transistors PMOS de la technologie T2 [29].

La Fig. III.29 compare les modes de mesure direct et inverse de la dégradation du courant en régime saturé pour V_{DD}, V_{DD} et $V_{DD}/2, V_{DD}/2$. La réduction du courant de Drain est plus forte à $V_{DD}/2$ qu'à V_{DD} . Ceci est lié à l'écrantage d'une partie des défauts par l'extension de la ZCE pour la polarisation V_{DD}, V_{DD} . En effet quand $V_{DS} = V_{GS} = V_{DD}/2$ l'extension de la zone de

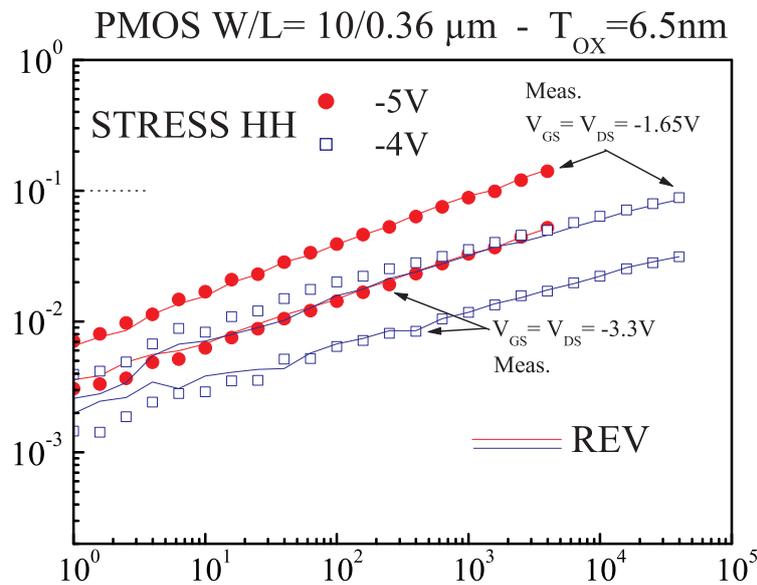


FIG. III.29 – Variation du courant de Drain mesuré en mode saturé et comparé en mode FWD et REV au cours d'un stress HH.

charge d'espace est faible (canal faiblement pincé), la partie active du canal est donc en contact avec une zone de défauts plus grande. Mais les similitudes entre les mesures REV et FWD nous permettent de conclure que la dégradation n'est pas localisée au Drain mais s'étale au dessous du canal vers la source, rendant la dégradation quasi uniforme.

Le cas IB est intéressant puisqu'il présente un retournement de cinétique au cours du vieillissement dû à la compétition entre deux mécanismes :

1. Aux temps courts, les électrons chauds générés sont injectés dans l'oxyde et piégés, donnant une charge N_{ot}^- et des états d'interface donneurs. La charge négative ainsi piégée, a pour effet de réduire la longueur effective électrique du canal. Ceci a pour effet d'accroître le courant (en valeur absolue), comme dans le cas HE (réalisé à $V_{DS}/V_{GS} \approx 5$). L'augmentation de la charge Q_{ot}^- réduit le champ électrique latéral ξ_m , et donc l'efficacité du piégeage au cours du temps.
2. Quand ξ_m n'est plus assez important, le piégeage diminue fortement, tendant rapidement vers zéro. Les trous chauds générés ont un effet qui devient dominant sur le piégeage d'électrons. Ces trous génèrent des états d'interface donneurs N_{it}^+ , soit par recombinaison avec les électrons, soit en dépassant les liaisons Si-H, ce qui a pour conséquence directe de réduire $|I_{DS}|$ et augmenter $|V_{Th}|$. Cet effet de retournement est bien visible sur la Fig. III.30.

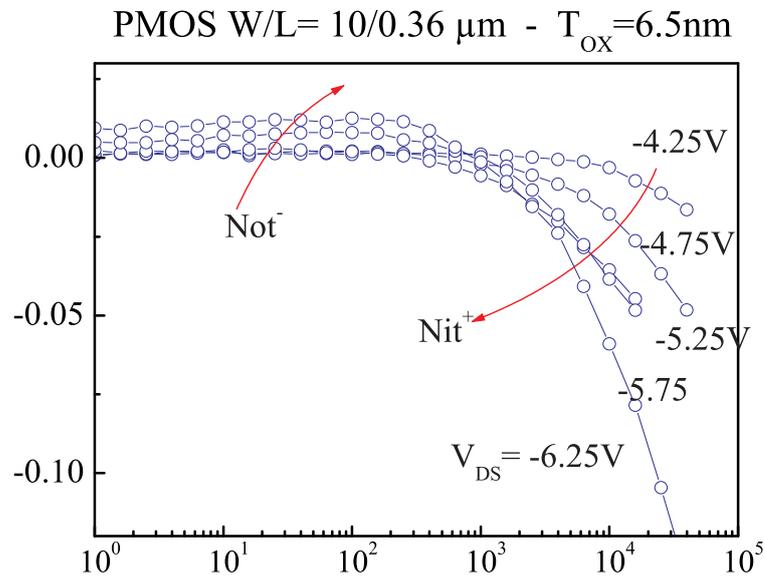


FIG. III.30 – Variations de la tension de seuil pour les transistors PMOS au cours d'un stress IB .

Dans cette section, nous venons de voir que pour la technologie T2, les transistors NMOS vieillissent plus rapidement lorsqu'ils sont soumis à des injections de porteurs chauds dans les conditions de polarisation du maximum de courant substrat. Ce sont les électrons qui semblent être à l'origine de la dégradation des caractéristiques électriques par la génération d'états d'interface. Le transistor PMOS est également le plus fortement dégradé dans ces conditions. Toutefois, le mécanisme n'est pas identique, puisqu'on observe un retournement : la génération d'états donneurs à l'interface (N_{it}^+) domine à long terme, la création de pièges accepteurs dans l'oxyde (N_{ot}^-) aux temps courts. Dans les deux cas, ce type de dégradation influe directement sur les vitesses de commutation des circuits logiques, lié à la réduction de I_{DS} .

III.2.4 Oxydes minces : 2.1nm

Dans cette section, nous allons décrire les mécanismes de vieillissement des dispositifs CMOS à oxyde de Grille ultra minces (2.1nm). Le régime porteurs chauds est caractérisé avec les courbes $I_{BS} - V_{GS}$ pour plusieurs valeurs de la tension de Drain, afin de déterminer les pics de courant Substrat associés à V_{GS} . La Fig. III.54 présente ces caractéristiques pour la filière GO1. Le PMOS présente un ratio classique V_{GS}/V_{DS} proche de 0.3, et pratiquement constant vis à vis de la tension de Drain. Le NMOS présente une valeur de ce ratio plus singulière dans la mesure où il n'est pas constant, puisqu'il varie de 1 pour $V_{DS} = V_{DD}$ à 0.5 quand $V_{DS} = 3.5V$.

En second lieu, on observe que le niveau du courant substrat mesuré est nettement plus important pour le NMOS, et ce pour des tensions accélératrices moins élevées que pour le PMOS. Ceci est dû à la mobilité des trous plus faible que celle des électrons, ce qui entraîne une cinétique suffisante pour enclencher l'ionisation par impact avec un seuil énergétique plus faible ($\phi_{i,e} = 1.3eV$) que pour les trous ($\phi_{i,h} = 2.54eV$) [30].

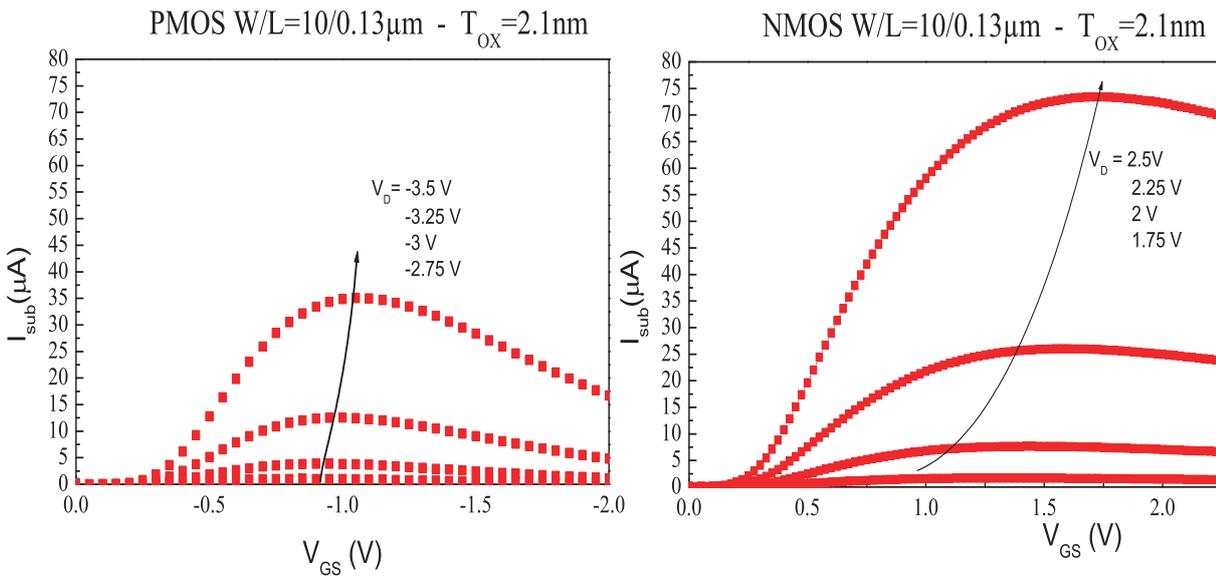


FIG. III.31 – Mesure du courant Substrat en régime porteurs chauds pour plusieurs valeurs de V_{DS} sur les transistors NMOS et PMOS de la technologie T3.

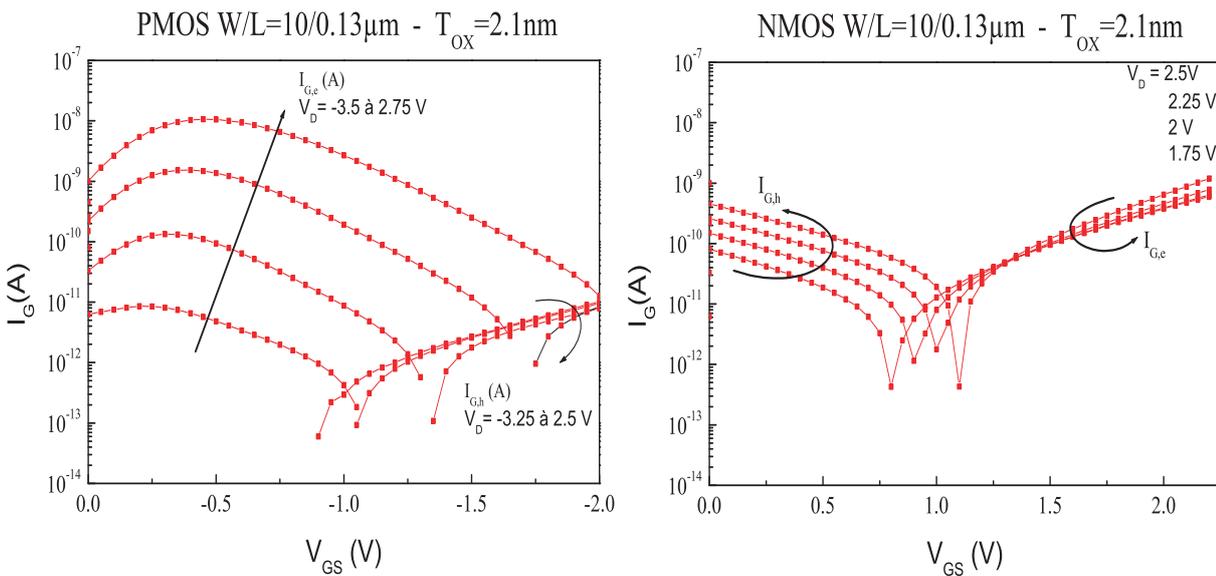


FIG. III.32 – Mesure du courant de Grille en régime porteurs chauds pour plusieurs valeurs de V_{DS} sur les transistors NMOS et PMOS de la technologie T3.

Avant de rentrer dans le détail des stress, notons que la finesse des oxydes de T3 et T4, implique un courant tunnel en mode direct et non plus en mode Fowler Nordheim. Ceci implique que les porteurs peuvent franchir la barrière sans avoir préalablement acquis une énergie cinétique importante. On peut en effet mesurer un courant de Grille significatif avec $V_{DS} = 0V$ (pas d'accélération) comme le montre la Fig. III.33. On peut y observer un courant substrat en l'absence de champ latéral, qui traduit une seconde composante dans le courant DT. Ces différentes composantes du courant tunnel direct sont illustrées sur le schéma de la Fig. III.34. A bas V_{GS} , dans les PMOS, des trous provenant de la bande de valence du canal en inversion (CI) sont injectés vers la grille (HVB). Pour des tensions de grille intermédiaires et supérieures, ce sont les électrons de la bande de valence qui sont injectés de la grille dans le substrat (EVB). Dans les NMOS soumis à de faibles V_{GS} , on observe un courant d'électrons du CI vers la Grille (ECB), puis, pour des tensions de Grille plus importantes, on observe l'apparition d'un courant d'électrons de la bande de valence (EVB) qui se traduit par un courant de trous dans le substrat.

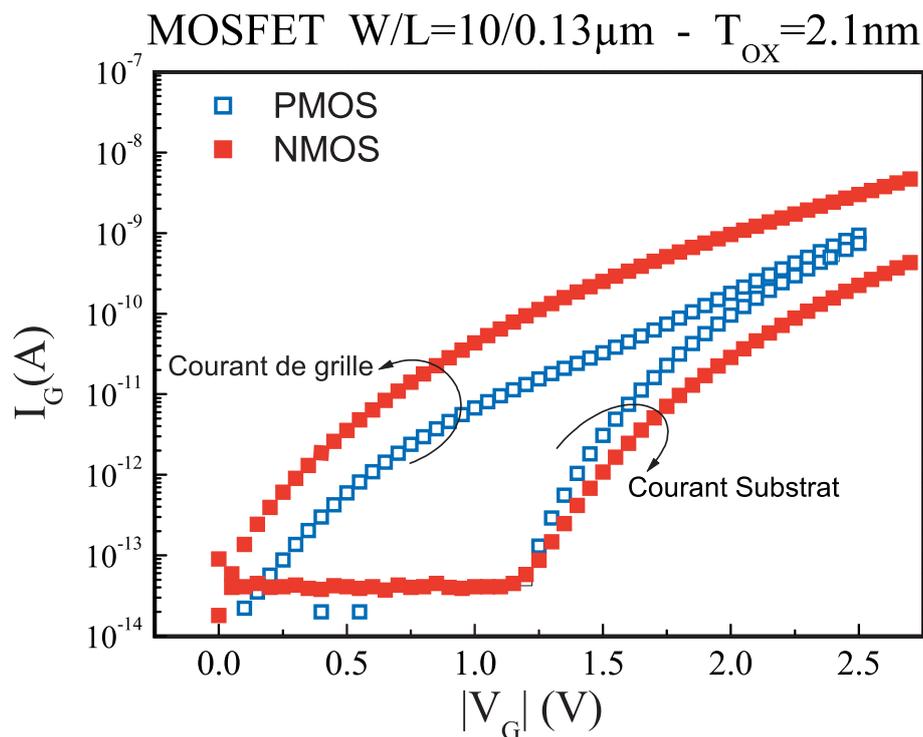


FIG. III.33 – Mesure des courants de Grille et Substrat en l'absence de tension de Drain pour les transistors NMOS et PMOS pour la technologie T1.

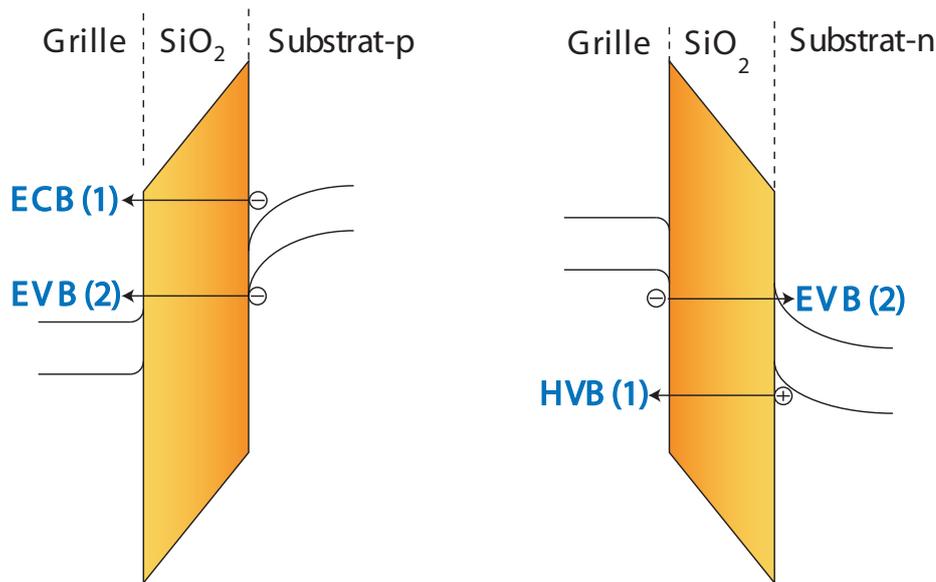


FIG. III.34 – schéma des composantes du courant tunnel direct.

a) Étude des transistors NMOS

Nous avons détaillé et caractérisé le régime porteurs chauds à l'aide des courants représentatifs de ces phénomènes. Nous allons maintenant comparer les trois types de dégradation standards pour mettre en évidence les différents mécanismes de génération de défauts :

1. stress sous injection d'électrons chauds (HE) ;
2. stress au maximum du courant substrat (IB) ;
3. stress uniforme à $V_{DS} = 0V$ (VG+), il ne s'agit pas d'un stress porteurs chauds mais la finesse de l'oxyde impose de se pencher sur ce cas.

i. Stress HE. En polarisant la structure à $V_{GS} = V_{DS}$ pendant le stress, on impose qu'en $x = L_{eff}$ le champ vertical vérifie $F_{OX} \approx 0V/cm$. La condition la plus favorable d'injection se situe près du Drain. La différence entre les cas HE et IB réside finalement dans le champ vertical imposé par la combinaison des potentiels V_{GS} et V_{DS} , et dans la proportion distincte d'électrons et de trous chauds supposée égale pour le cas IB par rapport au cas HE où il y a plus d'électrons chauds. Pour les fortes tensions de Grille ce sont les électrons qui sont injectés dans l'oxyde, car leur probabilité d'injection est plus importante avec V_{GS} dans le cas HE. Le fait que les deux types de stress soient assez proches en terme de niveau dégradation atteint, semble indiquer que l'injection des trous (réalisée dans le cas IB) est faible et a peu d'impact sur la qualité de l'interface.

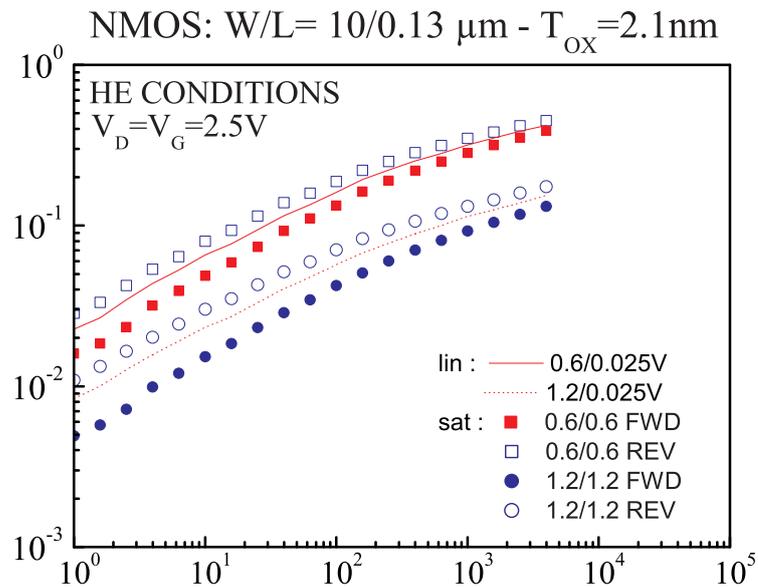


FIG. III.35 – Variation relative du courant de Drain d'un NMOS $W/L = 10/0.13\mu\text{m}$ lors d'un stress HE mesurée en régime linéaire comparé au régime saturé.

Les mesures de la variation du courant de Drain pour différentes polarisations de mesure sont tracées sur la Fig. III.35. Première constatation quant au régime saturé (V_{DD}, V_{DD}) : on observe une différence nette entre les mode REV et FWD ce qui donne une première indication de la localisation des dégradations au Drain. Regardons maintenant le régime linéaire ($V_{DS} = 25\text{mV}$). Quand V_{GS} est élevé (V_{DD}) le courant est nettement moins dégradé que lorsque la tension de Grille est plus faible ($V_{DD}/2$). Avec $V_{DS} = 25\text{mV}$ le canal n'est pas pincé. Tout le canal est donc sensible à l'intégralité de la partie dégradée sous le canal. Nous avons donc ici une forte baisse de la mobilité liée au champ vertical, causée par la rugosité de surface (défauts présents à l'interface) en hausse (augmentation de la résistance du canal à bas V_{GS}). Enfin, en comparant les régimes linéaire et saturé FWD pour la même tension de Grille prise à $V_{GS} = 0.6\text{V}$, on observe que les trois courbes (linéaire, saturé FWD et REV) sont assez similaires : l'écrantage de la zone des défauts par l'extension de la ZCE influe peu, impliquant un étalement de la dégradation plus vers le canal.

On a représenté sur la Fig. III.36 l'ensemble des cinétiques réalisées sur les NMOS de la technologie T3. Pour l'intervalle de tensions de stress comprises entre $V_{DS} = 1.75\text{V}$ et $V_{DS} = 2.5\text{V}$, les courbes $\Delta I_{DS}/I_{D0}$ ont la même pente $n \approx 0.4$ ce qui suggère que la dépendance temporelle de la génération de la dégradation reste proportionnelle à V_{DS} et ne présente pas de phénomène de saturation.

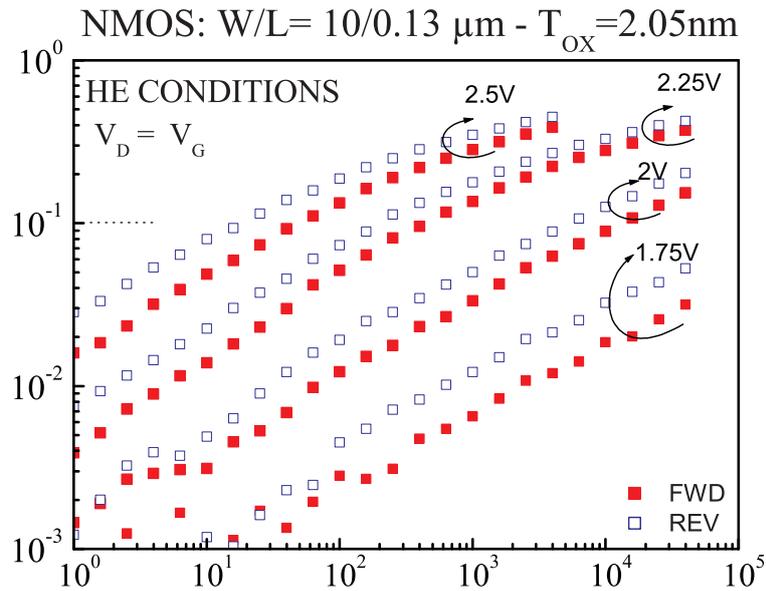


FIG. III.36 – Variation relative du courant de Drain d'un NMOS $W/L = 10/0.13\mu m$ lors d'une stress HE mesurée en régime saturé ($V_{GS} = V_{DS} = V_{DD}/2$).

ii. Stress IB. Les échantillons ont été stressés dans la gamme de tensions de Drain comprises entre $1.75V$ et $2.5V$ comme précédemment. Un premier enseignement à tirer de cette série de vieillissements est la localisation de la dégradation. Elle semble être plutôt étalée vers le milieu du canal. En effet, sur la Fig. III.37 on observe une faible différence entre les mesure du courant de Drain en régime linéaire ($V_{GS} = 0.6V$ et $V_{DS} = 25mV$) et saturé FWD ($V_{GS} = 0.6V$ et $V_{DS} = 0.6V$). Cette faible dépendance en V_{DS} signifie que la ZCE ne masque pas la zone des défauts pour $V_{DS} = 0.6V$. En revanche si on augmente la tension de Drain pour la mesure, la zone de défauts d'interface est en partie masquée et le courant de Drain semble moins affecté par la dégradation.

On observe donc un comportement similaire à ce que nous avons vu dans le cas HE (Fig. III.35). En effet on peut noter les points suivants :

- on observe les mêmes effets vis à vis des conditions de mesure V_{GS} , V_{DS} et du mode de mesure FWD ou REV ;
- on constate que les courbes de cinétiques présentent les mêmes pentes ($= 0.4 - 0.5$) dans les deux cas ;
- Un très faible décalage dans la base temporelle est observé.

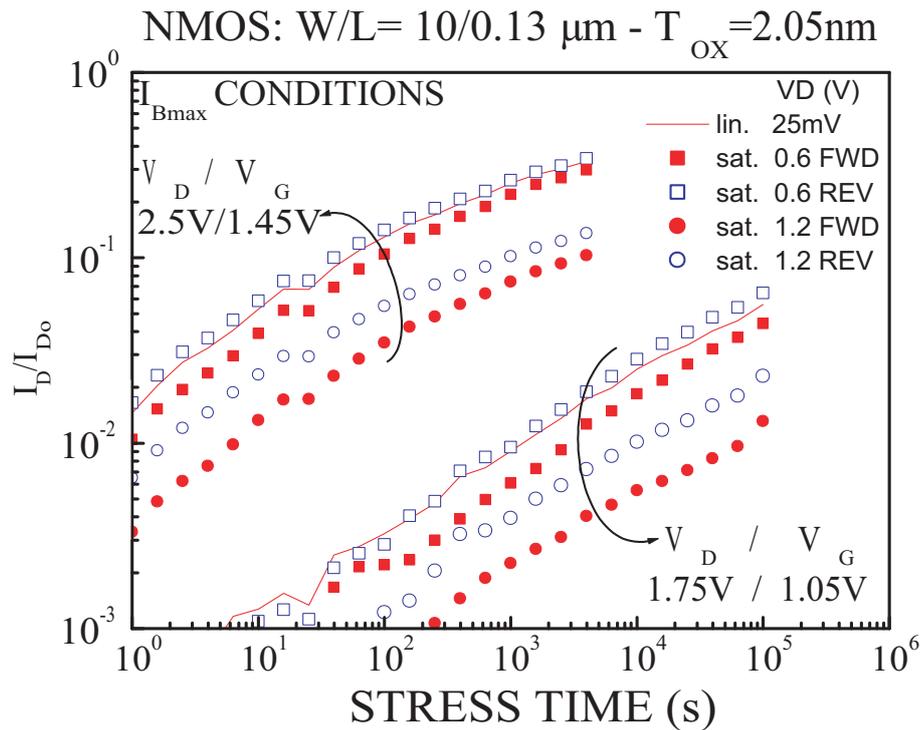


FIG. III.37 – Mesures à différentes polarisations de la dégradation du courant de Drain pendant l’injection de porteurs chauds au maximum du courant substrat sur des transistors NMOS de la technologie T3.

Ces trois points suggèrent que le même mécanisme de génération d’états d’interface intervient dans les cas de dégradations IB et HE, et que seul le nombre de défauts créés diffère.

iii. Stress VG+. Pour cette gamme d’oxyde de Grille ultra-mince, la probabilité pour les porteurs du canal d’être injectés en mode tunnel direct est importante. Le transport se fait en mode direct. Ainsi, des porteurs peu énergétiques vont pouvoir atteindre la Grille. C’est ce qui explique le très faible niveau de dégradation atteint sur les cinétiques présentées sur la Fig. III.38. On observe que malgré le fort champ appliqué ($V_{GS} = 3.5\text{V}$ ce qui correspond à un champ vertical F_{OX} proche de 15MV/cm), le courant est réduit de 10% qu’après 10^5s de stress. Pour les tensions plus faibles, même si le courant de Grille mesuré n’est pas négligeable, les porteurs n’ont pas suffisamment d’énergie pour générer des états d’interface. Notons enfin que les cinétiques FWD et REV sont identiques du fait de l’uniformité de la dégradation. Il est important de pouvoir distinguer la nature des défauts induits par le stress. L’analyse par mesure CP (Fig. III.39) a été effectuée pour la tension de stress $V_{GS} = 3\text{V}$. Les résultats montrent que le nombre d’états d’interface est faible pour la tension de stress $V_{GS} = 3\text{V}$ puisque la variation du courant pompé ΔI_{CP} vaut seulement 17pA . Cette augmentation du courant pompé correspond à une augmentation du nombre d’états d’interface relativement faible : $\Delta N_{it} = 7.08 \times 10^9\text{cm}^{-2}$. Cette valeur paraît d’autant plus faible que le courant tunnel dans le NMOS est important :

les électrons sont donc ici moins efficaces à générer des défauts que dans les technologies précédentes. Par ailleurs on ne note pas de décalage notable des courbes dérivées, ce qui signifie que la charge piégée est très faible : $\Delta V_{CP} = 40mV$. Cette valeur correspond à la limite de sensibilité de la mesure, dépendante de l'intervalle en V_{GS} , ce qui suggère que la finesse de l'oxyde rend impossible la mesure d'un éventuel piégeage de charges. Notons enfin que les courbes ne saturent pas parfaitement à cause du courant tunnel direct important qui se superpose au courant CP pour cette gamme d'épaisseur d'oxyde. Ceci n'influe pas sur l'amplitude de ΔI_{CP} , puisque la part du courant tunnel semble ne pas varier dans le temps, étant donné que les écarts des plateaux restent constant (les plateaux de saturation restent parallèles).

Dans cette partie nous avons fait l'étude des transistors nMOSFET de la technologie T3. Elle a mis en avant que pour cette gamme d'oxydes ultra-fins, le pire cas d'injection de porteurs chauds est la polarisation $V_{GS} = V_{DS}$, qui présente par ailleurs un mécanisme de dégradation très proche de la dégradation observée au maximum du courant substrat. La Fig. III.40 montre les taux de génération d'états d'interface dans les trois types de stress HE, IB et VG+. Les injections uniformes VG+ ont montré un taux de génération de défauts beaucoup plus faible même à champ vertical élevé. Ceci témoigne que le piégeage de charge dans l'oxyde n'est plus un mécanisme dominant dans le vieillissement de la structure, ceci a pu être démontré par les mesures CP. .

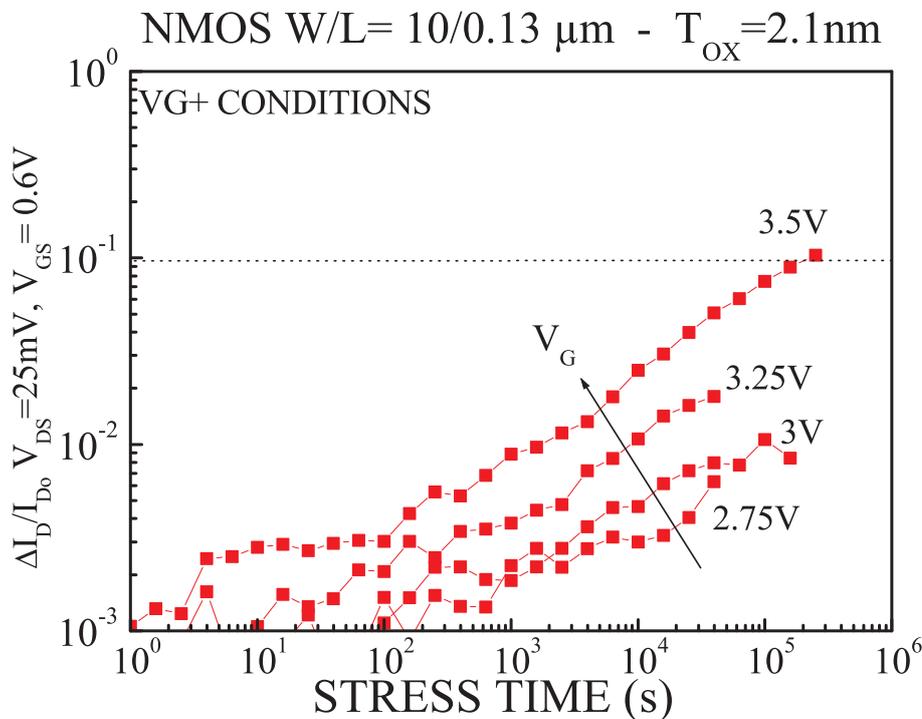


FIG. III.38 – Cinétique de réduction du courant de Drain pendant l'injection uniforme sur les transistors NMOS de la technologie T1.

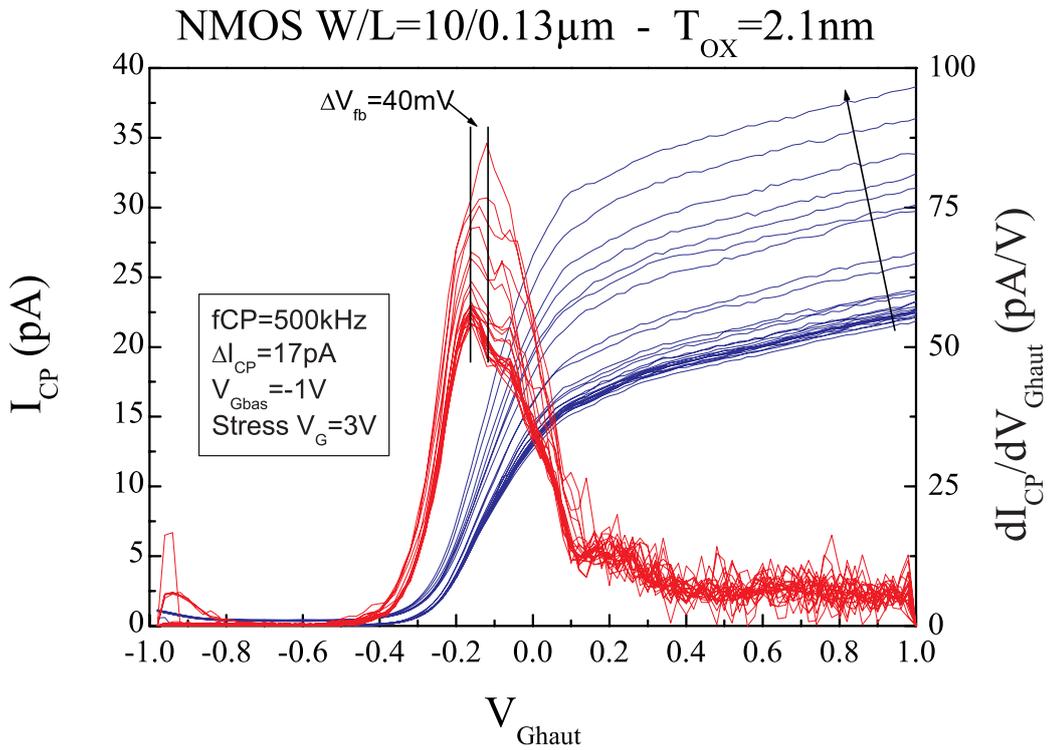


FIG. III.39 – Mesures par pompage de charges pendant le stress VG+ sur transistor NMOS de la technologie T3.

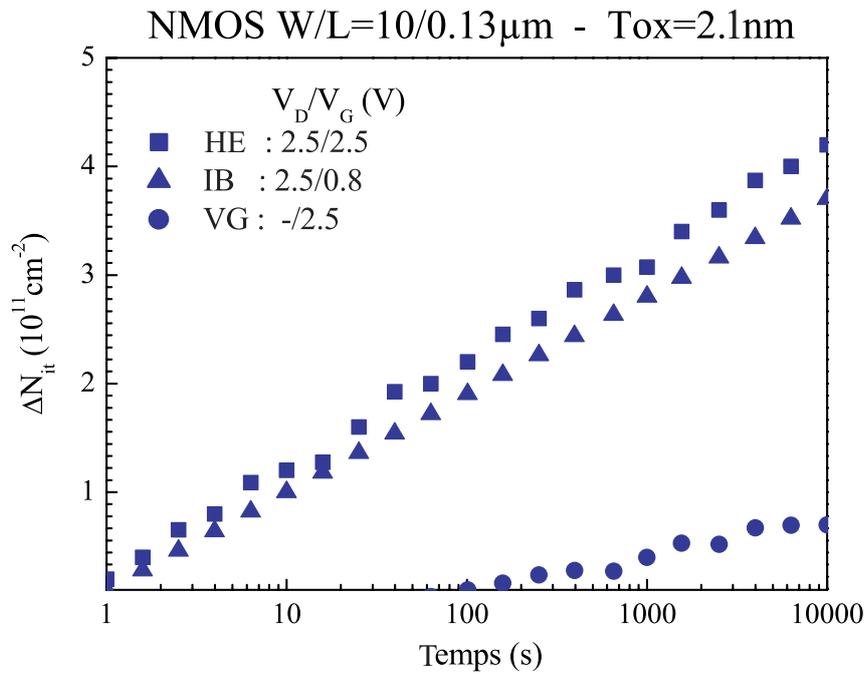


FIG. III.40 – Evolution de la quantité d'états d'interface au cours des stress HE, IB et VG+.

b) Etude des transistors PMOS.

Nous présentons dans cette partie l'étude des transistors PMOS à oxyde de Grille ultra-mince. Comme pour les transistors à canal n nous avons étudié le cas des injections uniformes. Nous allons voir que dans le cas du PMOS le courant tunnel (et donc l'injection de porteurs peu énergétiques) joue un rôle dans le vieillissement de la structure plus important que dans le cas du transistor NMOS.

Nous avons soumis des transistors PMOS de la technologie T3/T4, de longueur de Grille L comprise entre $0.1\mu m$ et $0.4\mu m$, aux stress de type :

1. Injections de trous chauds (HH) ;
2. Au maximum de courant substrat (IB) ;
3. Uniformes (VG-) à $V_{DS} = 0V$, correspondant à des injections de trous du canal.

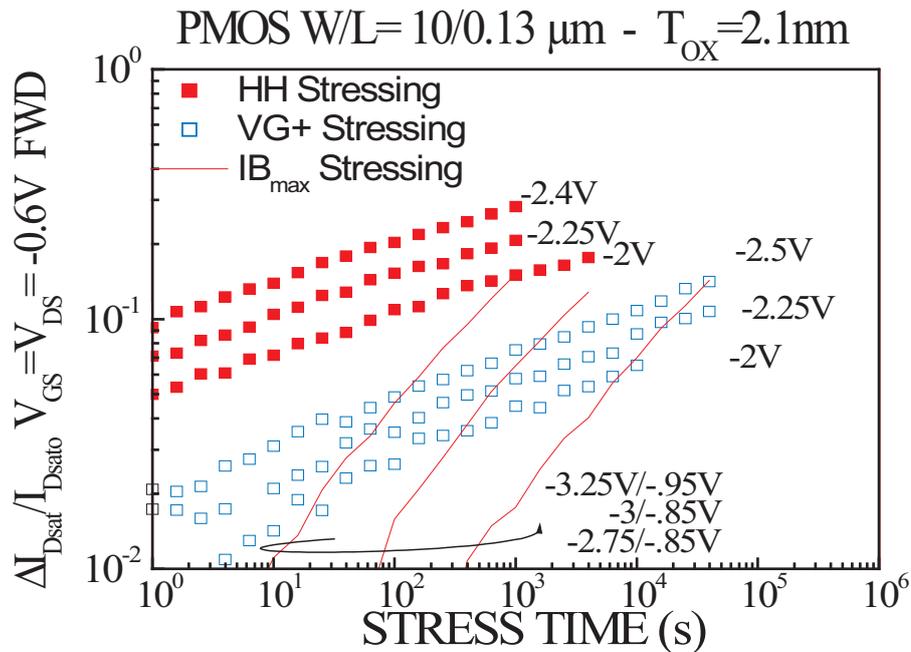


FIG. III.41 – Comparaison de la variation du courant de Drain pendant les trois types de stress HH, VG+ et IB mesurée en mode saturé à $(V_{DD}/2 - V_{DD}/2)$.

Les cinétiques des trois cas de stress standards sont réunis sur un même graphique (Fig. III.41) [31]. De cette figure se dégagent deux points importants :

- On peut observer que le mode HH est sans ambiguïté le pire cas de stress, ce qui a déjà été observé pour l'oxyde de $6.5nm$, mais de manière moins marquée (Fig. III.28). Ceci semble donc être propre à la technologie MDD. De plus, il présente les mêmes facteurs d'accélération (pente) que le mode VG+ mais à un niveau de dégradation plus faible [32] ;

- en revanche le mode IB est responsable d'un facteur d'accélération plus faible et nécessite des tensions de Drain plus fortes, si bien que l'on atteint avec un écart de deux décades en temps les 10% de dégradation de $\Delta I_{Dsat}/I_{Dsat0}$.

Cette hiérarchie dans le niveau de dégradation se retrouve dans l'étude de l'influence de la tension de Grille sur le type de mécanisme de génération de défauts. La Fig. III.42 montre le niveau de dégradation du courant de Drain (régime linéaire) au bout de $t = 4000s$ en fonction de la tension de Grille pour $V_{DS} = -2V$: le stress HH est bien le pire cas de polarisation. Nous allons maintenant détailler les mécanismes qui entrent en jeu dans chacun de ces trois cas de dégradation.

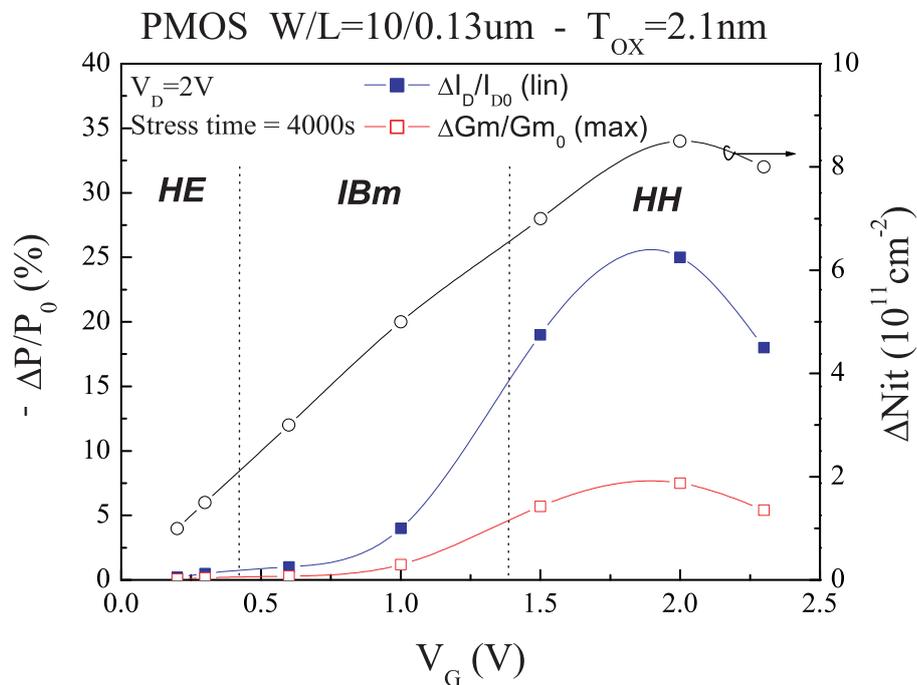


FIG. III.42 – Influence de la tension de Grille sur la dégradation du courant de Drain (linéaire) des transistors PMOS de la technologie T3 pour $V_{DS} = -2.5V$.

i. Stress HH. Nous avons soumis les transistors PMOS à oxyde ultra-mince au stress HH, qui se réalise à la condition $V_{GS} = V_{DS}$. Nous avons déterminé qu'il s'agit du pire cas de dégradation pour ce type de dispositifs. Les résultats marquants de cette première série de stress sont :

1. le niveau de dégradation est atteint dès la première seconde, preuve supplémentaire que la condition HH est le pire cas de dégradation ;
2. La dégradation reste visible et mesurable même à la tension d'alimentation V_{DD} .

La dépendance en V_{DS} de mesure est faible pour autant que la tension de Grille vaut $-0.6V$. La réduction du courant de Drain est donc ici attribuable à une réduction de la mobilité prononcée. D'autre part, cette absence de différences marquées entre les régimes saturé ($V_{DD}/2$) et linéaire, nous donne une première information sur l'étendue de la dégradation, qui semble donc s'étaler assez largement dans le canal. En augmentant les polarisations de mesure jusqu'à V_{DD} , la dégradation est réduite de pratiquement un facteur 10, ce qui est dû au déplacement du point de pincement du canal vers la Source. Considérons maintenant les dégradations mesurées en mode FWD et REV. La comparaison (Fig. III.43) des deux modes de mesures ne présente pas

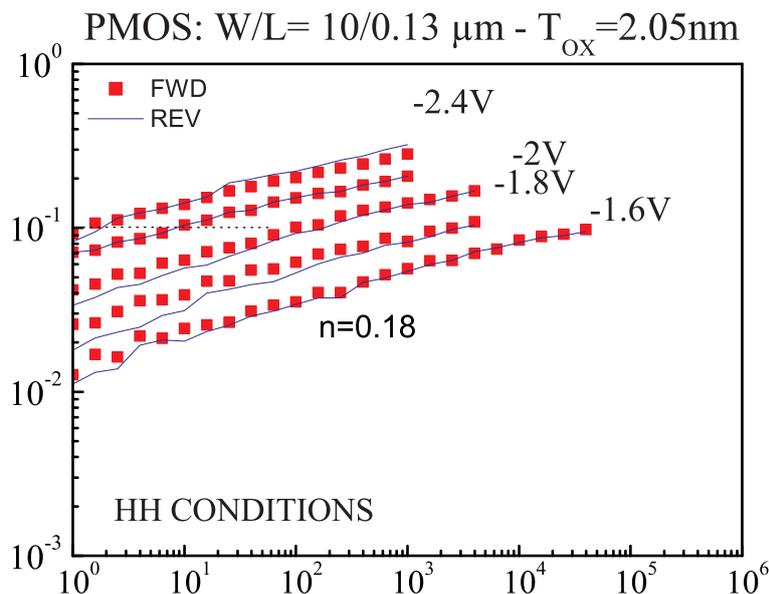


FIG. III.43 – Cinétiques de dégradation du courant de Drain mesuré en régime saturé à $V_{GS} = V_{DS} = V_{DD}/2$ en modes direct et inverse.

de différence, exactement comme si le dispositif avait été vieilli par une condition d'injection uniforme. Cette similitude s'obtient quelles que soient les conditions de mesure vis à vis des tensions de Drain et de Grille. Cette uniformité de la dégradation va pourtant à l'encontre du

stress HH, où l'on sait que les porteurs chauds sont générés près du Drain. Des porteurs sont donc injectés du côté de la source, où $\phi_C(x = 0)$, le potentiel au point x du canal, vaut V_{Th} . Par conséquent, dans cette zone le champ électrique présente une composante verticale importante comparée à sa composante latérale. Des porteurs faiblement énergétiques sont injectés dans l'oxyde donnant naissance à courant tunnel direct. Or nous avons vu sur la Fig. III.41 que ce type d'injection sous des tensions de Grille équivalentes sont génératrices de dégradations non négligeables (10% de ΔI_{DS}). Ces résultats sont confirmés par les mesures de pompage de charges pour la tension de Drain de stress $V_{DS} = -2V$ (Fig. III.44).

La dérivée de ces courbes ne permet pas de détecter de façon significative des charges piégées dans l'oxyde. En effet la technique se trouve limitée par l'importance du courant tunnel superposé au courant pompé. Ceci met en évidence les difficultés de la mise en œuvre de la mesure CP lorsque l'oxyde est très mince. La détermination de la charge piégée est limitée par le pas de la mesure. Typiquement nous avons choisi un pas de $50mV$ qui définit une limite de détection :

$$\Delta N_{ot} = \frac{C_{ox}}{q} \Delta V_{CP} \approx 5 \times 10^{11} cm^{-2} \quad (\text{III.12})$$

avec $C_{ox} = 1.64 \times 10^{-6} F/cm$. Cette valeur est importante, et pourrait être abaissée par la réduction du pas de mesure CP. Cependant, il faut ici considérer un autre facteur : les basculements successifs entre accumulation et inversion entraînent le vidage de la charge d'oxyde éventuelle (N_{ot}^- , N_{ot}^+) sur des pièges en volumes ou proches de l'interface. De plus la détection de la charge N_{ot}^+ requiert la mesure en $I_{CP} - V_{Ghaut}$, en polarisant vers l'accumulation où le courant de fuite DT est important. En effet, le cycle CP vers l'accumulation va décharger les trous piégés situés à la distance tunnel de l'interface, ce qui rend impossible leur détection au cours de la mesure, qui prend plusieurs secondes. Il convient donc de réduire au maximum le temps de mesure CP, et donc de minimiser le nombre de points de mesure. Il paraît ainsi raisonnable de faire l'hypothèse que seul les états d'interface sont créés durant le stress : les pics ne bougent pas (Fig. III.45).

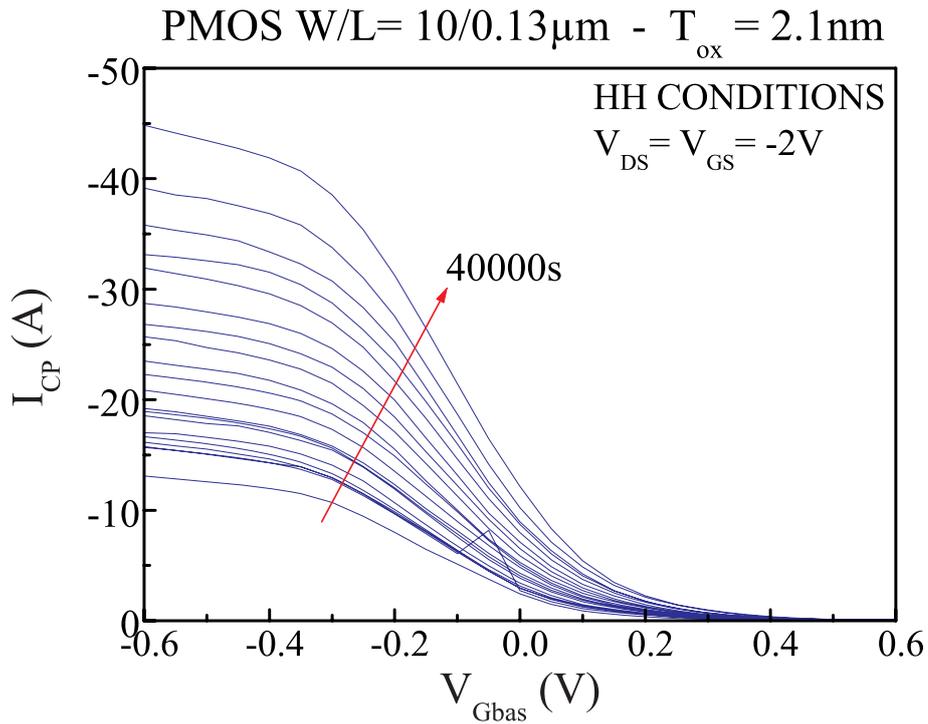


FIG. III.44 – Mesures de pompage de charges au cours d'un stress HH à $V_{DS} = -2V$ sur un transistor PMOS de la technologie T4.

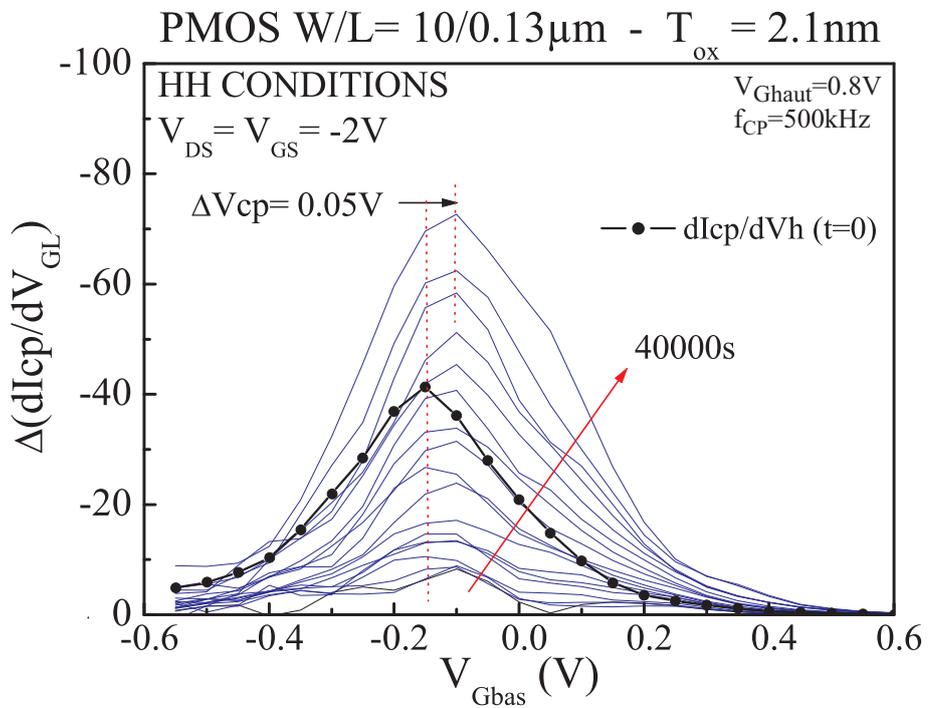


FIG. III.45 – Différence dans les dérivées des courbes I_{CP} obtenues pendant le stress HH sur un transistor PMOS de la technologie T4.

ii. **Stress IB.** Nous avons réalisé des stress au maximum de I_{BS} dans la gamme de tensions de Drain comprises entre $-2.2V$ et $-3.2V$, c'est à dire pour une condition proche de l'avalanche. Les cinétiques mesurées (Fig. III.46) en régime linéaire exhibent un retournement (illustré par le pointillé). A $t = 0s$ on note un pourcentage de dégradation sévère qui diminue dans un premier temps pour reprendre une augmentation "standard" en loi de puissance du temps, conséquence de la génération d'états d'interface. Ce comportement aux temps courts est imputable à un piégeage d'électrons dans l'oxyde. Rappelons que dans le cas du maximum de I_{BS} , des électrons et des trous chauds sont générés conjointement dans la zone où le champ latéral est maximal. Ainsi une partie de ces électrons injectés sont piégés dans l'oxyde, comme discuté pour l'oxyde GO2 (T2). Les autres sont collectés à la Grille comme le montre les mesures de $I_{G,e}$ (Fig. III.32) dans la gamme de tensions correspondantes aux pics de I_{BS} (Fig. III.31).

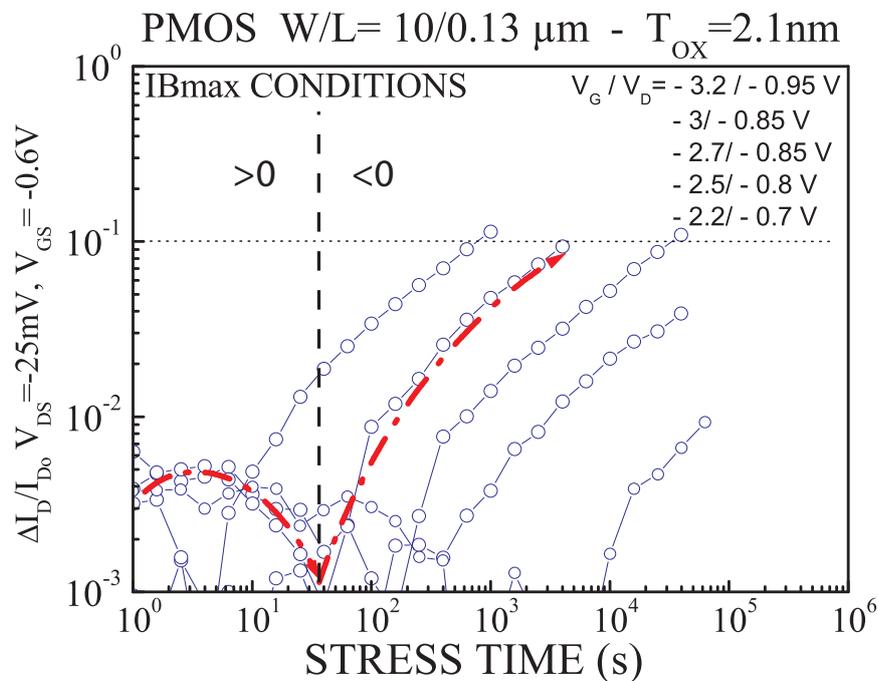


FIG. III.46 – Cinétiques de stress au maximum du courant substrat pour les transistors PMOS de la technologie T3, mesurées à différentes polarisations (régimes linéaire et saturé); le retournement est illustré par le trait pointillé épais.

Le point de retournement est d'autant plus long à apparaître que la tension de Drain pendant le stress est élevée, montrant l'augmentation du piégeage d'électrons avec V_{DS} dans la première phase du stress. La charge négative piégée a pour effet de réduire la longueur de canal, déplaçant de fait le pic de champ latéral vers le canal, impliquant une diminution de la charge piégée. Le retournement se produit quand la génération N_{it} domine le piégeage de la charge dans l'oxyde. Autre point important : on observe sur la Fig. III.47 l'influence des conditions

de mesure. Les modes direct et inverse ne présentent que très peu de différence, confirmant l'étalement prononcé de la dégradation le long du canal. L'influence de la tension de Drain a un comportement en cloche : à faible V_{DS} le courant de Drain est moins dégradé qu'à $V_{DD}/2$, et pour V_{DD} la diminution du courant de Drain est la plus faible.

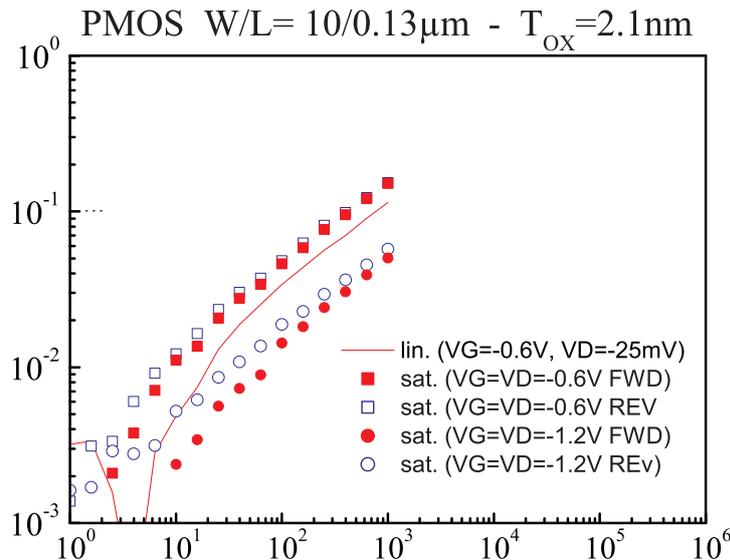


FIG. III.47 – Variation relative du courant de Drain durant un stress au maximum du courant Substrat mesurée en régime linéaire pour $V_{DS} = -25mV$ et $V_{GS} = -0.6V$.

iii. Stress VG- Nous avons étudié l'impact de l'injection uniforme de trous à travers l'oxyde sur une gamme de tensions de Grille comprise entre $-1.6V$ à $-2.25V$. La Fig. III.48 montre pour la tension de stress $V_{GS} = -2.25V$ que l'on atteint un niveau de dégradation supérieur à celui rencontré dans le NMOS à des tensions plus élevées. Il faut noter que le transistor PMOS n'a pas montré la même endurance que le NMOS vis à vis du stress uniforme : le claquage de l'oxyde intervient plus tôt que pour le NMOS, si bien qu'il ne nous a pas été possible de stresser les PMOS au delà de $-2.8V$, montrant une plus forte efficacité des trous à générer les défauts. Les 10% de réduction de $|I_{DS}|$ sont atteints en 40000s. Les régimes linéaire et saturé mesurés à $V_{GS} = -0.6V$ donnent les mêmes cinétiques, ce qui implique une forte réduction de mobilité. Logiquement, les mesures FWD et REV ne présentent pas de différence notable. En effet, l'injection est uniforme dans le cas VG-, et donc la dégradation également. Lorsque l'extension de la ZCE est maximale, lors de la mesure le courant semble moins réduit, à cause du phénomène d'écrantage d'une partie des défauts.

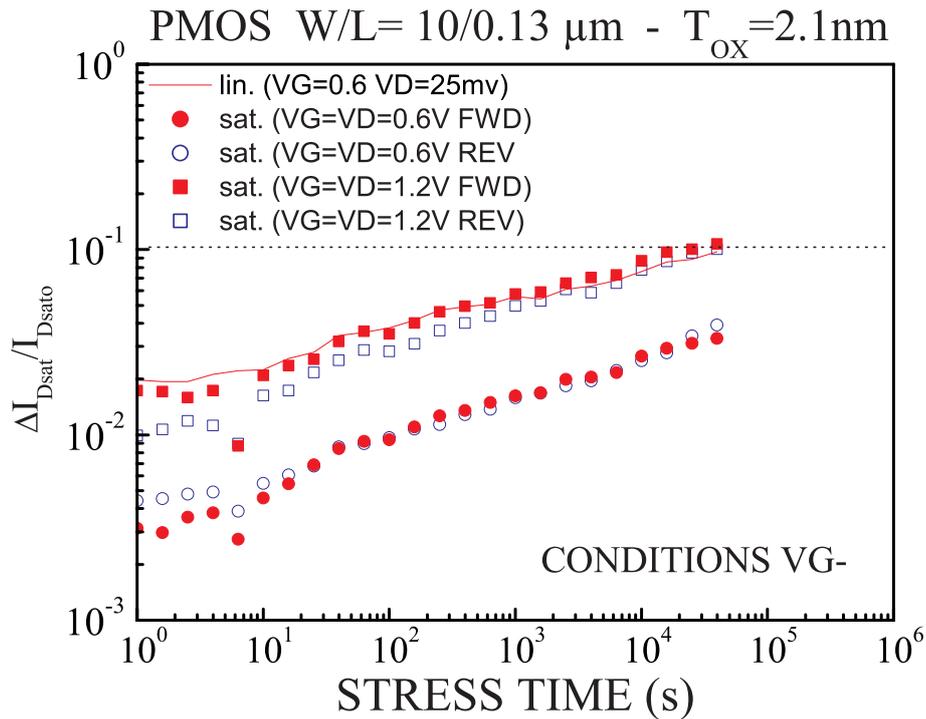


FIG. III.48 – Dégradation du courant au cours d'injections uniformes de trous sur des PMOS de la technologie T3 mesuré dans plusieurs polarisations (régimes linéaire et saturé).

c) Spécificité du cas HH dans le PMOS

Sur la Fig. III.49 on compare les courants de Grille mesurés à $V_{DS} = 0V$ et $V_{DS} = V_{GS}$, on observe un comportement similaire des courants I_G^{DT} (labels pleins) et I_G^{HH} (labels ajourés), avec une supériorité du courant tunnel quand $V_{DS} = 0V$. Ceci semble indiquer que très peu de porteurs chauds injectés atteignent la Grille. En revanche le courant substrat mesuré est plus important avec $V_{DS} = V_{GS}$ (HCI) que dans le cas où $V_{DS} = 0V$ (EVB). Le courant HCI est révélateur d'un fort taux de génération de porteurs chauds et le courant EVB est composé des électrons provenant de la bande de valence de la Grille par injection tunnel direct (DT).

Nous avons modélisé ce comportement en considérant une séparation entre les zones d'injections tunnels et d'injections de porteurs chauds. La première zone a pour longueur L_{tun} et la seconde L_{inj} et $L_{eff} = L_{tun} + L_{inj}$. Comme le montre la Fig. III.50, les dégradations générées par porteurs chauds s'étalent donc sur L_{inj} . Nous faisons ici l'hypothèse que la transition entre les deux modes d'injection est abrupte. Ceci nous permet d'écrire :

$$I_G^{HH} = I_G^{tun} + I_G^{inj} \quad (III.13)$$

où I_G^{tun} est le courant de trous provenant de la région de la Source et I_G^{inj} le courant dû à la fraction de trous chauds injectés qui ont pu atteindre la Grille. Notons alors α la proportion de

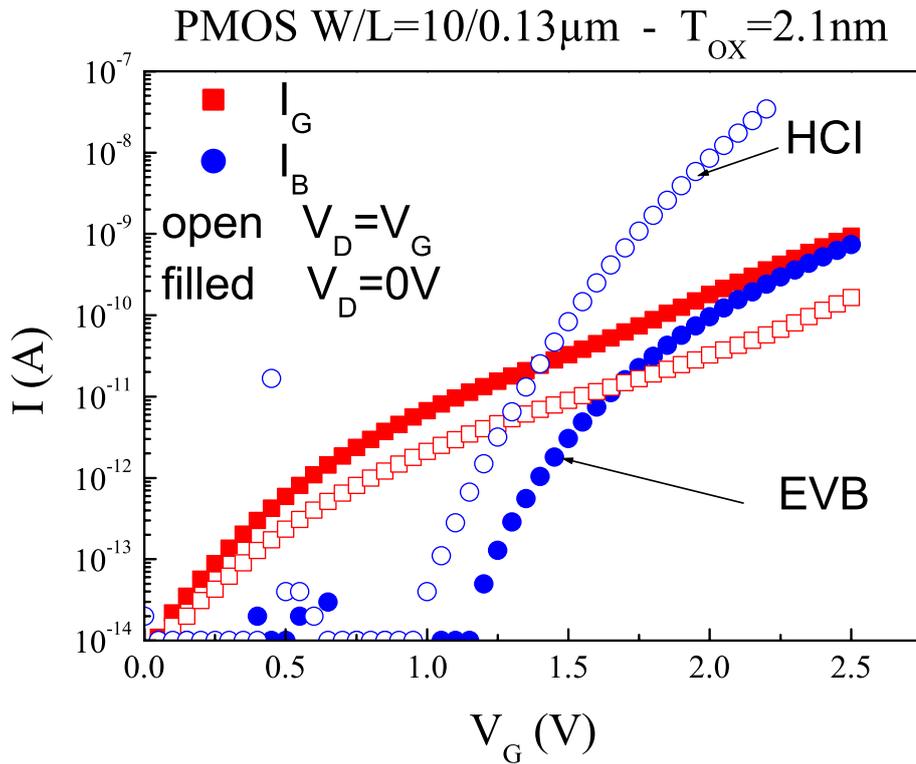


FIG. III.49 – Comparaison du courant de trous mesuré à la Grille pour $V_{GD} = 0V$ et $V_{GS} = V_{DS}$.

ces porteurs dans le courant mesuré I_G^{HH} , α est un paramètre dépendant de V_{DS} tel que l'on a :

$$I_G^{inj} = \alpha \times I_G^{HH} \quad ; \quad 0 < \alpha < 1 \quad (\text{III.14})$$

Ceci permet de réécrire la relation (III.13) indépendamment du courant de porteurs chauds :

$$I_G^{HH} = \frac{I_G^{tun}}{1 - \alpha} \quad (\text{III.15})$$

Pour comparer les modes d'injection VG- et HH, établissons le rapport des courants de Grille mesurés à $V_{DS} = 0V$ et $V_{GS} = V_{DS}$:

$$\frac{I_G^{HH}}{I_G^{DT}} = \frac{I_G^{tun}}{I_G^{DT}} \frac{1}{1 - \alpha} \quad (\text{III.16})$$

En raisonnant en terme de densité de courant, nous pouvons faire apparaître les surfaces des zones impliquées :

$$I_G^{tun} = J_G^{tun}(L_{tun} \times W_{eff}) \quad (III.17)$$

$$(III.18)$$

$$I_G^{DT} = J_G^{DT}(L_{eff} \times W_{eff}) \quad (III.19)$$

Nous avons supposé que les courants I_G^{tun} et I_G^{DT} sont de même nature, c'est à dire des courants tunnel direct, ce qui implique $J_G^{tun} = J_G^{DT}$. Le rapport de (III.17) et (III.19) se ramène donc simplement au rapport de leur surface d'injection : $L_{tun}W_{eff}/L_{eff}W_{eff}$. Nous pouvons finalement exprimer la relation entre les deux courants tunnels:

$$\frac{I_G^{HH}}{I_G^{DT}} = \frac{L_{tun}}{L_{eff}} \frac{1}{1 - \alpha} \quad (III.20)$$

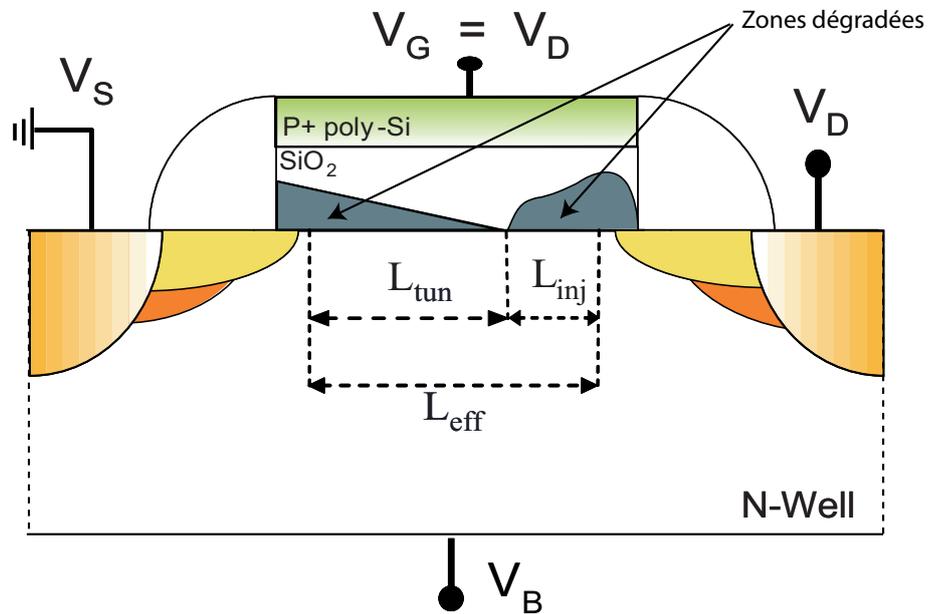


FIG. III.50 – Illustration de la représentation des courants de trous injectés dans la Grille durant les stress HH.

Nous faisons alors comme seconde hypothèse que la zone d'injection de trous chauds varie comme la longueur de la ZCE, que nous avons définie au premier chapitre (I.3.4) par :

$$L_{inj} = \Delta L_{eff} = \left[l \ln \left(\frac{(V_D - V_{def})/l + E_m}{\xi_{sat}} \right) \right] \quad (\text{III.21})$$

où δ_s est un paramètre fixé, qui permet d'ajuster la transition entre les régimes linéaire et saturé et V_{Def} est donnée par la relation [33] :

$$V_{Def} = V_{Dsat} - \frac{1}{2} \left[V_{Dsat} - V_D - \delta_s + \sqrt{(V_{Dsat} - V_D - \delta_s)^2 + 4\delta_s V_{Dsat}} \right] \quad (\text{III.22})$$

La combinaison de (III.16) et (III.21) nous permet finalement d'obtenir l'expression de séparation des courants :

$$\frac{I_G^{HH}}{I_G^{DT}} = \frac{1 - \frac{\Delta L}{L_{eff}}}{1 - \alpha} \quad (\text{III.23})$$

A l'aide des courbes de Fig. III.51, la valeur de α a été déterminée de manière à superposer les courbes expérimentales et simulées. Chaque courbe est obtenue pour une valeur du ratio V_{GS}/V_{DS} différente :

- quand $V_{GS}/V_{DS} = 1$, la structure est en mode saturé puisque $V_{DS} \geq V_{Dsat} \approx (V_{GS} - V_{Th})/\alpha_P$ et le canal est pincé ;
- pour 0.5 le point de pincement apparaît plus tôt avec les tensions négatives, pour $V_{GS} \approx 0.3V$;

Notons que $\Delta L(V_{DS})$ (qui est la zone d'injection des trous chauds) montre un bon accord avec (III.21) et (III.23). Le courant de trous chauds est plus faible que le courant tunnel. La détermination de α a montré une faible dépendance avec V_{DS} : on trouve $0.38 < \alpha < 0.42$. Remarque : ce modèle est simplifié car les deux populations de porteurs sont impliqués dans une zone intermédiaire où notre modèle n'est pas valable, du fait que l'on peut avoir un recouvrement des zones d'injections. Ceci justifie l'emploi du terme α qui correspond à cette zone de transition, où les deux populations de porteurs coexistent.

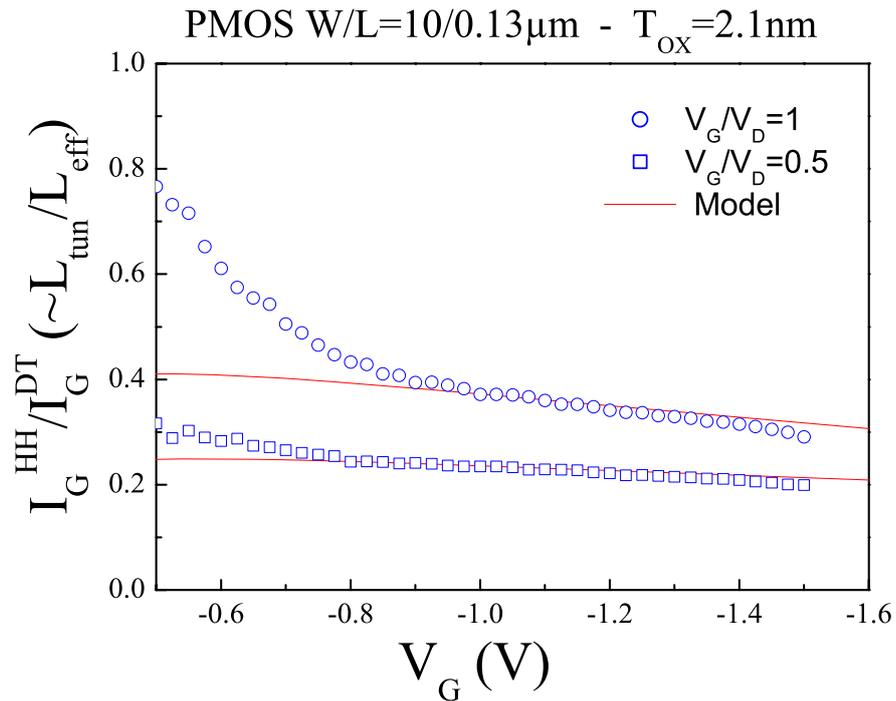


FIG. III.51 – Comparaison des rapports des courant de Grille (mesurés à $V_{DS} = 0V$ et $V_{GS} = V_{DS}$) et des surfaces d'injections (de trous chauds au Drain et froids à la source).

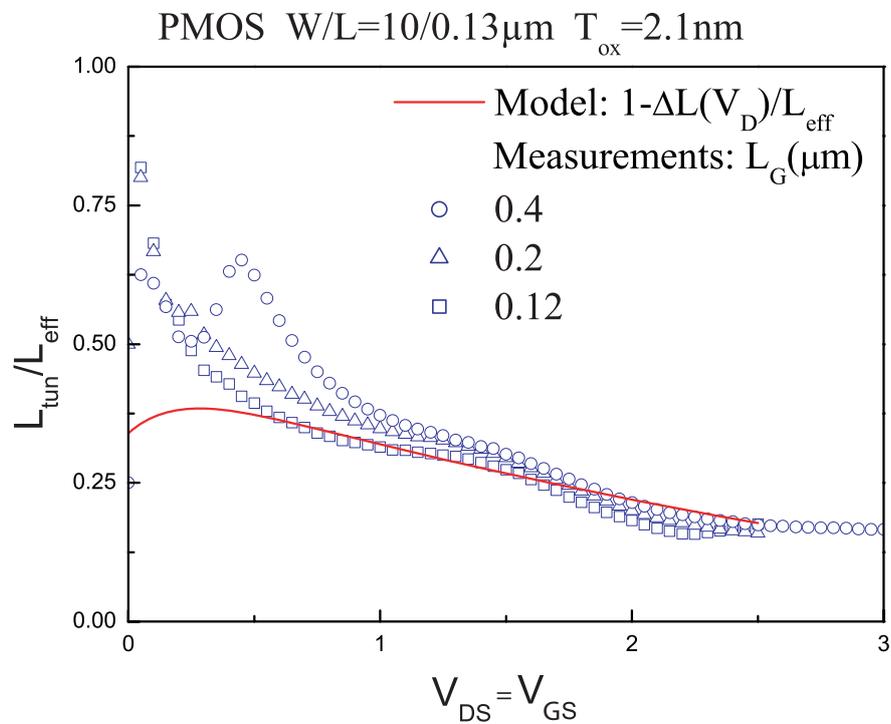


FIG. III.52 – Représentation du rapport des courants I_G^{HH} et I_G^{DT} (permettant d'évaluer $L_{\text{tun}}/L_{\text{eff}}$ à l'aide de (III.23)) pour différentes longueurs de canal pour la polarisation $V_{GS} = V_{DS}$.

Il est également intéressant de constater sur la Fig. III.52 que le rapport des courants de Grille est indépendant de sa longueur pour la polarisation $V_{GS} = V_{DS}$. La répartition des mécanismes semble donc indépendante de la longueur du canal. Ce modèle de séparation nous sera utile dans le chapitre prochain pour évaluer le rôle de chaque mécanisme dans l'estimation de la durée du dispositif.

Dans le but de valider la prédominance de la composante tunnel dans la mesure de I_{GS} à $V_{GS} = V_{DS}$, nous avons calculé la densité de courant tunnel dans le cas HH. La Fig. III.53 montre qu'on retrouve la même densité de courant que dans le cas DT avec le calcul en $L_{tun} = L_{eff} - \Delta L$ pour la modélisation. La courbe simulée démarre pour $V_{GS} \geq V_{Th}$, lorsque les conditions de polarisation autorisent le mode saturé, c'est à dire quand le point de pincement commence à se déplacer vers la Source.

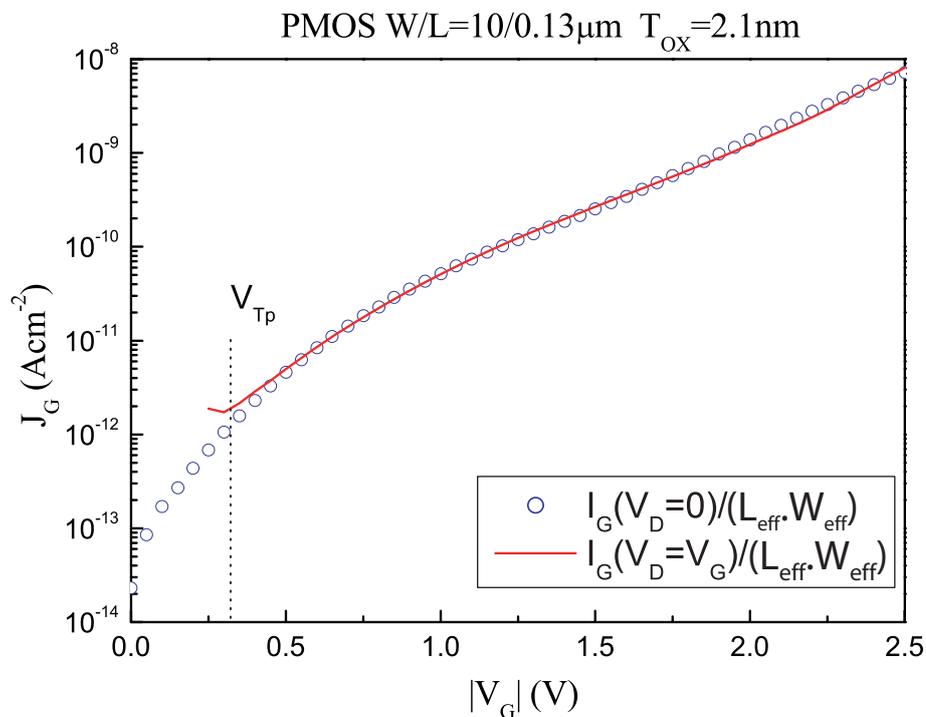


FIG. III.53 – Comparaison des densités de courant tunnel, calculée (en prenant I_{GS} pour $V_{GS} = V_{DS}$ ramené à la surface calculée $L_{tun} \times W_{eff}$) et mesurée (I_{GS} pour $V_{DS} = 0V$ ramené à la surface totale de la Grille).

d) Comparaison des dispositifs PMOS *Low Leakage* et *High Speed*

La différence majeure entre ces deux types de dispositifs réside dans la dose de dopant plus importante à la surface du canal pour les LL, dans le but d'augmenter V_{Th} et la résistance du canal. Ainsi la fonction "faibles fuites" est réalisée grâce à des valeurs de I_{off} inférieures dans les transistors LL, assurant une consommation statique moindre. Cette différence se retrouve dans le dopage moyen mesuré par les techniques utilisant l'effet substrat (I.4.2) : $N_D^{HS} = 4.8 \times$

10^{17}cm^{-3} et $N_D^{LL} = 1.5 \times 10^{18} \text{cm}^{-3}$.

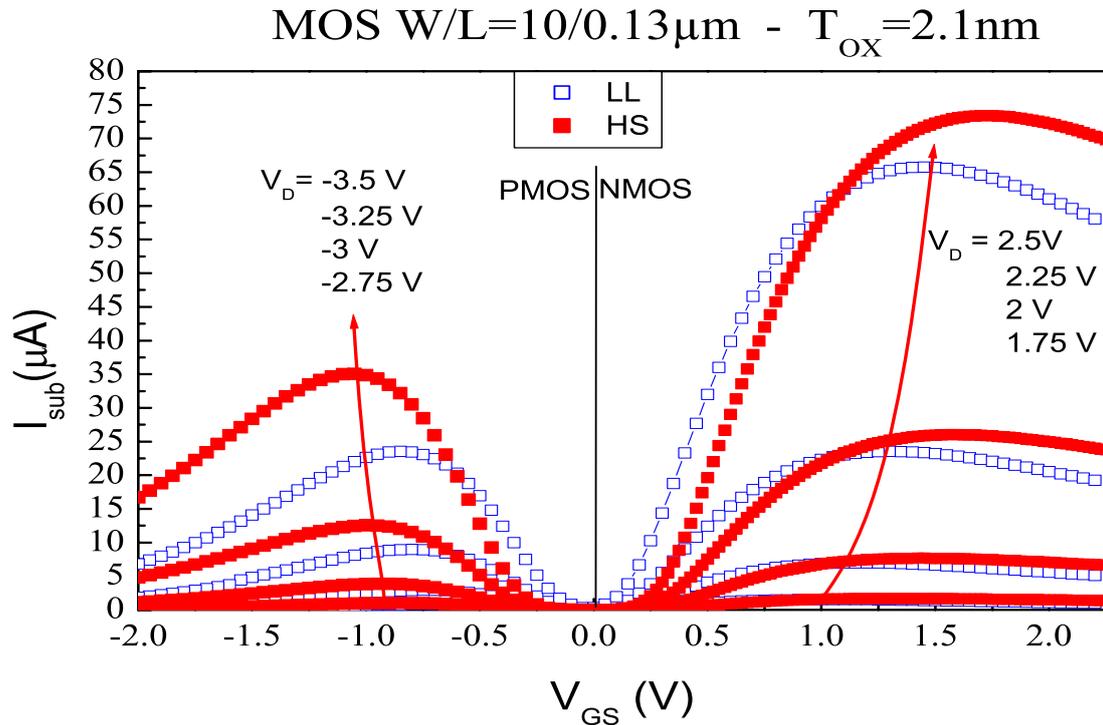


FIG. III.54 – Courbes de caractérisation du régime porteurs chauds pour la technologie T3 (filère 0.13 μm HS) pour les transistors NMOS et PMOS.

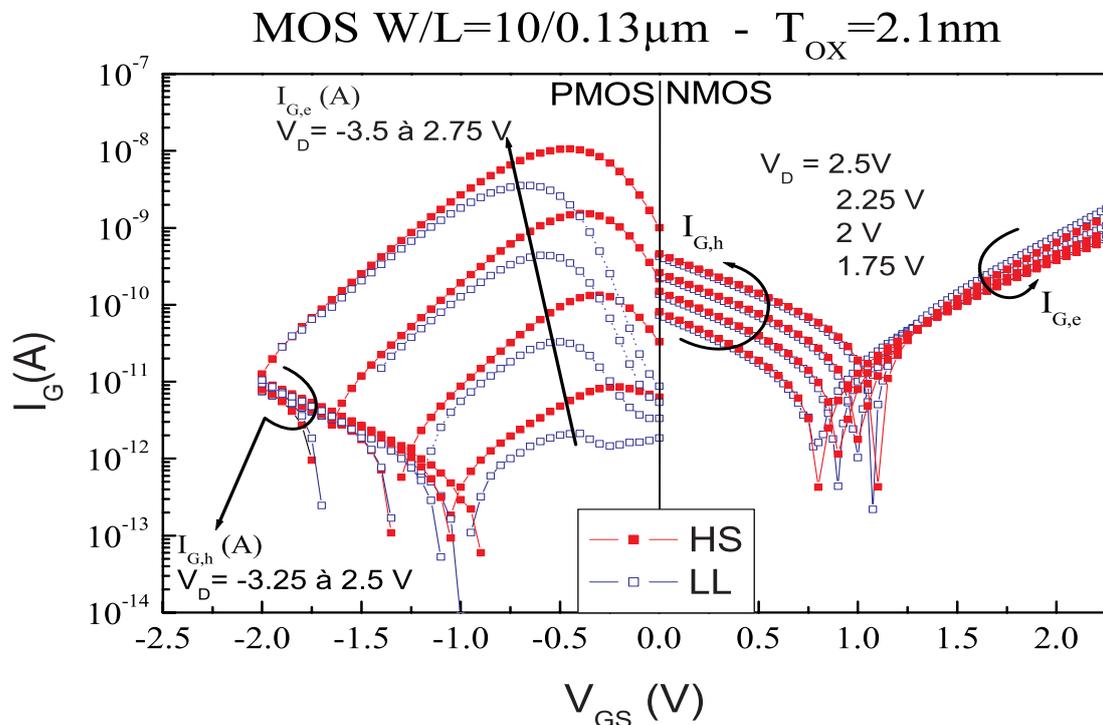


FIG. III.55 – Mesures du courant de Grille électronique pour les transistors PMOS de la technologie T3/T4 pour les deux types HS et LL.

La très forte valeur de dopage observée dans le LL est peu réaliste et suggère que la technique manque de sensibilité pour sonder le profil du dopage dans la profondeur du substrat. En effet la méthode qui utilise l'effet body (I.4.2), est plus sensible à la valeur en surface même si un éventuel changement de pente du tracé $\Delta V_{Th} = f(\sqrt{2\phi_F - V_{BS}})$ peut témoigner d'une forte variation de dopage en volume.

Les Fig. III.54, III.55 montrent respectivement les pics du courant substrat, et des courants de Grille sur les deux types de transistors des technologie T3/T4 : High Speed et Low Leakage. Les pics de courant substrat sont obtenus pour des ratio $V_{GS}/V_{DS} = 0.35$ pour les HS alors que pour les LL ce ratio s'établit autour de $0.22V$. De plus, les valeurs des pics de I_{BS} sont plus importants pour les HS, il en va donc de même pour les taux de génération de porteurs chauds. Dans le transistor PMOS, on observe que le courant de trous augmente avec les tensions de Grilles très négatives, avec une branche tunnel indépendante de $|V_{DS}|$ [32]. A bas $|V_{GS}|$ on retrouve le maximum de courant $I_{G,e}$ en corrélation au maximum de I_{BS} pour les dispositifs à bas $|V_{Th}|$. Les courbes de $I_{G,e}$ ont des formes en cloche avec des pics décalés entre les HS et les LL dû au décalage en V_{Th} . Dans le cas du NMOS, on observe la disparition des courbes en cloche dû au courant thermoionique, malgré la génération de porteurs chauds avec V_{DS} (III.54). On ne mesure plus que le courant tunnel direct Grille-Drain et un retournement. Cela est dû à la composante DT des électrons qui proviennent de la Grille dans la région Grille-Drain pour $V_{GS} \ll V_{DS}$, puis des électrons qui sont injectés du canal en mode DT pour $V_{GS} \geq V_{DS}$, correspondant à l'inversion du champ vertical.

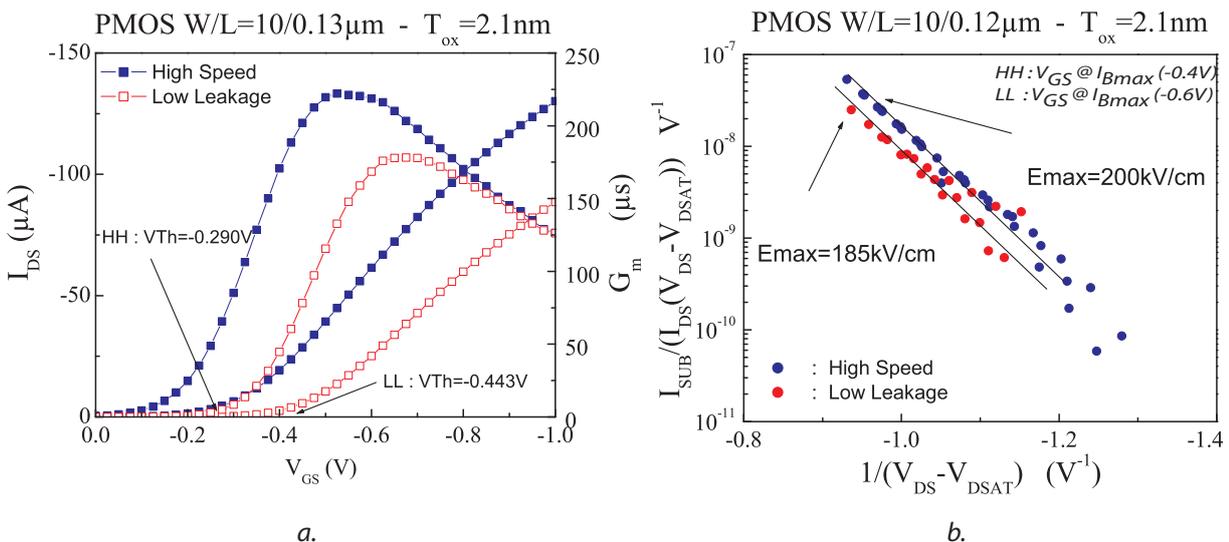


FIG. III.56 – a : Caractéristique $I_{DS} - V_{GS}$ et transconductance des dispositifs LL et HH. b : évaluation du taux d'ionisation pour extraire la valeur maximale du champ électrique latéral.

Concrètement l'augmentation de la tension de seuil des LL s'accompagne d'une nette dimi-

nution de transconductance (Fig. III.56.a). En revanche le champ électrique maximal varie peu (Fig. III.56.b) dans le PMOS, laissant présager des résultats similaires vis à vis de l'injection de porteurs chauds. Notons toutefois que les valeurs de pic de champs maximales ne sont pas obtenues pour les mêmes valeurs de la tension de Grille, et apparaissent avec des pics de courant substrat qui sont plus élevés pour le HS (Fig. III.54).

Les stress à la condition HE, ajustés aux pics de $I_{G,e}$ (c'est à dire pour chaque V_{GS} fixé au pic de $I_{G,e}$ et même V_{DS}), subis par les dispositifs PMOS HS et LL, nous ont permis d'observer sur la Fig. III.57 que ce mécanisme dans les transistors HS induit systématiquement un effet de perçage (punchthrough), c'est à dire un décalage positif net (lié au piégeage d'électrons) et un décalage négatif lié à la croissance rapide d'états d'interface donneurs chargés positivement. Le premier effet est connu comme l'effet "Hot Electron Induced Punchthrough" (HEIP) [34] aux champs verticaux élevés où la tension V_{PT} de perçage est abaissée avec la dose de dopant élevée de la région de LDD. Ceci mène par conséquent à un champ latéral élevé qui peut induire un piégeage important d'électrons, d'abord dans la région de recouvrement, avant que la région dégradée dans le canal actif prenne un rôle dominant avec l'augmentation du nombre d'états d'interface donneurs. La compétition entre ces deux mécanismes entraîne un effet de retournement comme nous l'avons souligné dans la section III.2.b.ii pour la condition IB. Ce phénomène en deux étapes ne s'observe pas dans les dispositifs LL [35] comme on peut le constater sur la Fig. III.57, démontrant l'absence de piégeage d'électrons au temps courts et la prédominance des états d'interfaces donneurs.

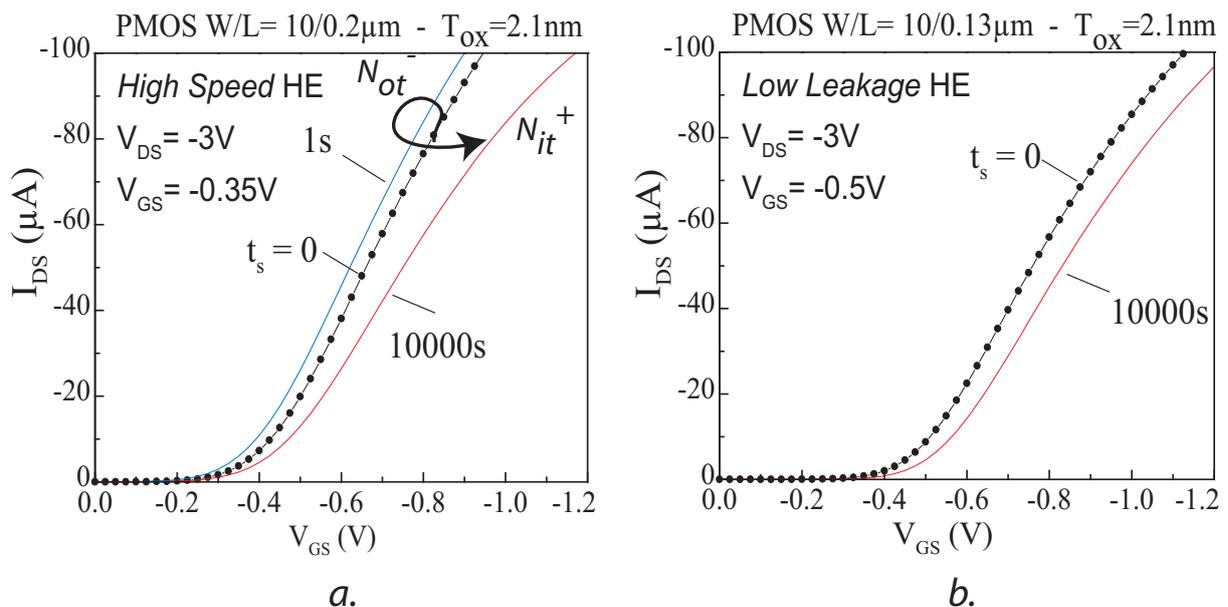


FIG. III.57 – Comparaison des effets de l'injection de porteurs chauds au maximum du courant de Grille électronique dans les dispositifs HS (a) et LL (b) [32].

Nous avons comparé l'effet de l'injection des trous chauds réalisée à $V_{DS} = V_{GS}$ et l'injection uniforme à $V_{DS} = 0V$ dans les deux types de structures HS et LL. Afin de représenter les résultats de manière plus synthétique nous les avons regroupés sur les Fig. III.59 et III.60. Concernant le stress HH, l'observation de ces résultats nous permet de conclure les points suivants :

- Le courant tunnel de "trous froids" de la bande de valence crée une dégradation observable sur les caractéristiques linéaires, c'est à dire à bas champ vertical : $F_{OX} = -5MV/cm$ ($= -6MV/cm$) dans des dispositifs HS (LL) ;
- L'extension de la région dégradée au-dessus du canal actif semble identique dans les deux types de dispositif comme le montre la même variation de ΔI_{DSat} FWD, et ΔI_{Dlin} pour les HS et LL (Fig. III.59) ;
- La légère différence observée entre les modes REV et FWD (Fig. III.60) montre que des dommages plus importants sont générés dans la zone de recouvrement Grille-Drain des dispositifs HS, car ΔV_{Th} est observé plus grand dans les dispositifs HS [32] en liaison avec le pic de champ latéral légèrement plus fort dans cette région (8%). Ceci peut avoir comme origine la dégradation de l'espaceur, pour cette condition de polarisation, dû en partie au champ latéral plus élevé dans cette zone pour le HS.

Ces observations permettent de clarifier les différences entre les deux structures. Si la valeur du pic de champ est plus grande dans les dispositifs HS, en revanche, la valeur maximale se situe plus en profondeur dans le substrat et sous la zone 2 (en référence à la Fig. II.1), alors que pour les structures LL, elle se situe plus près de l'interface et dans la zone 1 (Fig. III.58).

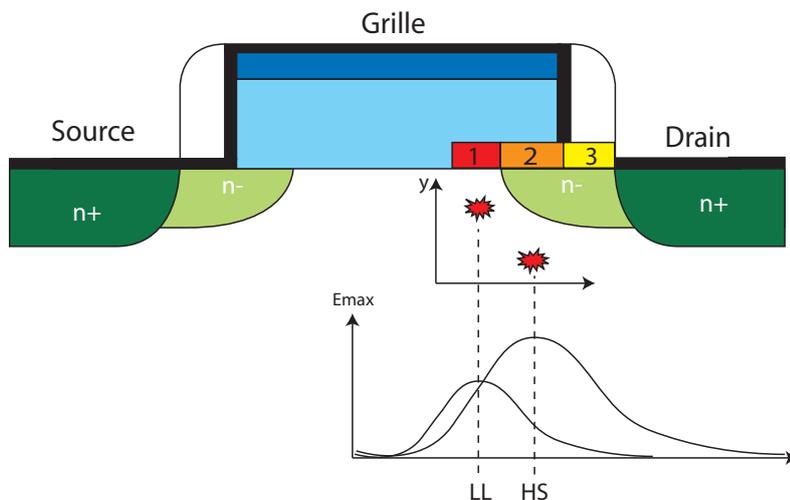


FIG. III.58 – Localisation des pics de champ dans les transistors PMOS des technologies T3 et T4.

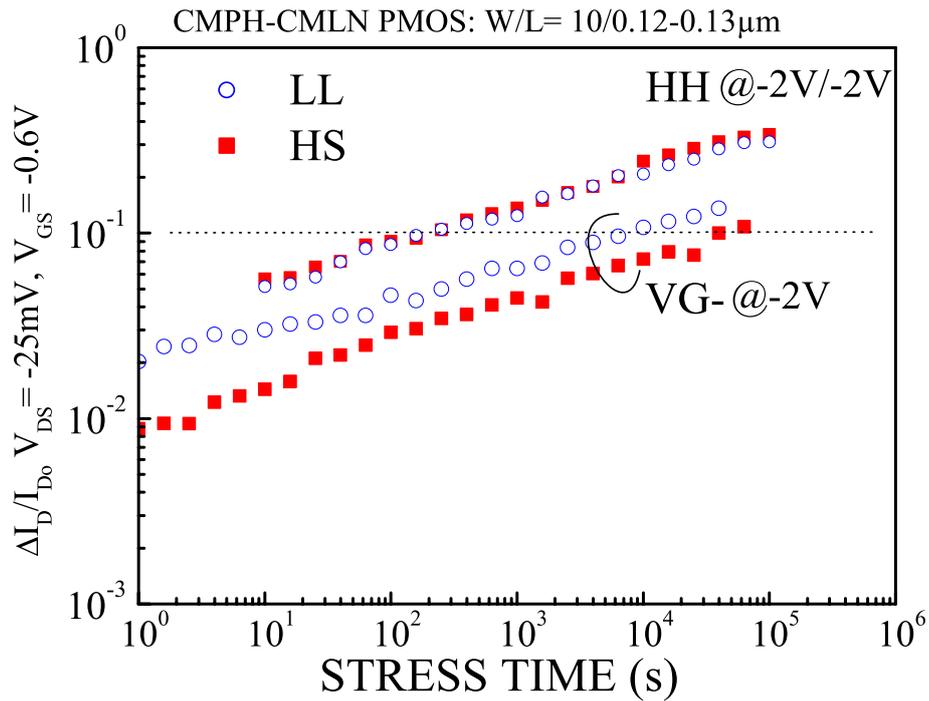


FIG. III.59 – Variation du courant de Drain pendant des stress HH et VG- sur les deux types de dispositifs HS (T3) et LL (T4).

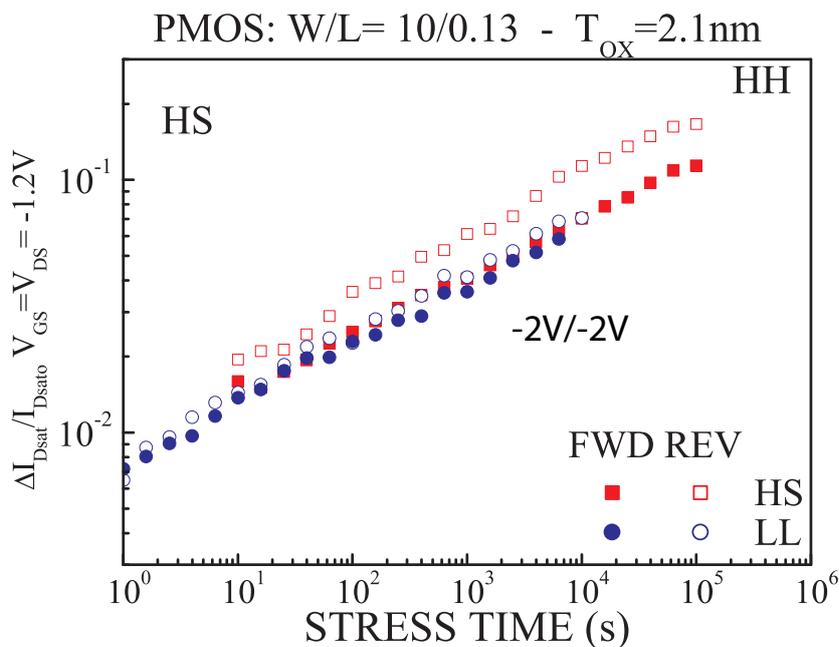


FIG. III.60 – Comparaison de la variation du courant de Drain pendant un stress HH sur les technologies T3 et T4.

III.3 Stress alternés

Nous appelons stress alternés, de courtes successions d'injections HH et HE. Nous avons utilisé ce type de manipulation dans le but de révéler l'existence de pièges neutres générés par l'injection de trous dans le transistor NMOS [36, 37]. D'autre part, cette technique a permis de mettre en évidence des pièges d'interface accepteurs chargés négativement. On distingue en effet dans la littérature des pièges lents et rapides [38]:

- les états lents sont des pièges de charges dans le volume de l'oxyde, mais à une distance proche de l'interface (dite longueur tunnel), pour qu'ils puissent communiquer avec la couche d'inversion par conduction tunnel [39];
- les pièges dits rapides correspondent aux pièges situés sur l'interface [40].

III.3.1 Influence des pièges neutres dans les oxydes épais

Nous avons étudié l'effet de l'alternance d'injections de trous chauds et d'électrons chauds sur transistor NMOS de la technologie T1. Nous avons commencé par une injection longue de trous chauds (phase 0 \rightarrow 1) réalisée pour $V_{DS} = 7.5V$ et $V_{GS} = 1.5V$ pendant 40000s suivie d'une injection courte (60s) d'électrons chauds (SEI) à $V_{GS} = V_{DS} = 7.5V$ (phase 1 \rightarrow 2) et d'une autre injection courte (60s) de trous chauds à $V_{DS} = 7.5V$ et $V_{GS} = 1.5V$ (phase 2 \rightarrow 3).

On observe sur la Fig. III.61 que cette succession d'injections, entraîne de nettes variations du courant de Drain en régime linéaire, avec des décalages des courbes non parallèles, donnant des variations de de la tension de seuil $\Delta V_{Th} = 0.3V$ puis $\Delta V_{Th} = -0.15V$. Rappelons que les décalages de la tension de seuil ont été décrits en II.1.2 par l'équation (II.5). Ceci indique donc une forte charge négative sur des pièges accepteurs [36, 37].

Par le passé, des preuves claires ont démontré les influences de la création des pièges d'électrons dans les structures MOS soumises à l'injection de trous chauds [36]. Ces défauts révélés par des injections de porteurs localisées ont été soigneusement mis en évidence par les dégradations des caractéristiques I-V et des simulations bidimensionnelles jusque dans la région de recouvrement Grille-Drain dans les transistors NMOS de type LDD grâce à l'emploi de la technique FG [37].

La Fig. III.61 illustre le comportement de type accepteur de ces pièges [36, 37] avec les courbes $I_{DS} - V_{GS}$ avant et après la contrainte HH (phase 0 \rightarrow 1) suivie de l'injection SEI qui décale la courbe vers les tensions de Grille positives (phase \rightarrow 2), alors qu'une injection SHI renvoie la caractéristique sur la courbe post-stress (phase 3). Différentes explications ont été données au sujet de ce comportement privilégiant la création de pièges neutres localisés au dessus de la zone recouvrement [41]. Des études plus récentes utilisant des expériences de charge

et de décharge des pièges [42] ont proposé à la fois la création de pièges fixes dans l'oxyde à quelques nanomètres de l'interface chargés négativement suite à l'injection des électrons chauds et de la neutralisation de la charge positive par les électrons dans une région plus étendue vers la Source.

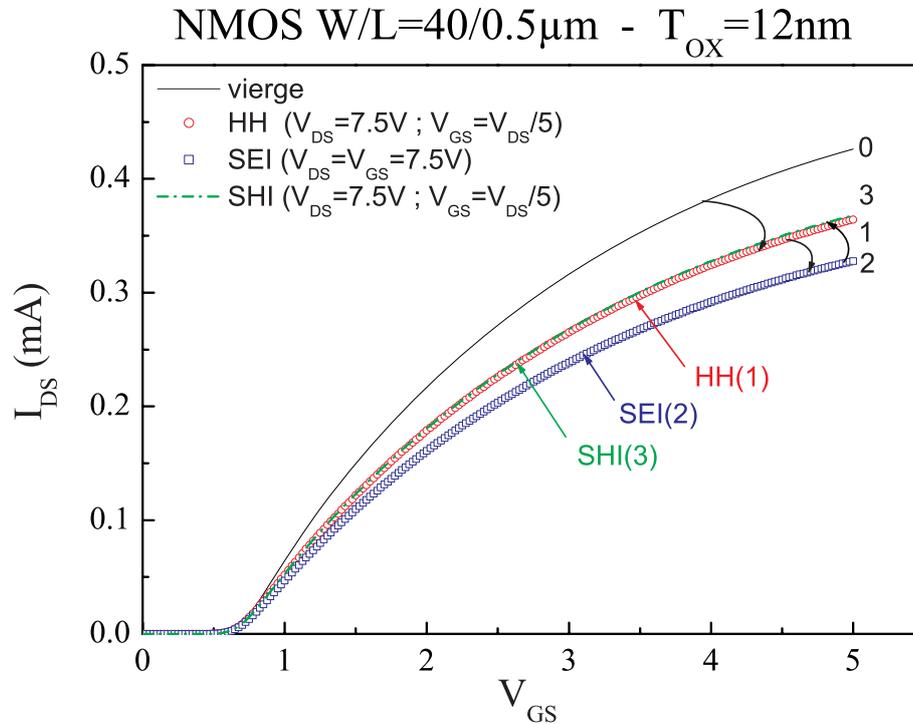


FIG. III.61 – Caractéristiques $I_{DS} - V_{GS}$ mesurées entre des alternances d'injections d'électrons chauds et de trous de chauds après un stress long (40000s) sur transistor NMOS de technologie T1.

III.3.2 Oxydes minces

Dans cette section nous avons réalisé le même type d'étude que sur les oxydes épais. Le dispositif à canal n a été soumis à une injection longue (40000s) de trous chauds (HH) effectuée à $V_{DS} = 2.2V$ et $V_{GS} = V_{DS}/5$ (phase 0 \rightarrow 1). Le transistor a ensuite subi deux injections rapides (60s) d'abord SEI à $V_{GS} = V_{DS} = 2.2V$ (phase 1 \rightarrow 2) puis SHI à $V_{DS} = 2.2V$ et $V_{GS} = V_{DS}/5$ (phase 2 \rightarrow 3). La Fig. III.62 montre les effets de cette séquence d'injections sur le courant de Drain linéaire. La phase HH dégrade nettement le courant (15% de réduction). Le passage à l'injection SEI aggrave encore la dégradation de 12%. Le basculement à l'injection SHI ne ramène pas la caractéristique $I_{DS} - V_{GS}$ comme dans le cas des oxydes épais. Ceci permet de conclure que l'injection HH ne crée plus de pièges neutres, mais seulement des états d'interfaces chargés négativement. En effet, on remarque que la phase SEI, de façon cumulative, a généré des états d'interface de même type mais qui ne sont pas neutralisables par l'injection SHI.

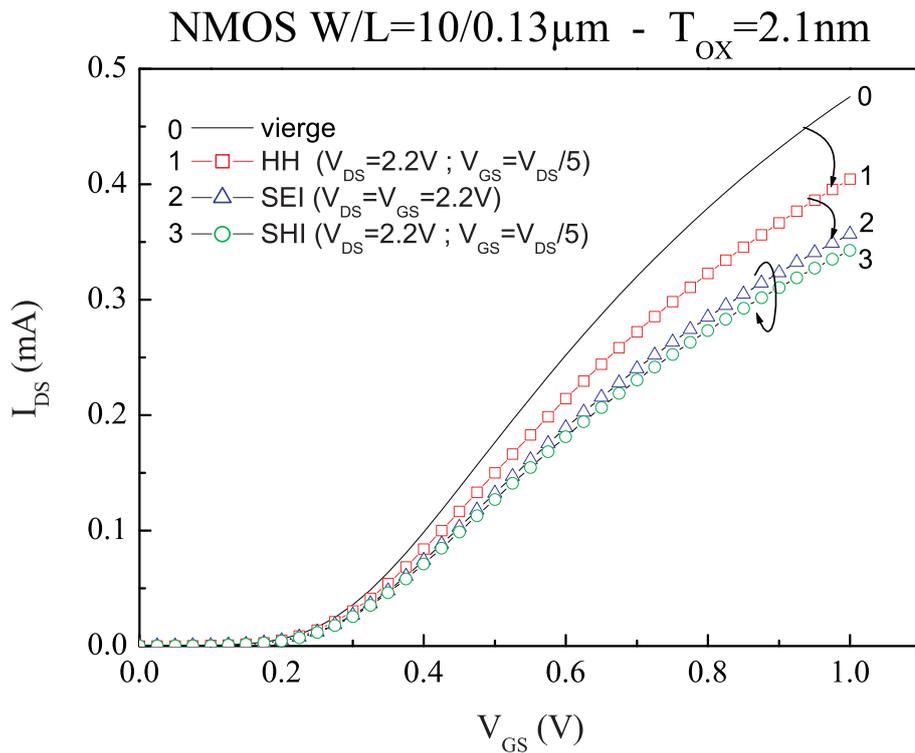


FIG. III.62 – Caractéristiques $I_{DS} - V_{GS}$ mesurées entre des alternances d'injections d'électrons chauds et de trous chauds après un stress long (40000s) sur transistor NMOS de la technologie T3.

Ce résultat est nouveau et met en évidence que l'emploi des oxydes minces nitrurés font disparaître l'influence des pièges neutres générés par les injections de trous chauds dans les transistors NMOS [43]. Dans les oxydes épais les pièges sont neutres et peuvent être remplis et vidés après le stress initial car ils sont situés à plusieurs nanomètres de l'interface sans effet possible du dépiégeage. Les pièges créés dans les oxydes minces sont de nature accepteurs et se vident très rapidement vers le substrat laissant seule l'influence des états d'interface.

On peut émettre une autre hypothèse : la disparition des pièges neutres peut s'expliquer par la diminution du nombre de trous chauds injectés. D'autre part la nitruration des oxydes de Grille implique une résistance plus élevée à la génération N_{it} . Enfin dans les oxydes minces, les pièges se déchargent plus rapidement car ils sont à une distance tunnel de l'interface.

III.4 Synthèse

Avant de conclure ce chapitre, faisons une synthèse sur l'évolution des mécanismes d'injections dans les structures MOSFETs. Pour les oxydes épais ($12nm$, technologie T1) le pire cas d'injection est la polarisation correspondante au maximum du courant de Grille électronique dans le PMOS ($V_{GS} = V_{DS}/4 - V_{DS}/5$). La structure est alors en avalanche au Drain (DAHE), où c'est le mécanisme dit "*channel shortening*" qui domine la dégradation du transistor PMOS. Pour le NMOS, c'est la polarisation du maximum du courant substrat qui réalise le pire cas de dégradation, avec un effet prononcé des électrons et des trous chauds du canal (CHE) ainsi générés, qui conduisent à la génération d'états d'interface.

Pour les oxydes d'épaisseur intermédiaire ($6.5nm$, technologie T2), c'est à la condition du maximum du courant substrat que la dégradation est la plus importante dans le NMOS. Pour réaliser ce type de vieillissement, il faut polariser à $V_{GS} = V_{DS}/2$. Le PMOS voit son pire cas de dégradation évoluer puisque c'est le cas HH ($V_{GS} = V_{DS}$) qui est le plus néfaste pour les transistors PMOS de cette technologie.

Sur les oxydes minces ($2.1nm$, technologie T3/T4), on a constaté une évolution notable du pire cas de dégradation puisque c'est à $V_{GS} = V_{DS}$ que le PMOS (injection de trous chauds du canal CHH) et les NMOS (injection d'électrons chauds du canal CHE) sont le plus fortement dégradés. Signalons que pour cette condition de polarisation à la tension nominale, le NMOS HS est au maximum du courant substrat. Enfin, notons que dans les dispositifs High Speed (T3) et Low Leakage (T4), les mécanismes sont identiques, seuls les niveaux de dégradation atteints varient, ainsi que les conditions de dégradations légèrement décalées vis-à-vis des tensions de stress V_{GS} et V_{DS} .

Ces éléments sont rassemblés sur la Fig. III.63 qui schématise les quantités de porteurs chauds injectés et les variations du courant de Drain vis à vis des conditions de stress.

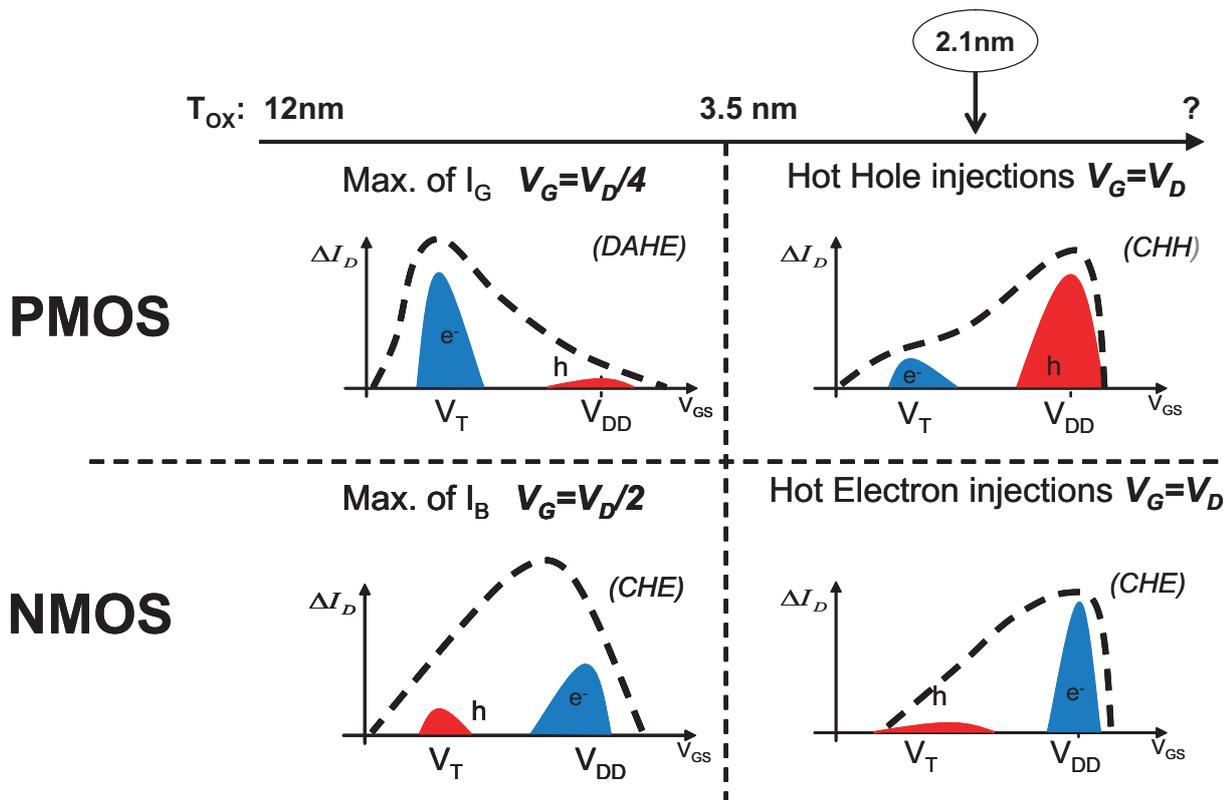


FIG. III.63 – Schéma récapitulatif des modes de dégradations porteurs chauds pour les différentes épaisseurs d'oxyde.

Conclusion

Dans ce chapitre nous avons comparé les dégradations suivant différentes configurations possibles d'injections de porteurs chauds, pour une gamme d'oxydes de Grille comprise entre $12nm$ et $2.1nm$. Nous avons réalisé de nombreux stress sur l'ensemble des trois technologies à notre disposition. L'analyse de ces expérimentations a pu être réalisée à l'aide de l'ensemble des techniques décrites au chapitre II. Cette étude nous a permis de conclure que la réduction de l'oxyde de Grille au fil de l'évolution des filières CMOS a pour conséquences principales:

- L'augmentation des courants de porteurs faiblement énergétiques à travers la Grille, qui s'avèrent fortement dégradants dans le cas des trous pour les oxydes ultra minces employés dans le PMOS ;
- L'augmentation de l'efficacité des trous chauds à générer des états d'interface dans les PMOS, avec pour corollaire la diminution de celle des électrons chauds à créer des charges fixes dans l'oxyde ;
- De même, dans les transistors NMOS, on constate une augmentation de l'efficacité des électrons chauds pour dégrader l'interface, accompagnée de la réduction de celle des trous ;
- Une très nette diminution de la charge piégée dans l'oxyde, que ce soit pour les NMOS ou les PMOS, avec dans ce dernier cas la disparition du phénomène de réduction de la longueur du canal lié à la charge négative piégée près du Drain.

Le cas des oxydes ultra-minces est rendu délicat, d'abord pour l'analyse des stress réalisés, mais aussi pour la compréhension des mécanismes mis en jeu. En effet, le fort courant tunnel direct, est présent dans toutes les mesures, que ce soit IV ou CP, mais aussi pendant les phases de vieillissement au cours desquelles son effet s'ajoute à celui des porteurs chauds. C'est dans le cas du PMOS que cette cohabitation est réellement gênante. En effet paradoxalement c'est dans le NMOS que ces courants sont les plus importants, mais c'est dans le PMOS qu'ils se révèlent le plus dégradant. Nous avons proposé un modèle simple pour séparer spatialement l'effet du courant tunnel et de l'injection de trous chauds, qui se manifestent simultanément dans les conditions de polarisation $V_{GS} = V_{DS}$ pour donner une dégradation répartie uniformément dans le canal.

Enfin, des expérimentations de stress alternés, nous ont permis de retrouver les résultats de la littérature pour les oxydes épais : les trous chauds génèrent des pièges neutres que l'on peut charger et décharger après stress. La réduction de l'oxyde a modifié ce comportement puisque les trous ne génèrent plus que des états d'interfaces que l'on ne peut vider par simple basculement du mode d'injection (absence de l'injection de trous chauds), ni par l'inversion du champ dans l'oxyde.

Références du Chapitre III

- [1] S. Tam, P. K. Ko, and C. Hu, “Lucky electron model of hot electron injections in MOSFET’s,” *IEEE Trans. Electron Devices*, vol. ED-31, p. 1116, 1994. [II.3.3](#), [III.1.1](#), [III.1.1](#), [III.1.1](#)
- [2] C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, and K. W. Terrill, “Hot-electron-induced MOSFET degradation - model, monitor, and improvement,” *IEEE Trans. Electron Devices*, vol. ED-32, pp. 375–385, 1985. [III.1.1](#), [III.1.1](#), [III.1.1](#), [III.1.1](#), [III.1.1](#), [III.1.1](#)
- [3] E. S. Yang, *Microelectronic Devices*, McGraw-Hill, Ed., New York, 1988. [III.1.1](#), [III.1.2](#)
- [4] Y. A. E. Mansy and D. M. Caugheyd, “Modeling weak avalanche multiplication in IGFETs and SOS transistor for CAD,” in *IEDM Tech Dig.*, pp. 31, 1975. [III.1.1](#)
- [5] T. H. Ning, C. M. Osburn, and H. N. Hu, “Emission probability of hot electron from silicon into silicon dioxide,” *J. App. Phys.*, vol. 48, pp. 286–293, 1977. [III.1.1](#), [b](#))
- [6] M. Shur, *Introduction to Electronic Devices*, J. Wiley and Sons, Eds., New York, 1996. [III.1.1](#), [III.1.2](#)
- [7] C. M. Svenson, *The defect structure Si/SiO₂ interface model based on trivalent silicon and its hydrogen compounds*. New York: Pergamon: S.T. Pantelides, Ed., ch. The physics of SiO₂ and its interface, pp. 329–333, 1978. [III.1.1](#)
- [8] F. C. Hsu and S. Tam, “Relationship between MOSFET degradation and hot-electron induced interface-state generation,” *IEEE Trans. Electron Devices*, vol. EDL-5, pp. 50–52, 1984. [III.1.1](#)
- [9] J. G. Simmons, “Richardson-schottky effect in solids,” *Phys. Rev. Letters*, vol. 15, pp. 967–968, dec 1965. [III.1.1](#)
- [10] G. V. D. Bosch, G. Groeseneken, and E. Maes, “Critical analysis of the substrate hot-hole injection technique,” *Solid-State Electronics*, vol. 37, no. 3, pp. 3393–3399, 1994. [III.1.2](#)
- [11] S. Wolf, *Silicon Processing for the VLSI Era - Volume III - The Submicron MOSFET*, L. Press, Ed., California, 1995. [III.1.2](#), [III.1.3](#), [2](#)
- [12] R. Woltjer, A. Hamada, and E. Takeda, “Time dependence of p-MOSFET hot-carrier degradation measured and interpreted consistently over ten orders of magnitude,” *IEEE Trans. Electron Devices*, vol. ED-40, pp. 392–401, 1993. [III.1.2](#)

- [13] S. Ogura, P. J. Tsang, W. W. Walker, D. L. Critchlow, and J. F. Shepard, "Design and characterization of the lightly doped drain-source (LDD) insulated gate field-effect transistor," *IEEE Trans. Electron Devices*, vol. ED-27, pp. 1359–1367, 1980. [III.1.2](#)
- [14] P. J. Tsang, S. Ogura, W. W. Walker, J. F. Shepard, and D. L. Critchlow, "Fabrication of high-performance LDDFET's with oxide sidewall-spacer technology," *IEEE Trans. Electron Devices*, vol. ED-29, pp. 590–596, 1982. [III.1.2](#)
- [15] R. Degraeve, "Oxide reliability," in *IEEE International Reliability Physics Symposium, Tutorial Notes*, 1997, pp. Topic 7, pp. 7.1–7.71., 1997. [III.1.3](#), [III.1.3](#)
- [16] J. McPherson and H. Mogul, "Disturbed bonding states in SiO₂ thin-films and their impact on time-dependent dielectric breakdown," in *IEEE International Reliability Physics Symposium Proc.*, pp. 47, 1998. [III.1.3](#)
- [17] D. J. Dumin, S. Mopuri, S. Vanchinatan, R. Scott, R. Subramanian, and T. Lewis, "High field emission related thin oxide wearout and breakdown," in *IEEE International Reliability Physics Symposium, Tutorial Notes*, pp. 143, 1994. [III.1.3](#)
- [18] R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, and H. E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," *IEEE Trans. Electron Devices*, vol. 45, no. 4, p. 904, 1998. [III.1.3](#)
- [19] E. Rosenbaum, "Oxide reliability," in *IEEE International Reliability Physics Symposium, Tutorial Notes*, pp. Topic 6a, pp. 6a.1–6a.27, 1996. [III.1.3](#)
- [20] D. Goguenheim, A. Bravaix, D. Vuillaume, F. Mondon, M. Jourdain, and A. Meinertzhagen, "Stress induced leakage currents in N-MOSFET's submitted to channel hot-carrier injections," *Journal of Non-Crystalline Solids*, vol. 245, p. 1999, 41. [III.1.3](#)
- [21] J. de Blauwe, J. Houdt, D. Wellekens, G. Groeseneken, and H. Maes, "SILC - related effects in flash E2PROM's - part i: A quantitative model for steady state SILC," *IEEE Trans. Electron Devices*, vol. 45, no. 8, p. 1745, 1998. [III.1.3](#)
- [22] D. J. DiMaria and E. Cartier, "Mechanism for stress-induced leakage currents in thin silicon dioxide films," *J. Appl. Phys*, vol. 78, pp. 3883–3894, 1995. [III.1.3](#)
- [23] A. Bravaix, T. D. Gilio, and D. Goguenheim, "Worst-case of the hot-carrier degradation between GO1-GO2 p-MOSFETs J203YB2 07A0 wafer," *Rapport Interne STMicroelectronics*, Juillet 2003. [a](#))
- [24] A. Bravaix, D. Goguenheim, N. Revil, and E. Vincent, "Hole injection enhanced hot-carrier degradation in PMOSFETs used for System On Chip applications with 6.5-2nm thick gate-oxide," *Microelectronic Reliability*, vol. 44, no. 1, pp. 65–77, 2004. [a](#)), [a](#))
- [25] A. Bravaix and D. Vuillaume, "analysis of the hot carrier degradation of deep-submicrometer large -angle tilt implanted drain (LATID) MOSFETs," *Solid State Electronics*, vol. 41, no. 9, pp. 1293–1302, 1997. [a](#))

-
- [26] A. Bravaix, "Etudes des dégradations du transistor PMOS soumis aux injections de porteurs chauds," Ph.D. dissertation, Univ. PARIS VII, Paris, 1991. **b)**
- [27] P. Heremans, R. B. ans G. Groeseneken, and H. E. Maes, "Consistent model for the hot-carrier degradation in n-channel and p-channel," *IEEE Trans. Electron Devices*, vol. 35, p. 2194, 1988. **b)**
- [28] N. G. Kwok and G. W. Taylor, "Effects of hot carrier trapping in- and p-channel MOSFET's," *IEEE Trans. Electron Devices*, vol. 30, p. 871, 1983. **b)**
- [29] A. Bravaix, D. Goguenheim, N. Revil, and E. Vincent, "Deep hole trapping effets in the degradation mechanisms of 6.5-2nm thick gate-oxide PMOSFETs," *Microelectronics Eng.*, vol. 72, pp. 106–111, 2004. **b)**, **III.28**
- [30] J. J. Tzou, C. C. Yao, R. Cheung, and H. W. K. Chan, "Hot-carrier-induced degradation in p-channel LDD MOSFET's," *IEEE Electron Device Lett.*, vol. 7, p. 5, 1989. **III.2.4**
- [31] T. D. Gilio and A. Bravaix, "Technique de détermination de la durée de vie des transistors MOSFET de longueur de grille ultra courte," in *Journées Nationales du réseau Doctoral de Micro-électronique, Marseille, 2004*. **b)**
- [32] A. Bravaix, D. Goguenheim, N. Revil, and E. Vincent, "Comparison of low leakage and high speed deep submicron PMOSFET's submitted to hole injection," in *IEEE International Reliability Workshop 02*, 2002, pp. 14–20, 2002. **b)**, **d)**, **III.57**, **d)**
- [33] X. Z. K. Y. Lim, "An analytical effective channel-length modulation model for overshoot in submicron MOSFETs based on energy-balance formulation," *Microelectronic reliability*, no. 42, pp. 1857–1864, 2002. **c)**
- [34] M. Koyanagi, A. Lewis, J. Zhu, R. Martin, T. Huang, and J. Chen, "Investigation and reduction of hot electron induced punchthrough (HEIP) effect in submicron PMOSFETs," in *IEDM proc. p722, 1986*. **d)**
- [35] A. Bravaix, L. Gauthé, and D. Goguenheim, "P-mosfet hot-carrier reliability in low leakage and high speed HCMOS9 technology," *rapport confidentiel de ISEM-Lab pour ST-Microelectronics Crolles*, 2002. **d)**
- [36] B. Doyle, M. Bourcerie, J.-C. Marchetaux, and A. Boudou, "Interface state creation and charge trapping in the medium-to-high gate voltage range ($v_d/2 \leq v_g \leq v_d$) during hot-carrier stressing of n-MOS transistors," *IEEE Trans. Electron Devices*, vol. EDL-37, pp. 744–754, 1990. **1**, **b)**, **III.3**, **III.3.1**
- [37] D. Vuillaume, J. Marchetaux, P.-E. Lippens, A. Bravaix, and A. Boudou, "A coupled study by floating-gate and charge-pumping techniques of hot-carrier-induced defects in submicrometer LDD n-MOSFET's," *IEEE Trans. Electron Devices*, vol. 40, no. 4, p. 773, 1993. **III.3**, **III.3.1**
- [38] D. Bauza and G. Ghibaud, "Analytical study of the contribution of fast and slow oxide traps to the charge pumping current in MOS structures," *Solid State Electronics*, vol. 39, no. 4, pp. 563–570, 1995. **III.3**
-

- [39] M. Declercq and P. Jespers, "Analysis of interface properties in MOS transistors by means of charge pumping measurements," vol. 9, 1974, pp. 244–253, 1974. [III.3](#)
- [40] S. Brugler and P. Jespers, "Charge pumping in MOS devices," *IEEE Trans. Electron Devices*, vol. 16, no. 3, pp. 297–302, 1969. [III.3](#)
- [41] P. Heremans, R. Bellens, G. Groeseneken, H. Maes, B. S. Doyle, M. Bourcerie, C. Bergonzoni, R. Benecchi, A. Bravaix, A. J. Marchetaux, and A. Boudou, "The generation and characterization of electron and hole traps created by hole injection during low gate voltage hot-carrier stressing of n-MOS transistors," *IEEE Trans. Electron Devices*, vol. 39, p. 758, 1992. [III.3.1](#)
- [42] W. Weber, M. Brox, R. Thewes, and N. Saks, "Hot-hole-induced negative oxide charges in n-MOSFETs," *IEEE Trans. Electron Devices*, vol. 42, pp. 1473–1480, 1995. [III.3.1](#)
- [43] D. S. Doyle and G. J. Dunn, "Recovery of hot-carrier damage in reoxidized nitrated oxide MOSFETs," *IEEE Trans. Electron Devices*, vol. 3, no. 1, pp. 39–40, 1992. [III.3.2](#)

CHAPITRE IV

Techniques d'estimation de la durée de vie des dispositifs MOSFET

Introduction

D'un point de vue industriel, les études de fiabilité ont pour but de déterminer la viabilité commerciale d'un produit, de s'assurer que le produit final répondra aux attentes des consommateurs, en terme de robustesse et de longévité. Les fondeurs ont donc besoin de connaître la durée de vie des dispositifs, évaluée suivant des critères d'usures bien définis. Bien évidemment, les impératifs temporels de développements rendent impossible de polariser la structure dans ses conditions normales de fonctionnement et d'attendre ses premiers signes de défaillance. C'est pourquoi on utilise les méthodes d'accélération présentées au chapitre précédent. Dans ce dernier chapitre nous allons présenter les modèles prédictifs empiriques qui permettent de déterminer les temps nécessaires pour atteindre un niveau de dégradation donné d'un paramètre pré-défini, en distinguant le cas du NMOS et celui du PMOS, en fonction des différentes filières technologiques. Par la suite, sera présentée l'influence du moniteur choisi pour suivre le vieillissement des structures MOSFET. Enfin nous nous pencherons sur le cas particulier du PMOS de la technologie T3.

IV.1 Expression de la durée de vie

Le but de toute étude de fiabilité est de connaître la durée de vie du produit. Mais comment peut-on définir la durée de vie d'un transistor MOSFET. Sur quel critère se baser pour affirmer qu'un dispositif MOS n'est plus viable ? Pour répondre à ces questions, il faut être capable de :

1. déterminer le critère capable de décrire le vieillissement de la structure de la manière la plus pertinente possible vis à vis du mécanisme responsable des dégradations ;
2. de mettre en place une méthodologie fiable pour extrapoler la durée de vie aux tensions nominales.

Ceci est une tâche complexe car les mécanismes de dégradation évoluent au fil des technologies successives. Le Tab. IV.1 montre de manière synthétique les résultats du Chap. III : la localisation de la dégradation et les mécanismes qui en sont à l'origine évoluent de manière significative avec l'épaisseur de l'oxyde.

	Oxydes épais (12nm)		Oxydes "moyen" (6.5nm)		Oxydes fins (2.1nm)	
	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
zone	1 - 2 - 3	1 - 2	1 - 2	1 - 2	1 - 2	1
défauts	N_{it}	N_{it}, N_{ot}^-	$N_{it}, N_{ot}^-, N_{ot}^+$		$N_{it} \nearrow, N_{ot}^- \searrow, N_{ot}^+$	
pire cas	IB	IGe	IB	HH	HE	HH

TAB. IV.1 – Localisation des dégradations dans le canal (référéncée par rapport à la Fig. II.1) et mécanismes de dégradation (pire cas) pour les dispositifs NMOS et PMOS des technologies T1, T2 et T3.

Pour chacun de ces cas DC, la polarisation responsable du vieillissement est différente, ainsi que le mécanisme de dégradation qui en résulte. En régime AC, les cycles de polarisation sont assimilables à une succession dans le temps de cas DC [1, 2]. On comprend donc la difficulté à choisir la méthode de prédiction de la durée de vie mais aussi son mode de représentation.

La définition la plus courante de durée de vie pour un composant correspond au niveau de dégradation (qui est donc notre critère) atteint par un paramètre donné. Le critère courant actuel de l'industrie est 10% de dégradation à la tension nominale. Concrètement, on considèrera que la durée de vie est atteinte lorsque le paramètre choisi a varié de 10% au fonctionnement nominal. Ce critère correspond donc plutôt à une réduction des performances qu'à la notion de défaillance complète du dispositif. Le choix du paramètre d'étude est crucial et doit être en adéquation avec les conditions d'utilisation les plus sévères et surtout le plus révélateur possible du mécanisme de dégradation étudié. A partir des stress accélérés ($V_{Dstress} \approx 1.3 - 1.4V_{DD}$)

que nous avons effectué au Chapitre III, il faut déduire le niveau de dégradation à la tension nominale.

IV.1.1 Première modélisation du vieillissement des dispositifs MOSFET

A partir de constatations empiriques, il a été observé d'une manière générale, que la variation d'un paramètre électrique P des transistors MOSFET soumis à l'injection de porteurs chauds suit une loi de puissance du temps [3] :

$$\Delta P \propto At^n \quad (\text{IV.1})$$

Le paramètre P peut représenter les variations de la tension de seuil (ΔV_{Th}), de la transconductance ($\Delta Gm/Gm_0$), du courant de Drain ($\Delta I_D/I_{D0}$) ou d'autres paramètres transistor représentatifs de leur performance. n est la pente dans un graphe Log-Log et dépend des valeurs de V_{GS} et V_{DS} pendant le stress. n est donc une caractéristique du mécanisme d'injection.

L'autre aspect important de (IV.1) est le paramètre A qui représente finalement l'amplitude de la dégradation à $t = 0$. A est fortement dépendant de V_{DS} et assez peu de V_{GS} . Il a été démontré expérimentalement que la dépendance de A avec la tension de Drain peut se mettre sous la forme [3] :

$$A \propto \exp\left(\frac{-\alpha}{V_{DS}}\right) \quad (\text{IV.2})$$

où α est simplement la pente de (IV.2) dans un graphe avec une échelle Lin-Log. En effet cette équation se transforme en :

$$\ln(A) \propto -\alpha(1/V_{DS})$$

Nous avons représenté l'évolution de A en fonction de l'inverse de la tension de Drain pendant des stress de type IB sur les trois épaisseurs d'oxyde à notre disposition.

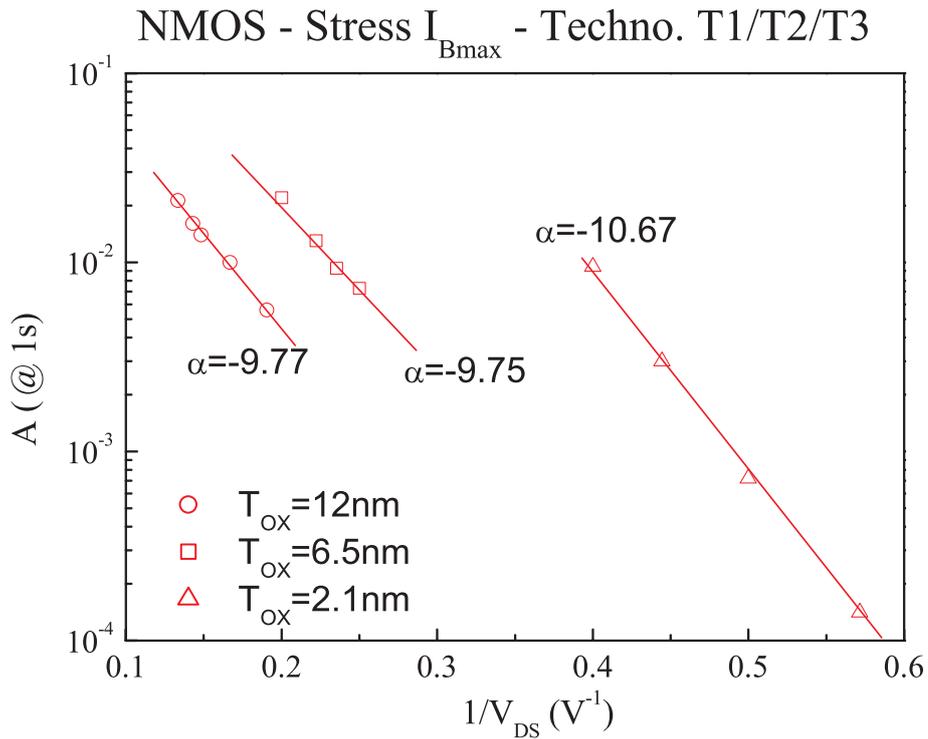


FIG. IV.1 – Relation entre le paramètre A de (IV.2) et la tension de Drain pendant les stress IB effectués sur des NMOS des technologies T1, T2 et T3 ; A est extrait des cinétiques de dégradation du courant de Drain mesuré en régime linéaire pendant le stress.

Par ailleurs, il a été démontré analytiquement que les pics de courant du courant substrat peuvent être décrits par [4] :

$$I_{BS}^{max} \propto \exp\left(-\frac{\beta}{V_{DS}}\right) \propto A^{\frac{\beta}{\alpha}} \quad (IV.3)$$

A partir de (IV.1), (IV.2) et (IV.3) on peut déduire une première relation simplifiée¹ de la durée de vie τ :

$$\tau \propto \exp\left(\frac{b}{V_{DS}}\right) \quad (IV.4)$$

On choisit pour cela un critère, ce qui revient à fixer la valeur limite pour ΔP . La valeur standard est 10%. Dans cette dernière relation, b est simplement le rapport n/α .

De la même façon, Takeda [3] a établi une relation de la durée de vie en fonction du courant substrat :

1. Cette expression découle en fait de la relation du courant substrat au champ latéral : b/V_{DS} a remplacé $b/(V_{DS} - V_{DSat})$ (I.113) dans I.4.5

$$\tau \propto (I_{BS}^{max})^{-l} \quad (IV.5)$$

avec $l = -1/n \cdot (\alpha/\beta)$ qui a été évalué autour de 3.2 – 3.4 pour une gamme d'épaisseur d'oxyde comprise entre 20nm et 6nm [3].

La Fig. IV.2 montre la durée de vie des trois technologies T1, T2 et T3 pour les stress IB, la méthodologie sera décrite dans la prochaine section. On constate que (IV.5) décrit parfaitement les trois courbes, qui sont alors des droites superposées dans l'échelle Log-Log.

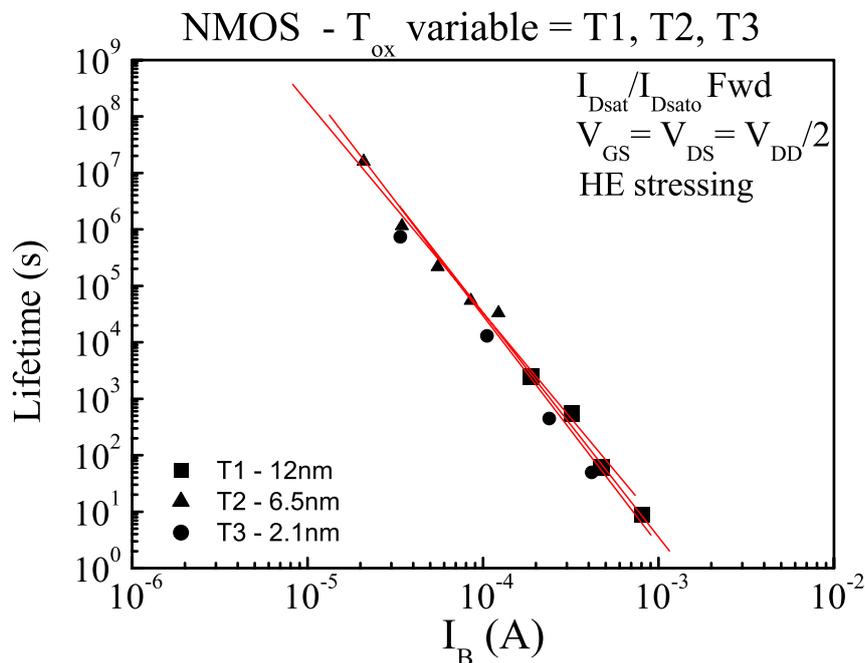


FIG. IV.2 – Comparaison des durées de vie exprimées en fonction du courant Substrat des transistors PMOS de la technologie T1, T2 et T3 soumis au stress IB.

Les relations (IV.4) et (IV.5) sont les premières expressions qui permettent de déterminer clairement la durée de vie des structures MOSFETs en fonction du champ électrique (proportionnel à V_{DS}) et du nombre de porteurs chauds (proportionnel au maximum du courant substrat) générés lors de l'ionisation par impact.

IV.1.2 Détermination pratique de la durée de vie (Transistor NMOS)

Afin de confirmer ou d'infirmer les modèles de prédiction du vieillissement des structures MOSFETs que nous venons d'établir, nous avons besoin de pouvoir construire des graphes de durée de vie à partir des données expérimentales acquises au Chap. III. Pour réaliser cette tâche trois points de départ sont essentiels :

- Choisir un paramètre transistor (appelé P dans le modèle) le plus représentatif de la dégradation occasionnée (il serait par exemple peu judicieux de choisir la pente sous seuil dans le cas d'une dégradation uniforme) ;
- Choisir un niveau de dégradation (appelé critère de durée de vie) pour lequel on considérera que la durée de vie est atteinte par le dispositif (on prend généralement $\Delta P = 10\%$ qui est un critère industriel répandu) ;
- et enfin il faut choisir un moniteur en fonction duquel sera exprimé la durée de vie ($1/V_{DS}, I_{BS} \dots$).

Le point de départ est d'effectuer une série de stress pour plusieurs polarisations (autour de $1.3 - 1.4 \times V_{DD}$). L'étude des cinétiques de variation pendant le stress du paramètre choisi permet de relever le temps nécessaires pour atteindre le niveau de dégradation correspondant au critère fixé. On aura pris soin de noter les valeurs des courants I_{DS} , I_{BS} et I_{GS} au temps $t = 1s$ pour chaque polarisation (le programme de contrôle de notre banc expérimental stocke ces données dans le même fichier que les variations de tous les paramètres pendant le stress).

A partir de ces relevés on est en mesure de construire le graphe de durée de vie en plaçant les points $t_{crit}(V_{DS}) - 1/V_{DS}$ dans un tracé Lin-Log, ou bien les points $t_{crit}(V_{DS}) - I_{BS}(V_{DS})$ dans un tracé Log-Log (Fig. IV.4). La durée de vie est l'intersection de la droite définie par $X = 1/V_{DD} = 0.83$ ou bien $X = I_{BS}(@V_{DD})$ et la droite obtenue.

Sur la Fig. IV.4, on a tracé la durée de vie des transistors NMOS de la technologie T1 soumis à des stress au maximum du courant Substrat. Les points sont alignés dans l'échelle Log-Log et la durée de vie extrapolée par prolongation de la droite jusqu'à la valeur de I_{BS} mesuré à $V_{DS} = V_{DD}$ et $V_{GS} = 2.15V$ (condition I_{BSmax}) donne $\tau = 2 \times 10^9 s$, soit plus de dix ans, pour la dégradation du courant saturé (FWD).

La Fig. IV.5 montre la durée de vie exprimée en fonction du courant Substrat pour différents modes de fonctionnement pour un stress au maximum du courant Substrat. La durée de vie augmente avec les polarisations de mesure, en accord avec l'écrantage des défauts (localisés près du Drain) par l'extension de la ZCE : $\tau_{Lin} = 1.3 \times 10^{11} s$, $\tau_{Sat} = 1.6 \times 10^{11} - 2.36 \times 10^{15} s$

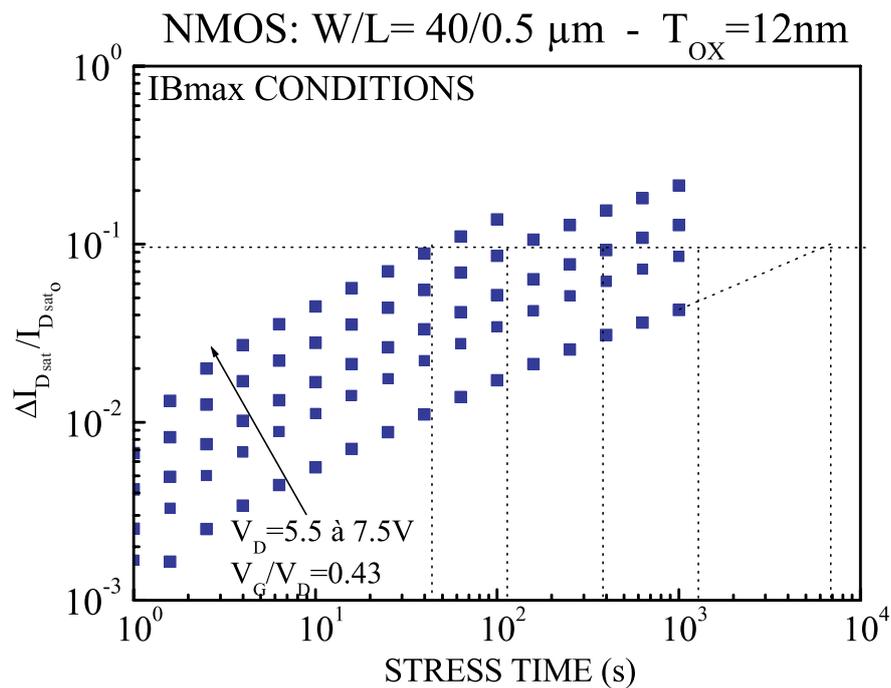


FIG. IV.3 – Illustration du relevé des temps nécessaires pour atteindre le critère de 10%, le paramètre est dans cet exemple le courant saturé à $V_{\text{GS}} = V_{\text{DS}} = V_{\text{DD}}/2$ FWD.

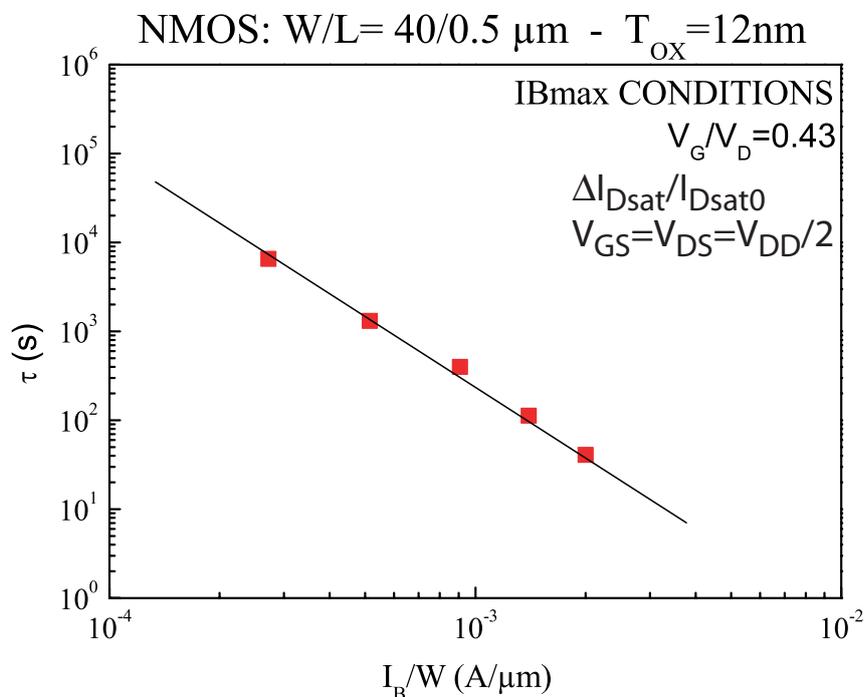


FIG. IV.4 – Durée de vie des transistors NMOS de la technologie T1 soumis à des injections de porteurs chauds au maximum du courant Substrat, extrapolée pour le mode de fonctionnement saturé ($V_{\text{DD}}/2, V_{\text{DD}}/2$) (FWD) en fonction du courant Substrat.

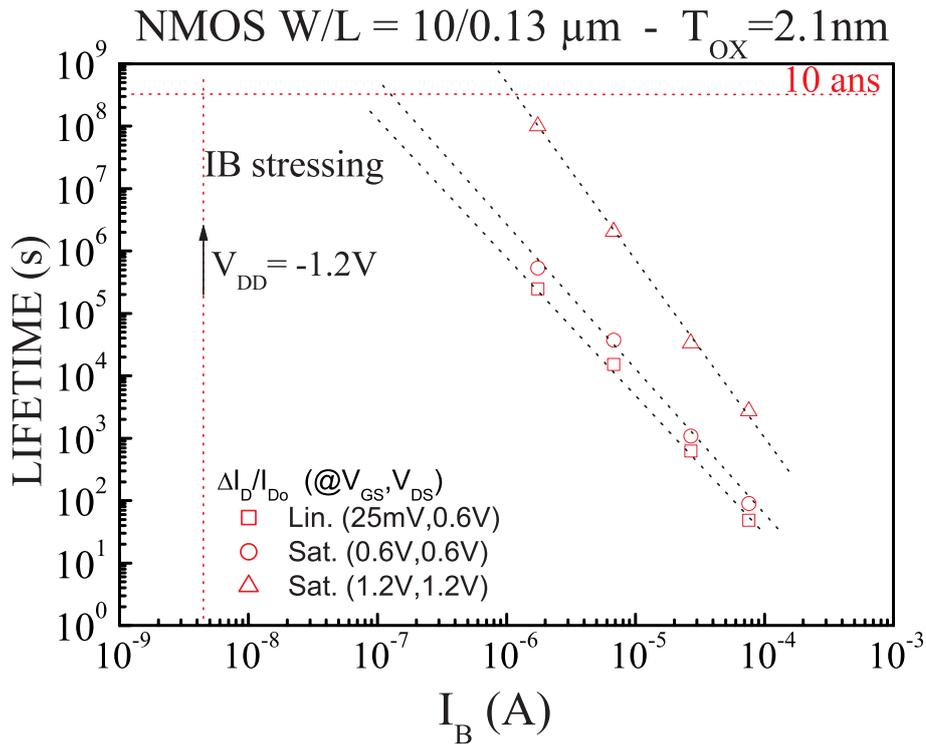


FIG. IV.5 – Durée de vie exprimée en fonction du courant Substrat pour différentes conditions de fonctionnement d'un transistor NMOS de la technologie T3.

IV.1.3 Expression de la durée à l'aide du "Lucky Electron Model"

a) Durée de vie en fonction de I_{BS} (transistor NMOS)

Le modèle analytique appelé "modèle de l'électron chanceux" ou "Lucky Electron Model" (LEM) [5] exprime la probabilité pour les électrons du canal de générer des défauts à l'interface Si-SiO₂ ou dans le volume de l'oxyde de Grille. Avant d'établir le lien entre ce modèle et la durée de vie des dispositifs MOSFET, rappelons les principes du LEM. Ces auteurs ont utilisé la relation qui lie le courant Substrat à la probabilité pour un électron de générer l'ionisation par impact [6] :

$$I_{BS} = C_1 I_{DS} \exp\left(-\frac{\phi_{i,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.6})$$

Nous avons déjà défini les quantités λ_e , $\phi_{i,e}$ et C_1 comme le libre parcours moyen, l'énergie que doit acquérir un électron pour provoquer ionisation par impact et un paramètre faiblement dépendante de ξ_m . Hu [7] a obtenu la relation liant ΔN_{it} au temps à partir du taux de génération G des défauts :

$$\frac{d\Delta N_{it}}{dt} = A.G(\Delta N_{it})\frac{I_{DS}}{W} \exp\left(\frac{-\phi_{it,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.7})$$

où A est une constante. Si on assume que I_{DS} et ξ_m ne sont pas affectés par la dégradation, (IV.7) s'intègre en écrivant :

$$\int \frac{d\Delta N_{it}}{G(\Delta N_{it})} = A\frac{I_{DS}}{W}t \exp\left(\frac{-\phi_{it,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.8})$$

Cette dernière relation permet d'obtenir l'expression de ΔN_{it} comme une fonction f du temps :

$$\Delta N_{it} = f\left[t\frac{I_{DS}}{W} \exp\left(\frac{-\phi_{it,e}}{q\lambda_e\xi_m}\right)\right] \quad (\text{IV.9})$$

Cette expression est de la forme $\Delta = C_0t^p$, ainsi la relation entre le temps et la formation des états d'interface s'écrit indépendamment de leur mécanisme de formation [7] :

$$\Delta N_{it} = C_2 \left[t\frac{I_{DS}}{W} \exp\left(\frac{-\phi_{it,e}}{q\lambda_e\xi_m}\right) \right]^p \quad (\text{IV.10})$$

On définit alors τ comme le temps nécessaire pour atteindre une quantité de défaut ΔN_{it} prédéfinie et C_2 inclue C_0 . (IV.10) se transforme donc comme :

$$\tau = \underbrace{\left(\frac{\Delta N_{it}}{C_2}\right)^{1/p}}_{C_3} \frac{W}{I_{DS}} \exp\left(\frac{\phi_{it,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.11})$$

En combinant (IV.11) (IV.6) on obtient l'expression :

$$\tau = C_3 \frac{W}{I_{DS}} \left(\frac{I_{BS}}{I_{DS}}\right)^{\left(\frac{-\phi_{it,e}}{\phi_{i,e}}\right)} \quad (\text{IV.12})$$

ce qui nous permet d'arriver à la relation usuelle [7, 8] :

$$\frac{\tau I_{DS}}{W} = C'_3 \left(\frac{I_{BS}}{I_{DS}} \right)^{-k} \quad (\text{IV.13})$$

La constante C'_3 contient la constante C_2 et la quantité ΔN_{it} à la puissance $1/p$. (IV.13) est exprimée en fonction des paramètres physiques des réactions de dégradation : k est le rapport des termes dans les exponentielles de (IV.3) et (IV.5), c'est à dire le rapport des énergies nécessaires pour générer un état d'interface d'une part, et nécessaire à un électron pour générer l'ionisation par impact d'autre part. $\phi_{it,e}$, représente la somme de la hauteur de barrière pour les électrons et de l'énergie de la liaison $\equiv \text{Si-H}$ (modèle standard pour la génération des états d'interface [7]), égale à $0.3eV$.

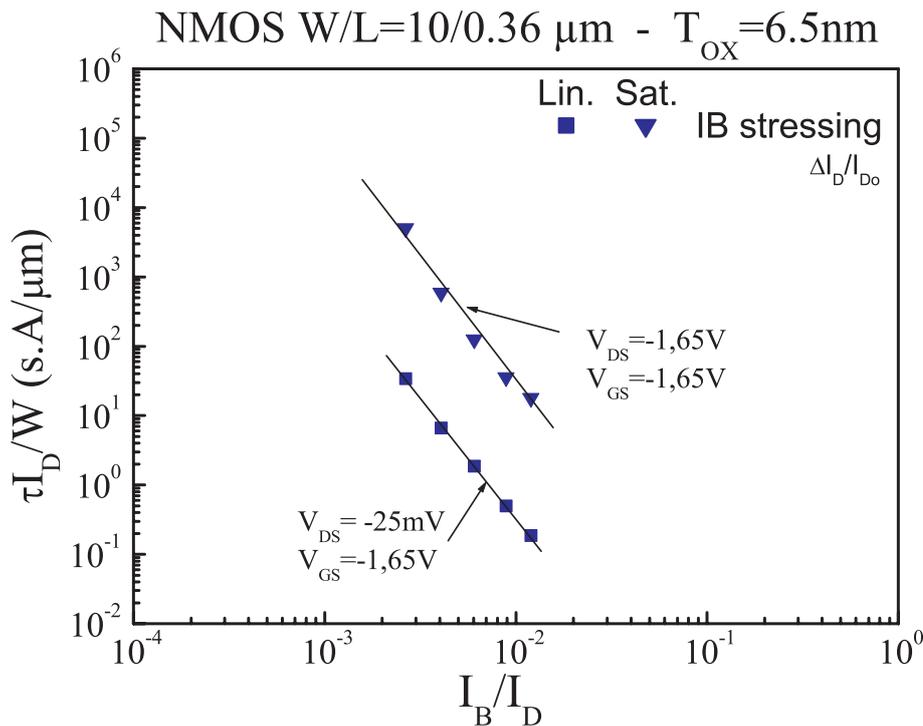


FIG. IV.6 – Durée de vie des transistors NMOS de la technologie T2 exprimée à partir de la modélisation déduite du LEM, pour les stress IB.

Sur la Fig. IV.6 nous avons tracé les durées de vie des dispositifs à canal n de la technologie T2 soumis au stress IB (cas standard). Nous avons utilisé le paramètre $\Delta I_D / I_{D0}$ avec un critère standard de 10% pour les modes linéaire et saturé ($V_{DD}/2$, $V_{DD}/2$). Pour ce mécanisme d'injection on trouve la même pente $k = 2.9$ à 3.1 . Ceci nous permet d'affirmer que dans les deux cas, ce sont les électrons qui sont injectés et qui sont responsables de la génération d'états

d'interface. En effet, la pente $k = \phi_{it,e}/\phi_{i,e} \approx 2.84$. D'autre part, le coefficient d'ionisation $\phi_{i,e}$ étant pris égal $1.3eV$, on obtient dans cette technologie une énergie critique pour les électrons $\phi_{it} = 3.7$ à $3.8eV$, ce qui correspond aux valeurs de la littérature [5, 7] pour les électrons.

b) Durée de vie en fonction de I_{GS} (transistors NMOS)

Comme nous l'avons exposé dans le chap. III, les transistors à oxydes de Grille épais sont fortement dégradés dans les conditions de polarisation correspondant au maximum du courant de Grille électronique. Ce mécanisme d'injection engendre une charge piégée dans l'oxyde ΔN_{ot}^- importante [9], qui n'est pas prise en compte par la modélisation en I_{BS} . Cette charge piégée est la conséquence de l'injection d'électrons chauds dans l'oxyde, dont la quantité est évaluée par le courant de Grille $I_{G,e}$.

Il paraît donc commode et réaliste de tenter de représenter la durée de vie à l'aide du courant de Grille I_{GS} . Toujours d'après le LEM, on peut exprimer le courant de Grille à l'aide de la probabilité qu'un électron surmonte la hauteur de la barrière de potentiel de l'interface [7]:

$$I_{GS} = C_3 I_{DS} \exp\left(\frac{-\phi_{b,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.14})$$

Nous en avons déjà parlé précédemment (III.1.1), cette relation, associée à l'expression du courant substrat (IV.6), permet d'écrire [7]:

$$\frac{I_{GS}}{I_{DS}} = C_4 \left(\frac{I_{BS}}{I_{DS}}\right)^l \quad (\text{IV.15})$$

L'exposant l est le rapport entre la hauteur de barrière (pour les électrons) et l'énergie d'ionisation : $l = \phi_{b,e}(F_{ox})/\phi_{i,e}$. C_4 est une constante qui a été évaluée proche de 2×10^{-3} pour $V_{GS} > V_{DS}$ [5]. Cependant, cette constante est à déterminer expérimentalement car elle dépend de F_{ox} et donc des conditions $V_{GS} - V_{DS}$ ou V_{GS}/V_{DS} . Il s'agit donc d'un paramètre dépendant de la technologie testée.

On peut ainsi trouver l'expression de la durée de vie en fonction du courant de Grille en combinant (IV.12) et (IV.14). On obtient l'équation [10]:

$$\frac{\tau I_{DS}}{W} = C_5 \left(\frac{I_{GS}}{I_{DS}}\right)^{-u} \quad (\text{IV.16})$$

Pour arriver à cette relation, nous avons considéré que le mécanisme dominant est le piégeage de charges négatives dans l'oxyde (stress HE). Il faut donc remplacer ΔN_{it} par ΔN_{ot} dans la constante de (IV.11). Ainsi l'exposant u est le rapport entre l'énergie nécessaire pour les électrons injectés d'être piégés dans l'oxyde ($\phi_{ot,e}$) et la hauteur de barrière ($\phi_{b,e}(F_{ox})$): $u = \phi_{ot,e}/\phi_{b,e}(F_{ox})$.

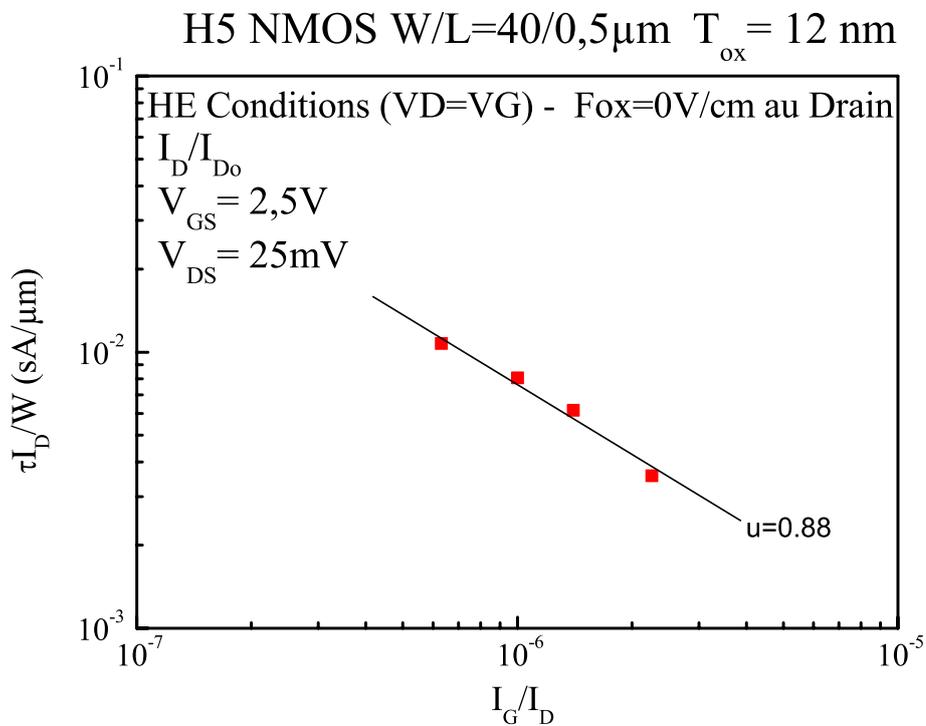


FIG. IV.7 – Durée de vie suivant les courants de Drain et de Grille d'un transistor NMOS de la technologie T1 soumis à un stress HE, le critère est 10% du courant de Drain en régime linéaire.

Nous avons représenté sur la Fig. IV.7 le tracé de la durée de vie d'un transistor NMOS pour un stress HE avec ce mode de représentation. L'exposant de la droite obtenue est $u = 0.88$. En prenant $\phi_{b,e} = 3.2eV$ (puisque $F_{ox} \approx 0V/cm$ au Drain), on trouve pour l'énergie de piégeage $\phi_{ot,e} = 2.88eV$ en accord avec le LEM [10].

IV.1.4 Equivalence pour les transistors PMOS

a) Durée de vie en fonction de I_{BS}

L'ensemble des expressions que nous venons d'établir sont appliquées au transistor NMOS, pour lequel a été développé le LEM. Pour transposer ces résultats au transistor PMOS, il faut reconsidérer l'analyse puisque les mécanismes de vieillissement sont différents.

Les résultats de stress que nous avons présentés au chapitre précédent (III.2.2.b) ont mis en avant les spécificités du transistor PMOS vis à vis des injections de porteurs chauds. Nous avons montré que le PMOS se voit fortement dégradé sous injection d'électrons chauds (au maximum du courant de Grille électronique). Il présente une forte charge négative piégée dans l'oxyde de Grille [11], qui se traduit par une variation des paramètres en logarithme du temps [12]. La nécessité de mettre en place de nouvelles méthodologies pour la détermination de la durée de vie des PMOS est renforcée par la multiplicité des mécanismes qui peuvent intervenir dans le vieillissement de la structure : N_{it} , N_{ot}^- et N_{ot}^+ [13].

Cas HE Dans le cas de l'injection HE (maximum du courant de Grille électronique, à bas $|V_{GS}|$), le mécanisme dominant pour les oxydes épais est le piégeage de charges négatives dans la région du Drain, conduisant à une réduction de la longueur active du canal [14]. Dans ces conditions la structure est dans le mode DAHE. L'analogie entre ce mécanisme et la dégradation au maximum du courant substrat dans le PMOS nous permet d'écrire la probabilité P pour un électron chaud d'être injecté et piégé dans l'oxyde [1]:

$$P \propto I_{BS} \exp\left(\frac{-\phi_{ot,e}}{q\lambda_e\xi_m}\right)$$

$\phi_{ot,e}$ représente l'énergie minimale pour créer ou remplir un défaut d'oxyde. (IV.11) se transforme ainsi en :

$$\tau = C_3 \frac{W}{I_{BS}} \exp\left(\frac{\phi_{ot,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.17})$$

Dans le PMOS le courant Substrat s'obtient par le produit du nombre de trous disponibles dans le canal ($\propto I_{DS}$) et la probabilité d'ionisation par impact ($\propto \exp[-\phi_{i,h}/(q\lambda_h\xi_m)]$):

$$I_{BS} = C_1' I_{DS} \exp\left(-\frac{\phi_{i,h}}{q\lambda_h\xi_m}\right) \quad (\text{IV.18})$$

Cette expression du courant Substrat est identique dans sa forme à celle obtenue dans le cas des transistors NMOS. Ici $\lambda_h \approx 5nm$ [15], ce qui correspond au libre parcours moyen des phonons optiques. En combinant (IV.18) et (IV.17) on arrive alors à l'expression de durée de vie pour l'injection HE (oxyde épais) [11] :

$$\frac{\tau I_{BS}}{W} = C_3 \left(\frac{I_{BS}}{I_{DS}}\right)^{-j} \quad (\text{IV.19})$$

L'exposant j est donné par $\frac{\phi_{ot,e}\lambda_h}{\phi_{i,h}\lambda_e}$.

Cas HH Le Chap. III a montré que la réduction de l'épaisseur de l'oxyde de Grille modifie profondément le pire cas de dégradation des transistors PMOS. Ainsi la polarisation de stress la plus néfaste devient $V_{GS} \approx V_{DS}$ (III.2.3.b) pour des oxydes $T_{ox} \leq 6.5nm$. Le mécanisme dominant devient le cas HH (injection de trous chauds), le piégeage de charges négatives est fortement réduit et son influence devient négligeable (III.2.4.b). On suppose donc que le mécanisme dominant est la génération d'états d'interface ΔN_{it} . On considère ainsi la probabilité P pour un trou de générer un état d'interface :

$$P \propto I_{DS} \exp\left(\frac{-\phi_{it,h}}{q\lambda_h\xi_m}\right)$$

qui permet de reprendre (IV.12) :

$$\tau = \frac{C_6}{I_{DS}} \exp\left(\frac{\phi_{it,h}}{q\lambda_h\xi_m}\right) \quad (\text{IV.20})$$

La combinaison de (IV.18) et (IV.20) donne finalement la relation de durée de vie pour le cas HH :

$$\frac{\tau I_{DS}}{W} = C_6 \left(\frac{I_{BS}}{I_{DS}} \right)^{-k'} \quad (\text{IV.21})$$

où $k' = \frac{\phi_{it,h}}{\phi_{i,h}}$. Cette expression découle de l'hypothèse que ΔN_{it} est le mécanisme dominant. Cependant, dans certains cas comme à partir du stress au maximum de I_{BS} (III.2.3.b)), l'injection des trous chauds entraîne le piégeage d'une charge piégée négative au temps courts. Comme les trous sont à l'origine de ce mécanisme, il suffit de remplacer $\phi_{it,h}$ par $\phi_{ot,h}$, qui est l'énergie pour créer ou remplir un piège dans l'oxyde, dans les équations (IV.18), (IV.20) et dans l'exposant de (IV.21). Cette approche, où l'injection de trous chauds génèrent des états d'interface dans les PMOS au temps longs, est très similaire au LEM pour les NMOS. On parle ainsi de LHM pour "*Lucky Hole model*" [16].

b) Durée de vie en fonction de I_{GS}

Comme nous venons de le voir, au fil des générations technologiques, il a fallu adapter les modèles d'extrapolation de la durée de vie pour suivre l'évolution des mécanismes de dégradation dominants. Pour les oxydes épais, ce sont les électrons qui dégradent la structure avec pour conséquence un fort piégeage de charges négatives. Ce mécanisme est caractérisé à bas V_{GS} par un courant électronique $I_{G,e}$. Pour les oxydes plus fins ($< 6.5nm$), la plus grande dégradation est observée pour des fortes valeurs de V_{GS} ($\approx V_{DS}$), où les trous chauds génèrent une grande quantité d'états d'interface. Ce mécanisme est quant à lui caractérisé par un courant de trous à la Grille $I_{G,h}$. Notons ici que pour les oxydes ultra minces ($< 3nm$) le courant mesuré n'est plus uniquement thermoionique, mais résulte du transport combiné avec le mode tunnel direct de porteurs faiblement énergétiques au travers de l'oxyde.

Etablissons maintenant les durées de vie suivant le courant de Grille $I_{G,e}$. Raisonnons en terme de probabilité : pour injecter un électron dans la Grille d'un PMOS, il faut réaliser les deux évènements suivants :

- un trou du canal a une énergie cinétique suffisante pour la génération de la paire électron-trou (probabilité proportionnelle à $\exp(-\phi_{i,h}/q\lambda_h\xi_m)$)
- un électron issu de ces paires e/h a une énergie supérieure à la hauteur de barrière $\phi_{b,e}(F_{ox})$ (probabilité proportionnelle à $\exp(-\phi_{b,e}(F_{ox})/q\lambda_e\xi_m)$)

La probabilité P d'injecter un électron chaud est donc donnée par le produit [1]:

$$P \propto \exp\left(-\frac{\phi_{i,h}}{q\lambda_h\xi_m}\right) \times \exp\left(-\frac{\phi_{b,e}(F_{ox})}{q\lambda_e\xi_m}\right)$$

La dépendance en F_{ox} de la hauteur de barrière ne peut pas être négligée puisqu'elle conditionne directement la probabilité d'injection. En utilisant cette probabilité appliquée à la quantité d'électrons près de la zone d'ionisation :

$$I_{G,e} = C'_1 I_{DS} \exp\left[-\frac{1}{q\xi_m} \left(\frac{\phi_{i,h}}{\lambda_h} + \frac{\phi_{b,e}(F_{ox})}{\lambda_e}\right)\right] \quad (\text{IV.22})$$

Cette expression du courant de Grille, qui diffère de celle obtenue dans le cas des transistors NMOS, donne une nouvelle relation du courant de Grille au courant Substrat pour le PMOS [1, 17]:

$$\begin{cases} \frac{I_{GS}}{I_{DS}} = C'_4 \left(\frac{I_{BS}}{I_{DS}}\right)^{l'} \\ l' = 1 + \frac{\phi_{b,e}(F_{ox})}{\phi_{i,h}} \times \frac{\lambda_h}{\lambda_e} \end{cases} \quad (\text{IV.23})$$

Sur la Fig IV.8 on a tracé cette relation pour $V_{GD} = V_{GS} - V_{DS}$ constant, ce qui correspond à une valeur de champ d'oxyde constante dans la zone d'injection. Le ratio λ_h/λ_e vaut 0.724 [11]. L'abaissement de la hauteur de barrière est calculé à l'aide de (III.2), et on obtient $\phi_b(F_{ox}) = 2.35eV$ pour $V_{GS} - V_{DS} = 5V$. La pente l' est évaluée à 1.6, ce qui permet d'obtenir l'énergie d'ionisation $\phi_{i,h} = 2.8eV$ donnant une valeur proche de la littérature [11].

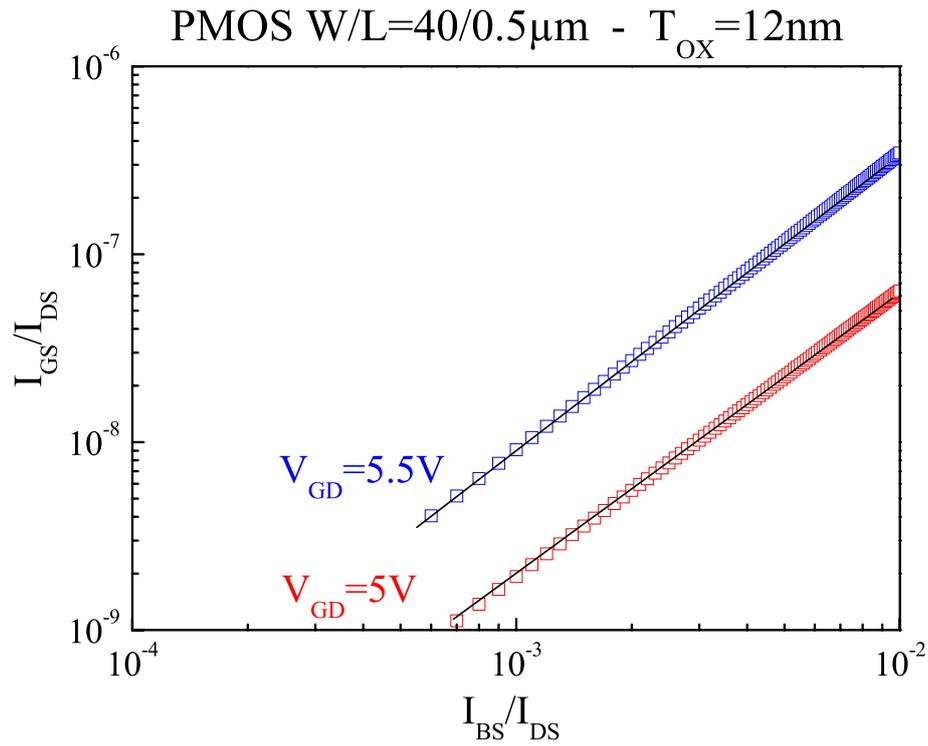


FIG. IV.8 – Relation entre le courant Substrat et le courant de Grille pour les transistors PMOS de la technologie T1.

Partant de l'expression générale de la durée de vie dans le cas où ΔN_{ot}^- est le mécanisme dominant [7, 11] :

$$\tau = \frac{C_5}{I_{BS}} \exp\left(\frac{\phi_{ot,e}}{q\lambda_e\xi_m}\right) \quad (\text{IV.24})$$

En substituant le terme exponentiel de cette équation par le rapport I_{GS}/I_{DS} de l'équation (IV.22) nous obtenons la formule donnant la durée de vie pour un taux de dégradation donné exprimé à l'aide du courant de Grille électronique :

$$\begin{cases} \frac{\tau I_{BS}}{W} = C_5 \left(\frac{I_{GS}}{I_{DS}}\right)^{-r} \\ r = \frac{\phi_{ot,e}}{\phi_{b,e}(F_{ox}) + \phi_{i,h} \frac{\lambda_e}{\lambda_h}} \end{cases} \quad (\text{IV.25})$$

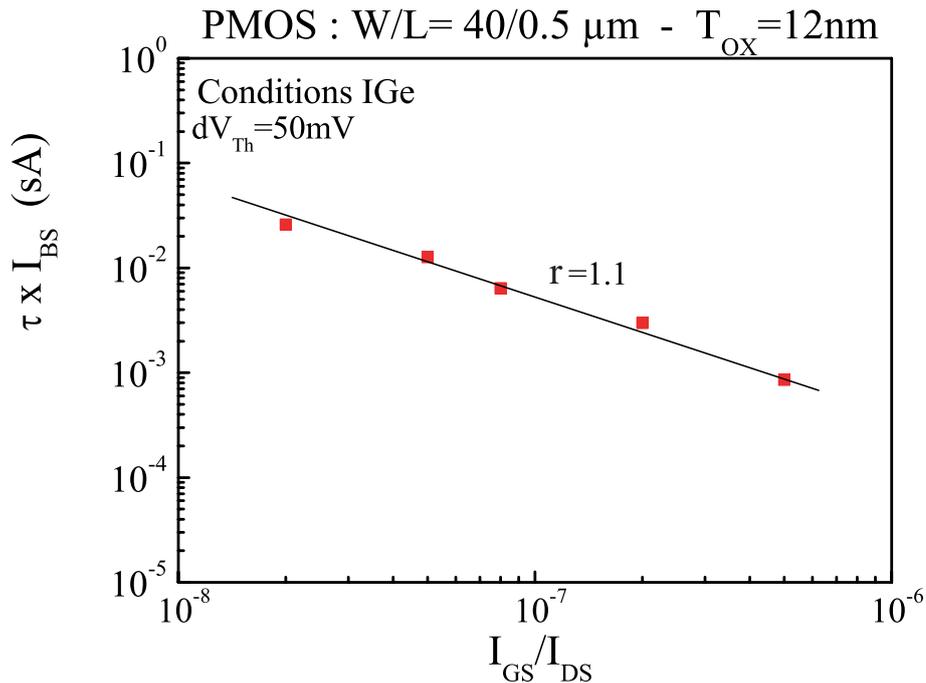


FIG. IV.9 – Tracé de la durée de vie avec le modèle $\tau/I_{BS} - I_{GS}/I_{DS}$ pour les transistors PMOS de la technologie T1 soumis au stress HE

La Fig. IV.9 montre la durée de vie tracée avec ce modèle pour un stress au maximum du courant de Grille électronique pour 50mV de variation de la tension de seuil. La pente est de $r = 1.1$. Le calcul de l'énergie de création des pièges dans l'oxyde donne $\phi_{ot,e} = 6.75\text{eV}$ qui est très proche de la valeur obtenue par Ning (6.5eV) [18]. Cette énergie est nettement supérieure à l'énergie requise pour rompre la liaison Si-H 3.1eV, ou les liaisons de type O-H (4.4eV), mais inférieure à celle de Si-O (7.98eV) [19, 20]. Ce n'est donc pas ce dernier mécanisme (rupture de la liaison Si-O) qui est retenu pour la création du piège dans l'oxyde des transistors PMOS de la technologie T1 soumis au stress HE.

IV.1.5 Synthèse sur les modèles LEM et LHM

Avant d'exposer les méthodologies qui permettent d'extraire les durées de vie à partir des données expérimentales, faisons une première synthèse sur les relations exprimant les courants de porteurs chauds, mesurés au Substrat et à la Grille, ainsi que sur les relations de durée de vie qui en découlent. Dans un but de clarté et de synthèse les résultats sont rassemblés dans deux tableaux (IV.2, IV.3), le premier regroupe les expressions des courants de porteurs chauds et le second les durées de vie exprimées avec ces derniers.

Les relations du Tab. IV.2 sont établies à partir de considérations probabilistes, à la base du développement du modèle LEM puis LHM, par extrapolation aux transistors PMOS. Des observations empiriques ont permis de relier le taux de génération des défauts au temps, et à ces

	NMOS	PMOS
I_B	$I_{BS} = C_1 I_{DS} \exp\left(-\frac{\phi_{i,e}}{q\lambda_e\xi_m}\right)$	$I_{BS} = C'_1 I_{DS} \exp\left(-\frac{\phi_{i,h}}{q\lambda_h\xi_m}\right)$
I_G	$I_{GS} = C_4 I_{DS} \exp\left(\frac{-\phi_{b,e}}{q\lambda_e\xi_m}\right)$	$N_{ot}^- : I_{G,e} = C'_1 I_{DS} \exp\left[-\frac{1}{q\xi_m}\left(\frac{\phi_{i,h}}{\lambda_h} + \frac{\phi_{b,e}(F_{ox})}{\lambda_e}\right)\right]$ $N_{it} : I_{G,h} = C'_4 I_{DS} \exp\left(\frac{-\phi_{b,h}}{q\lambda_h\xi_m}\right)$

TAB. IV.2 – Courants de porteurs chauds mesurés au Substrat et à la Grille, obtenus à l'aide du LEM pour le transistor NMOS et le transistor PMOS. Dans le cas du NMOS le courant de trous à la Grille est peu probable, pour le PMOS il faut distinguer les deux courants de trous et électronique.

courants de porteurs chauds à la Grille et au Substrat [7, 1, 11, 10]. Les relations de durée de vie en fonction de I_{GS} et I_{BS} ont pu être établies (Tab. IV.3).

Avant d'aborder la durée de vie des dispositifs soumis à des injections uniformes, examinons la Fig. IV.10. On y a représenté les durées de vie extrapolées des transistors NMOS et PMOS pour les trois technologies T1, T2 et T3, pour les trois cas de stress $V_{GS} \approx V_{GS}$, $V_{GS} \approx V_{DS}/2$, et bas V_{GS} .

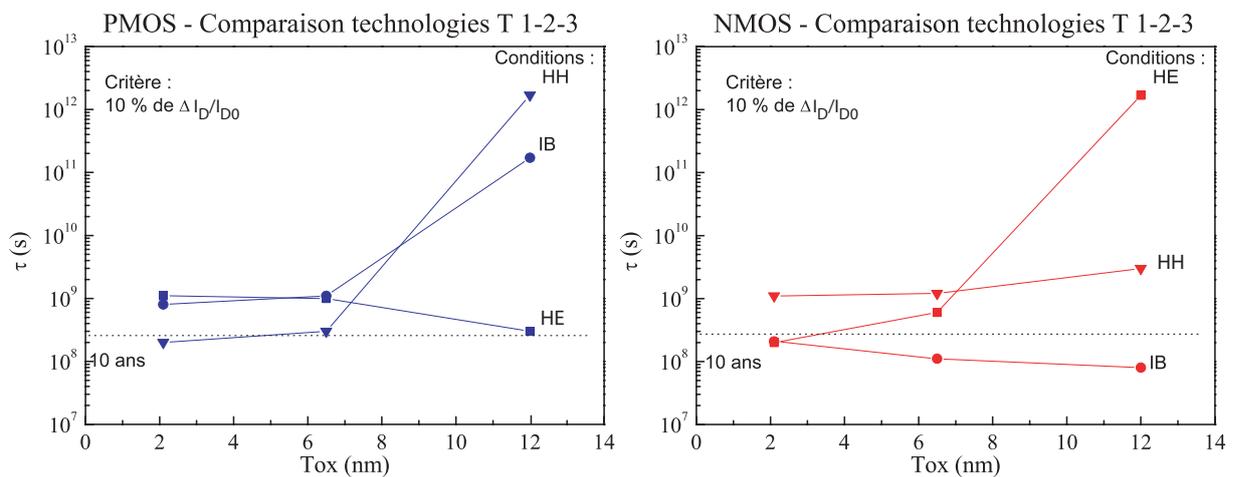


FIG. IV.10 – Durées de vie extrapolées pour les NMOS et PMOS des technologies T1, T2 et T3 pour les conditions de stress $V_{GS} \approx V_{GS}$, $V_{GS} \approx V_{DS}/2$, et bas V_{GS} ; le critère choisi est une variation de 10% du courant de Drain (régime linéaire).

Pour le PMOS, tous les cas de stress donnent lieu à une durée de vie supérieure aux 10 ans (3.15×10^8 s) préconisés. Le cas HH devient le pire cas dès $6.5nm$, dans les structures MDD

Polarisation	NMOS	PMOS
	$\tau \propto \exp\left(\frac{b}{V_{DS}}\right)$	$\tau \propto \exp\left(\frac{b'}{V_{DS}}\right)$
$V_{GS} \approx V_{DS}/5$	HH $\tau = (I_{BS})^{-t}$	HE $\frac{\tau I_{BS}}{W} = C_3 \left(\frac{I_{BS}}{I_{DS}}\right)^{-j}$ $\frac{\tau I_{BS}}{W} = C_4 \left(\frac{I_{GS}}{I_{DS}}\right)^{-r}$
$V_{GS} \approx V_{DS}/2$	IB $\frac{\tau I_{DS}}{W} = C_3 \left(\frac{I_{BS}}{I_{DS}}\right)^{-k}$	IB $\frac{\tau I_{BS}}{W} = C_3 \left(\frac{I_{BS}}{I_{DS}}\right)^{-j}$
$V_{GS} \approx V_{DS}$	HE $\frac{\tau I_{DS}}{W} = C_5 \left(\frac{I_{GS}}{I_{DS}}\right)^{-u}$	HH $\frac{\tau I_{DS}}{W} = C_6 \left(\frac{I_{BS}}{I_{DS}}\right)^{-k'}$

TAB. IV.3 – Relations donnant les durées de vie d'après les différents moniteurs : l'inverse de la tension de Drain au cours du stress, courant Substrat, courant de Grille. Les deux premières lignes sont des relations classiques de durée de vie, les deux suivantes se déduisent des courant de porteurs chauds prédits par les modèle LEM et LHM. Dans ces quatre formules, les coefficients k , j , k' , u , t et r sont définis par $k = \phi_{it,e}/\phi_{i,e}$, $j = \phi_{ot,e}\lambda_h/\phi_{i,h}\lambda_e$, $k' = \phi_{it,h}/\phi_{i,h}$, $u = \phi_{ot,e}/\phi_{b,e}(F_{ox})$ et $r = \phi_{ot}/(\phi_{b,e}(F_{ox}) + \phi_{i,h}\frac{\lambda_e}{\lambda_h})$.

pour les entrées/sorties I/O et le cœur logique. Le courant de Grille électronique est de moins en moins sévère avec la réduction de T_{ox} .

Pour le NMOS, les 10 ans ne sont pas atteints dans le cas IB pour les trois épaisseurs d'isolant. Le cas HE devient le pire cas, et atteint le même niveau de dégradation que le maximum du courant Substrat. Cependant, nous avons montré que les ratio V_{GS}/V_{DS} à fort champ ne correspondent plus aux mêmes conditions d'injections à la tension nominale.

IV.1.6 Expression de la durée de vie dans le cas des injections uniformes

Les stress uniformes ne font pas intervenir de porteurs chauds et ne représentent pas les pire cas de dégradation des structures MOSFETs. Néanmoins il peut s'agir d'un mode d'injection auquel peuvent être confrontés des circuits tels que les inverseurs ou les points mémoires. Le fait que la différence de potentiel latérale, le long du canal, soit nulle ($V_{DS} = 0$), implique que la probabilité d'injection des porteurs dans l'oxyde est indépendante de la position dans le canal. Cette dépendance avec la tension de Grille implique une dépendance avec le champ

dans l'oxyde F_{ox} . En considérant le Tab. III.7 on constate que $V_{FB} - \psi_S \approx 0.2V$ ce qui permet d'approximer :

$$F_{ox} \approx \frac{V_{GS} - 0.2}{T_{ox}} \quad (IV.26)$$

Cette relation met en évidence l'autre paramètre dont dépend l'injection uniforme : l'épaisseur de l'oxyde. C'est en effet T_{ox} qui va déterminer le mode de transport : Fowler Nordheim (FN) ou direct (DT). L'épaisseur transitoire entre les deux mécanismes se situe autour de $3nm$.

a) Durée de vie en fonction de la tension de Grille

La façon la plus simple de représenter la durée de vie dans le cas des injections uniformes est suivant la tension de Grille appliquée pendant le stress. Que ce soit en mode FN ou DT, l'expression du courant de Grille (et du nombre de porteurs injectés) s'exprime sous forme d'une exponentielle de la tension dans l'oxyde V_{ox} , très dépendante de la tension de Grille. Ainsi, on peut, à la manière du modèle en $1/V_{DS}$ dans le cas des porteurs chauds, exprimer la durée de vie comme :

$$\tau \propto \exp\left(\frac{C_{11}}{V_{GS}}\right) \quad (IV.27)$$

Sur la Fig. IV.11, on a tracé la durée de vie en fonction de $|V_{GS}|$ des PMOS et NMOS des technologies T1 et T3. Ceci permet d'étudier l'influence de l'épaisseur de l'oxyde sur l'influence des injections uniformes. Dans le cas des oxydes épais ($12nm$), il est nécessaire d'appliquer de fortes tensions à la grille pour observer un vieillissement significatif : pour $|V_{GS}| \approx 10V$, l'injection se fait en mode FN. Le NMOS montre une durée de vie plus importante que le PMOS ($\tau_N = 3.6 \times 10^{13}s$ et $\tau_P = 4.3 \times 10^9s$). Pour les dispositifs de la technologie T3 à oxyde ultra mince ($\tau_P = 6 \times 10^9s$), l'injection est en mode tunnel direct. Le NMOS requiert des tensions de Grille élevées pour observer une dégradation nette dans des temps raisonnables. On observe deux comportements vis-à-vis de la tension de Grille : pour des faibles tensions de stress la pente est différente. Ceci s'explique par des gammes de champ dans l'oxyde différentes, créant des quantités de défauts différentes.

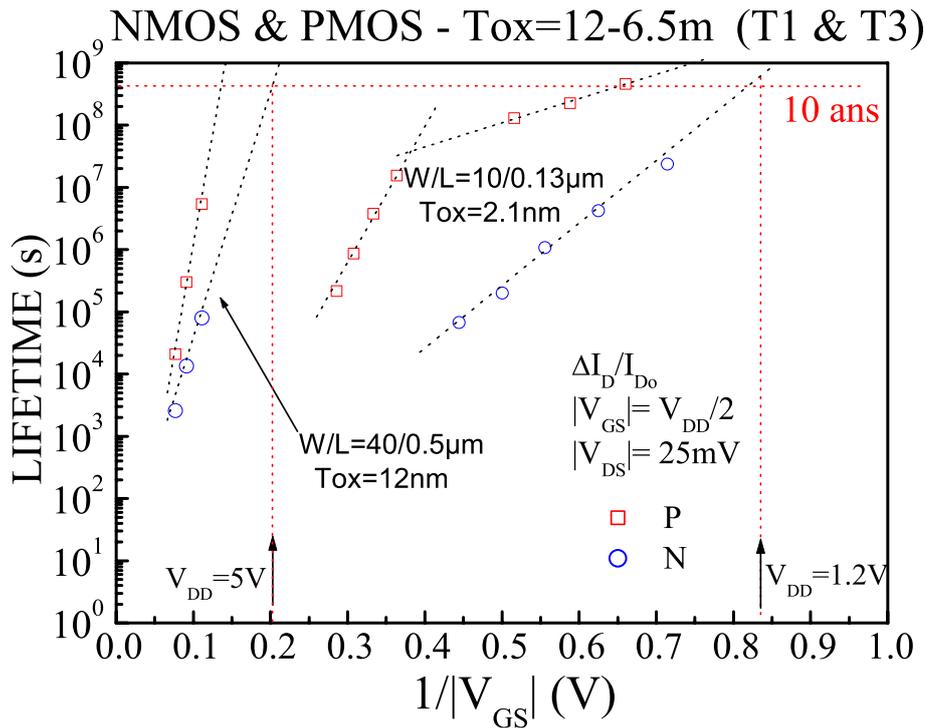


FIG. IV.11 – *Durée de vie en fonction de la tension de Grille des transistors à canal n et p des technologies T1 et T3 pour le stress uniforme.*

b) Durée de vie en fonction du courant de Grille

En l'absence de champ latéral, la tension de Grille nécessaire à l'injection des porteurs dans l'oxyde est d'autant plus importante que l'oxyde est épais. On évalue la quantité de porteurs injectés en mesurant le courant de Grille, ou encore la densité de courant dans la Grille. La Fig. IV.12 montre les densités de courant mesurées pour les trois épaisseurs d'oxyde 2.1nm, 6.5nm et 12nm.

L'oxyde mince montre une forte densité de courant qui s'explique par le fait que les porteurs ont besoin d'une faible énergie pour être injectés. Pour les autres épaisseurs d'oxyde, les tensions V_{GS} sont beaucoup plus importantes, mais les densités plus faibles, avec une pente moins prononcée pour l'oxyde le plus épais. La densité de courant, et donc le courant de Grille, sont donc révélateurs du mécanisme d'injection pour les stress uniformes. Ceci justifie de tracer la durée de vie en fonction du courant de Grille, pour lequel on trouve une relation de la forme :

$$\tau \propto (I_{GS})^{-m'} \quad (IV.28)$$

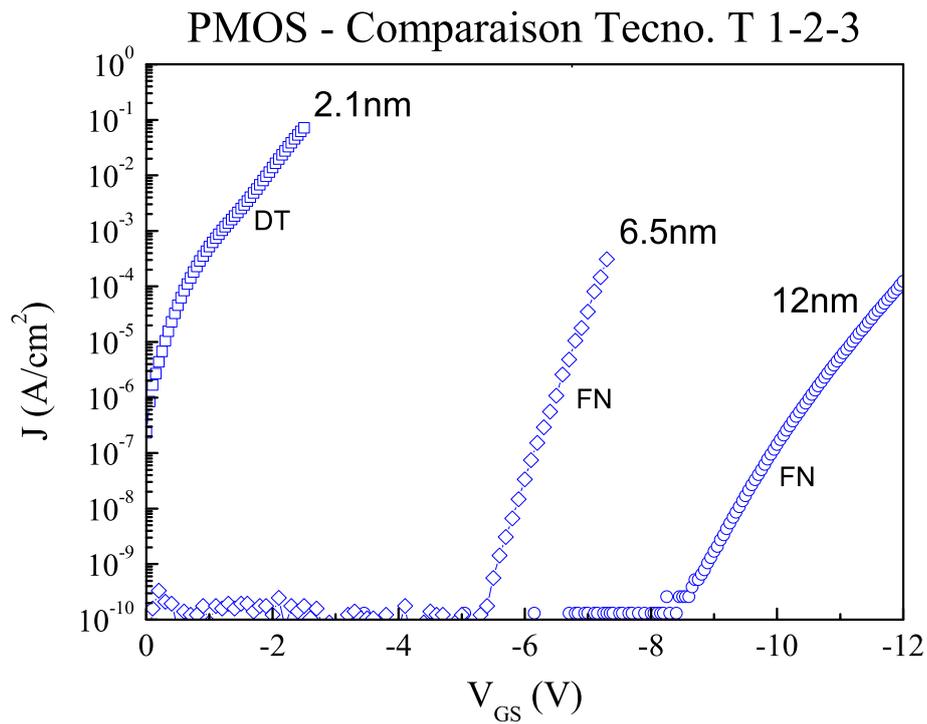


FIG. IV.12 – Densités de courant de Grille ($V_{DS} = 0V$) pour les trois technologies T1, T2 et T3.

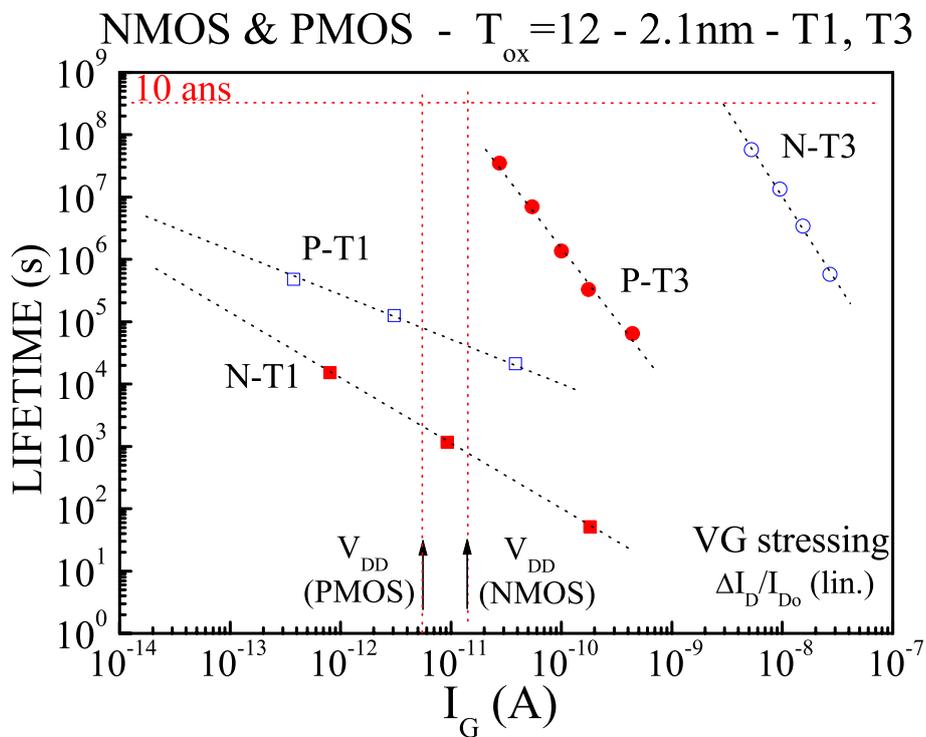


FIG. IV.13 – Durée de vie en fonction du courant de Grille des transistors à canal n et p des technologies T1 et T3 pour le stress uniforme.

Sur la Fig. IV.11 on observe la durée de vie pour les stress $V_{G\pm}$ mais exprimée à l'aide du

courant de Grille. Pour les oxydes minces, les durées de vie extrapolées à V_{DD} ont des valeurs en accord avec celles de la Fig. IV.13 : $\tau_N = 1.1 \times 10^{14} s$ et $\tau_P = 4.3 \times 10^9 s$. Les valeurs de courant aux quelles on extrapole la durée de vie ne sont pas pas les mêmes pour les PMOS et les NMOS. Pour les oxydes épais, il est difficile de donner la durée de vie extrapolée à V_{DD} du fait que pour ces tensions de Grille, I_{GS} est inférieur à la sensibilité du banc expérimental $\approx 10^{-13} A$. Pour les oxydes minces, le courant de Grille est beaucoup plus important (mode DT) et se mesure facilement à V_{DD} . Enfin sur la figure n'apparaît pas le fait que les gammes de tension de Grille appliquées pour les stress sont très différentes : 11V à 15V pour T1 contre 2V à 3V pour T3/T4.

IV.2 Durée de vie des transistors à oxydes ultra minces soumis aux injections de porteurs chauds

Dans les modélisations de la durée de vie basées sur les modèles LEM et LHM, τ est extrapolée à l'aide de la dépendance des courants de porteurs chauds, injectés à partir du canal. Les expressions qui découlent de ces considérations ne tiennent pas compte du transport des porteurs par courants tunnels directs, qui deviennent importants avec l'augmentation de V_{GS} , pour des oxydes de Grille tels que $T_{ox} < 3nm$. Nous allons voir ici quelle est l'influence de ces phénomènes tunnels sur le calcul de la durée de vie.

IV.2.1 Les oxydes minces et le concept des "*porteurs chanceux*"

Le pire cas de dégradation pour les NMOS et PMOS de la technologie T3 est l'injection à $V_{GS} = V_{DS}$, cas HE et HH, respectivement. Dans le NMOS le courant de Grille est bien plus important que dans le PMOS mais nettement moins dégradant, à cause de la mobilité supérieure des électrons et de leur hauteur de barrière plus faible que pour les trous. La Fig. IV.14 montre la durée de vie des transistor PMOS de la technologie T3, soumis à la dégradation HH. Les trois longueurs représentées présentent la même pente de -1.78 , ceci reflète le fait que le mécanisme responsable est identique pour les trois géométries : ΔN_{it} dominant. Par contre son impact varie proportionnellement à la longueur du canal. Ceci est la conséquence de la réduction de l'intensité du champ électrique latéral (et donc du nombre de porteurs chauds générés) avec la longueur du canal (au sein d'une même technologie).

IV.2.2 Modèle de séparation des porteurs chauds et froids

Dans cette partie, nous allons nous attarder sur le calcul de la durée de vie pour le stress $V_{GS} = V_{DS}$ dans le PMOS. Nous avons vu que dans ces conditions d'injection, la combinaison des mécanismes d'injections de porteurs chauds au Drain et d'injections de porteurs moins énergétiques à la source, donne un schéma de dégradation qui n'est pas trivial à analyser.

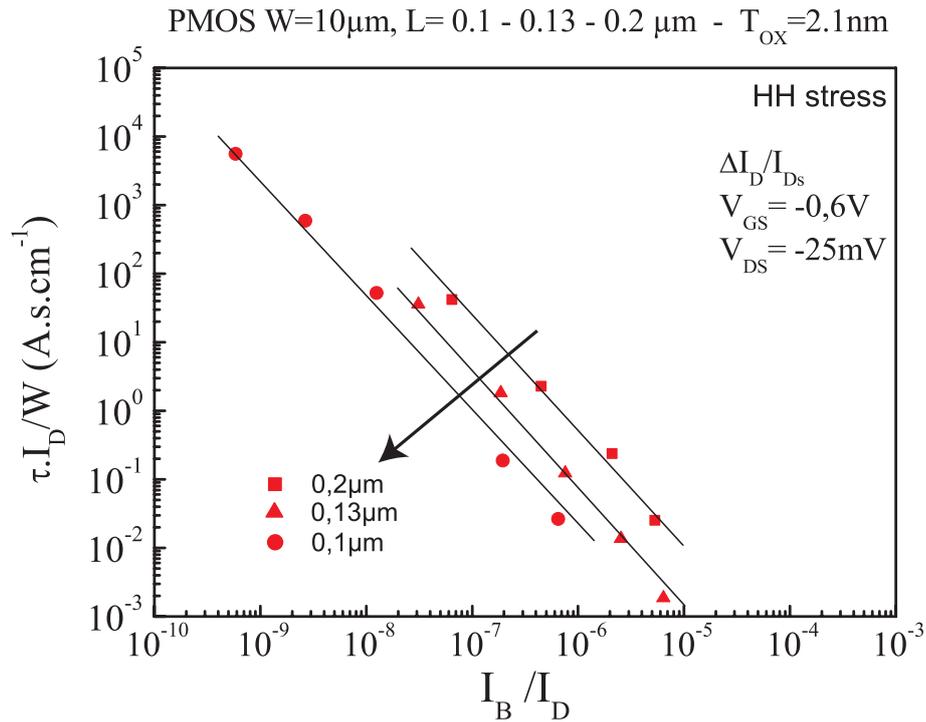


FIG. IV.14 – Représentation de la durée de vie des transistors PMOS de la technologie T3/T4 soumis au stress HH, représentée en fonction des courants de Drain et de Substrat.

Concernant le NMOS, les stress VG+ ont montré (III.2.4.b.iii) un niveau de dégradation faible même à des tensions élevées. Ceci implique que le courant tunnel influe peu sur la dégradation globale résultant de la contrainte à $V_{GS} = V_{DS}$ dans le NMOS. Finalement le cas HE pour le NMOS donne une quantité de défauts essentiellement localisés au Drain.

En revanche, dans le transistor PMOS, le flux de trous qui traversent l'oxyde par effet tunnel direct, s'avère être générateur d'une quantité de défauts à l'interface assez importante. La comparaison des modes de stress HH et VG- confirme leur similitude. Nous avons décrit au Chapitre précédent (III.2.4.c), un modèle simple qui permet de séparer le courant tunnel du courant de trous chauds. Ce modèle nous a conduit à l'équation :

$$\frac{I_G^{HH}}{I_G^{DT}} = \frac{1}{1 - \alpha} \left(1 - \frac{\Delta L}{L_{eff}} \right) \quad (IV.29)$$

où I_G est le courant mesuré à la grille dans les conditions $V_{GS} = V_{DS}$ et I_G^{DT} celui mesuré quand $V_{DS} = 0V$. α est un paramètre faiblement dépendant de V_{GD} . Le modèle est basé sur la mesure des courants de Grille. Nous allons donc maintenant relier la durée de vie en $\tau - I_{GS}$ à notre modèle de séparation. La section précédente a montré que la durée de vie peut être mise sous la forme :

$$\tau = BI_G^{-n} \quad (\text{IV.30})$$

où B est une constante et n la pente dans l'échelle Log-Log. Ayant considéré que les deux mécanismes (DT et HH) étaient indépendants, nous pouvons écrire :

$$\tau_{tun} = B_1(I_G^{tun})^{-n_1} \quad (\text{IV.31})$$

$$(\text{IV.32})$$

$$\tau_{inj} = B_2(I_G^{inj})^{-n_2} \quad (\text{IV.33})$$

où τ_{tun} et τ_{inj} sont les durées de vie du transistor PMOS s'il subissait uniquement les injections tunnel et de trous chauds pendant le stress à $V_{GS} = V_{DS}$. B_1 , B_2 sont des constantes, et n_1 et n_2 sont les pentes sur une échelle Log-Log. On peut exprimer le courant en fonction du temps nécessaire pour atteindre le niveau de dégradation choisi :

$$I_G = \left(\frac{\tau}{B}\right)^{-1/n} \quad (\text{IV.34})$$

et ainsi reformuler (III.13) comme :

$$\left(\frac{\tau}{B}\right)^{-1/n} = \left(\frac{\tau_{tun}}{B_1}\right)^{-1/n_1} + \left(\frac{\tau_{inj}}{B_2}\right)^{-1/n_2} \quad (\text{IV.35})$$

On peut finalement exprimer la durée pour le stress HH en fonction des durées de vie des deux mécanismes pris indépendamment :

$$\tau = B \left[\left(\frac{\tau_{tun}}{B_1}\right)^{-1/n_1} + \left(\frac{\tau_{inj}}{B_2}\right)^{-1/n_2} \right]^{-n} \quad (\text{IV.36})$$

τ_{tun} et τ_{inj} s'obtiennent en traçant la durée de vie avec en abscisse les courants de Grille correspondant aux deux mécanismes, grâce à la technique présentée au chapitre III (III.2.4.c). La Fig.

IV.15 montre ces tracés et les compare au stress uniforme (VG-). On constate que le mécanisme tunnel tun , pris isolé, donne une durée de vie proche du cas VG-. Le mécanisme correspondant à l'injection de trous chauds (inj) implique la durée de vie la plus faible. Dans le tableau IV.4

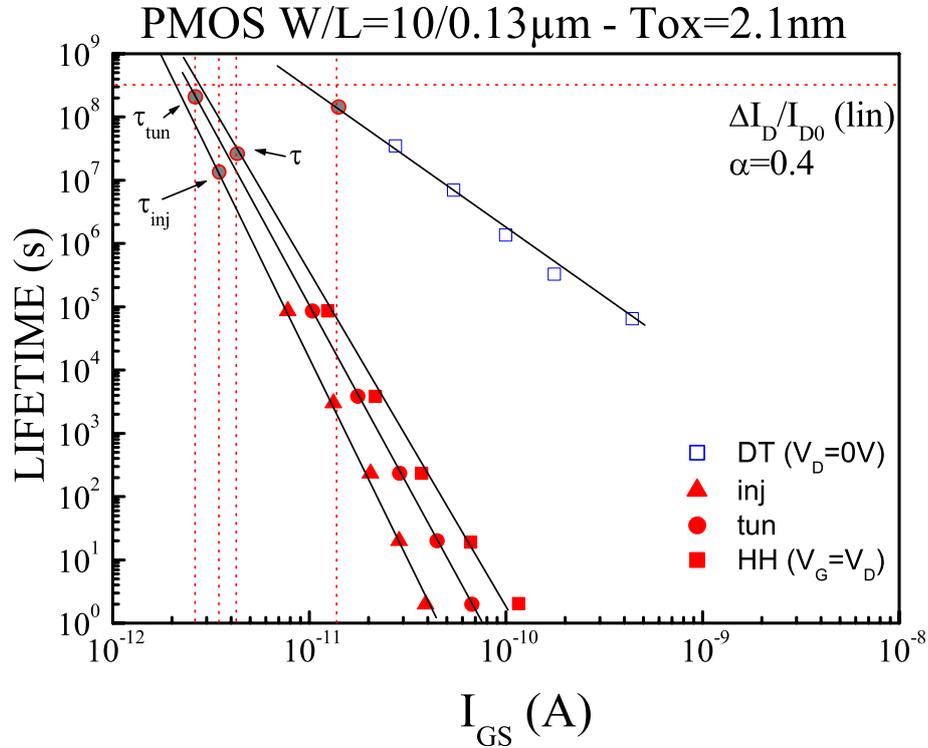


FIG. IV.15 – Comparaison des durées de vie exprimées en fonction du courant de Grille pour les transistors PMOS de la technologie T3 soumis au stress HH d'une part et VG- d'autre part.

nous avons réuni les données obtenues sur la Fig. IV.15.

mécanisme :	inj	tun	HH	VG-
τ (s)	1.2×10^7	1.9×10^8	2.7×10^7	1.45×10^8
B (sA^{-1})	10^{-67}	10^{-57}	10^{-47}	10^{-16}
n	6.63	5.7	4.7	2.3

TAB. IV.4 – Durée de vie pour les stress HH et VG-, et paramètre des courbes obtenues par le modèle de séparation, pour le transistor PMOS de la technologie T1.

Pour le stress HH la durée de vie est plus faible que pour VG-, conformément à nos observations précédentes : $\tau_{HH} = 2.7 \times 10^7 s$ et $\tau_{DT} = 1.45 \times 10^8 s$. Le rapport entre ces deux quantités est de 0.26, soit la valeur de $(L_{eff} - L_{tun})/(1 - \alpha)$. Le calcul de τ , c'est à dire la durée de vie globale pour le stress HH, avec (IV.36) donne une valeur de 3.5×10^7 . Cette valeur est proche de celle obtenue à partir des données expérimentales (2.7×10^7).

Ces résultats permettent de valider notre modèle simple de séparation des porteurs. Les injections de trous chauds, bien qu'elles soient en quantité inférieure, est le mécanisme dominant. C'est ce que nous avons observé en comparant les injections uniformes ($V_{DS} = 0V$) et localisées ($V_{GS} = V_{DS}$) au chapitre III (III.4.2).

IV.3 Influence du choix du moniteur

Un certain nombre de techniques pour modéliser la durée de vie des transistors MOSFET vient d'être exposé. C'est autant de moniteurs possibles pour déterminer la durée de vie. Chacun d'eux modélise un mécanisme de dégradation spécifique. Un choix doit être fait pour valider le développement des nouvelles technologies de fabrication. En effet, selon les standards de l'industrie, un procédé de fabrication est dit viable si les dispositifs qui en découle atteignent dix ans de fonctionnement (soit $3.1536 \times 10^8 s$) sans que ses caractéristiques ne varient de plus de 10%. Cette barrière des dix ans tend aujourd'hui à être réduite du fait de la difficulté pour l'atteindre. Il n'est pas rare pour les technologies les plus avancées de ne pas dépasser les 5 – 6 ans. C'est donc un critère qui va témoigner de la réduction des performances des dispositifs en fonction du temps de fonctionnement.

Dans ce contexte, le choix du moniteur devient crucial pour la caractérisation de la durée de vie. En effet, comme nous allons le voir des disparités nettes peuvent apparaître entre des résultats obtenus avec des moniteurs différents. Bien que ceci donne l'opportunité de disposer de méthodes d'extrapolation plus ou moins sévères, ces variations peuvent nuire à la bonne compréhension de l'influence des différents mécanismes (ΔN_{it} , ΔN_{ot}^{\pm}) à l'origine du vieillissement.

Les Tab. [IV.5](#), [IV.6](#), [IV.7](#) et [IV.8](#) rassemblent les calculs de durée de vie pour les transistors PMOS et NMOS de la filière $0.13\mu m-2.1nm$ d'oxyde de Grille, soumis à des injections de porteurs chauds dans les conditions de polarisations correspondantes aux cas HH et IB. Dans ces tableaux nous avons comparé les méthodes en $\tau - 1/V_{DS}$ et $\tau - I_{BS}$, qui constituent le moyen le plus rapide et pratique pour l'extraction de la durée de vie, et les méthodes basées sur les LEM et LHM, qui nécessitent un traitement plus long. Les coefficients C et m sont respectivement les abscisses à l'origine et les pentes des tracé $\tau I_{BS} - I_{BS}/I_{DS}$.

PMOS : $W/L = 10/0.2\mu m$ - Stress IB					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{BS} - I_{BS}/I_{DS}$	C	m
a	$4.54 \times 10^{22} s$	$5.18 \times 10^{17} s$	$1.07 \times 10^{24} s$	3.8×10^{-05}	-2.20024
b	$3.15 \times 10^{19} s$	$2.30 \times 10^{15} s$	$4.50 \times 10^{20} s$	1.67×10^{-04}	-1.81407
c	$5.06 \times 10^{19} s$	$3.27 \times 10^{15} s$	$7.42 \times 10^{20} s$	1.49×10^{-04}	-1.83978
d	$1.14 \times 10^{23} s$	$1.47 \times 10^{18} s$	$2.61 \times 10^{24} s$	1.58×10^{-04}	-2.17785
e	$1.35 \times 10^{22} s$	$3.19 \times 10^{17} s$	$2.56 \times 10^{23} s$	3.49×10^{-04}	-2.04801
PMOS : $W/L = 10/0.13\mu m$ - Stress IB					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{BS} - I_{BS}/I_{DS}$	C	m
a	$1.53 \times 10^{20} s$	$3.18 \times 10^{18} s$	$6.93 \times 10^{20} s$	2.79×10^{-06}	-2.34321
b	$4.60 \times 10^{18} s$	$1.38 \times 10^{17} s$	$2.19 \times 10^{19} s$	4.37×10^{-06}	-2.15973
c	$1.63 \times 10^{17} s$	$7.25 \times 10^{15} s$	$7.80 \times 10^{17} s$	1.30×10^{-05}	-1.95181
d	$1.45 \times 10^{18} s$	$6.43 \times 10^{16} s$	$6.35 \times 10^{18} s$	1.30×10^{-04}	-1.94205
e	$2.74 \times 10^{17} s$	$1.54 \times 10^{16} s$	$1.39 \times 10^{18} s$	1.96×10^{-04}	-1.85151
PMOS : $W/L = 10/0.1\mu m$ - Stress IB					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{BS} - I_{BS}/I_{DS}$	C	m
a	$1.77 \times 10^{17} s$	$3.63 \times 10^{14} s$	$9.60 \times 10^{15} s$	6.08×10^{-06}	-2.09664
b	$1.25 \times 10^{18} s$	$1.42 \times 10^{15} s$	$4.74 \times 10^{16} s$	2.03×10^{-06}	-2.24299
c	$3.72 \times 10^{16} s$	$1.02 \times 10^{14} s$	$2.38 \times 10^{15} s$	6.52×10^{-06}	-2.01715
d	$5.33 \times 10^{27} s$	$5.15 \times 10^{22} s$	$1.33 \times 10^{25} s$	1.40×10^{-08}	-3.57106
e	$5.07 \times 10^{15} s$	$3.48 \times 10^{13} s$	$5.34 \times 10^{14} s$	1.88×10^{-04}	-1.75303

TAB. IV.5 – *Durée de vie des transistors PMOS de longueur $L = 0.20 - 0.13 - 0.1\mu m$ de la technologie T3 soumis au stress IB. Signification des polarisations: **a** - $V_{GS} = -0.6V$, $V_{DS} = -25mV$; **b** - $V_{GS} = V_{DS} = -0.6V$ FWD; **c** - $V_{GS} = V_{DS} = -0.6V$ REV; **d** - $V_{GS} = V_{DS} = -1.2V$ FWD; **e** - $V_{GS} = V_{DS} = -1.2V$ REV.*

Dans ce premier tableau, on observe que la modélisation en I_{BS}/I_{DS} est la plus optimiste, la plus pessimiste étant la représentation en $\tau - I_{BS}$. Le bon compromis semble donc d'exprimer la durée de vie en fonction de l'inverse de la tension de Drain, directement proportionnel au champ latéral dans le canal. Ces observations valent pour les trois longueurs présentées. Remarquons également que la durée de vie extrapolée à V_{DD} croît avec la longueur du canal, dû au champ électrique moins élevé.

PMOS : $W/L = 10/0.2\mu m$ - Stress HH					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS} - I_{BS}/I_{DS}$	C	m
a	$1.57 \times 10^{12} s$	$6.37 \times 10^{09} s$	$4.37 \times 10^{09} s$	$2.35 \times 10^{13} s$	-1.78902
b	$1.77 \times 10^{11} s$	$7.89 \times 10^{08} s$	$5.46 \times 10^{08} s$	$7.63 \times 10^{14} s$	-1.7492
c	$1.85 \times 10^{11} s$	$9.41 \times 10^{08} s$	$6.57 \times 10^{08} s$	$2.83 \times 10^{13} s$	-1.70221
d	$4.15 \times 10^{13} s$	$2.52 \times 10^{11} s$	$1.77 \times 10^{11} s$	$1.35 \times 10^{10} s$	-1.67851
e	$2.01 \times 10^{13} s$	$1.33 \times 10^{11} s$	$9.43 \times 10^{10} s$	$2.40 \times 10^{10} s$	-1.62815
PMOS : $W/L = 10/0.13\mu m$ - Stress HH					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS} - I_{BS}/I_{DS}$	C	m
a	$2.54 \times 10^{09} s$	$3.86 \times 10^{07} s$	$4.48 \times 10^{07} s$	$5.65 \times 10^{13} s$	-1.84998
b	$2.83 \times 10^{09} s$	$3.62 \times 10^{07} s$	$4.21 \times 10^{07} s$	$1.23 \times 10^{13} s$	-1.92166
c	$6.26 \times 10^{08} s$	$1.24 \times 10^{07} s$	$1.42 \times 10^{07} s$	$2.31 \times 10^{12} s$	-1.72494
d	$4.34 \times 10^{11} s$	$7.04 \times 10^{09} s$	$8.15 \times 10^{09} s$	$1.79 \times 10^{10} s$	-1.82284
e	$2.47 \times 10^{12} s$	$2.51 \times 10^{10} s$	$2.97 \times 10^{10} s$	$5.34 \times 10^{12} s$	-2.05798
PMOS : $W/L = 10/0.1\mu m$ - Stress HH					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	I_{BS}/I_{DS}	C	m
a	$1.16 \times 10^{09} s$	$7.56 \times 10^{07} s$	$7.76 \times 10^{07} s$	$8.47 \times 10^{11} s$	-1.63514
b	$5.10 \times 10^{08} s$	$3.56 \times 10^{07} s$	$3.66 \times 10^{07} s$	$1.53 \times 10^{10} s$	-1.56942
c	$6.72 \times 10^{08} s$	$4.61 \times 10^{07} s$	$4.72 \times 10^{07} s$	$1.09 \times 10^{10} s$	-1.5986
d	$1.02 \times 10^{11} s$	$6.64 \times 10^{09} s$	$6.80 \times 10^{09} s$	$5.46 \times 10^{09} s$	-1.65021
e	$2.93 \times 10^{11} s$	$1.74 \times 10^{10} s$	$1.79 \times 10^{10} s$	$7.06 \times 10^{09} s$	-1.68488

TAB. IV.6 – Durée de vie des transistors PMOS de longueur $L = 0.20 - 0.13 - 0.1\mu m$ de la technologie T3 soumis au stress HH. Signification des polarisations : **a** - $V_{GS} = -0.6V$, $V_{DS} = -25mV$; **b** - $V_{GS} = V_{DS} = -0.6V$ FWD; **c** - $V_{GS} = V_{DS} = -0.6V$ REV; **d** - $V_{GS} = V_{DS} = -1.2V$ FWD; **e** - $V_{GS} = V_{DS} = -1.2V$ REV.

Pour ce second tableau, nous étudions la durée de vie pour le stress HH des transistors PMOS en fonction des modèles d'accélération. Ici la représentation $\tau - 1/V_{DS}$ est la plus optimiste, alors que les modèles en I_{BS} et I_{BS}/I_{DS} donnent des résultats très proches pour l'ensemble des polarisations de mesures. Seule la représentation suivant la tension de Drain semble en accord avec la longueur du canal (diminution de la durée de vie avec la longueur de Grille), les deux autres représentations donnent des résultats très semblables pour les deux plus petites longueurs considérées ici.

NMOS : $W/L = 10/0.2\mu m$ - Stress IB					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS} - I_{BS}/I_{DS}$	C	m
a	$4.73 \times 10^{14} s$	$1.96 \times 10^{14} s$	$1.42 \times 10^{16} s$	$1.51 \times 10^{06} s$	-2.63049
b	$2.73 \times 10^{15} s$	$1.04 \times 10^{15} s$	$9.20 \times 10^{16} s$	$1.57 \times 10^{06} s$	-2.74755
c	$8.60 \times 10^{13} s$	$3.75 \times 10^{13} s$	$2.28 \times 10^{15} s$	$1.34 \times 10^{06} s$	-2.52086
d	$1.14 \times 10^{17} s$	$4.23 \times 10^{16} s$	$4.19 \times 10^{18} s$	$2.10 \times 10^{05} s$	-2.82609
e	$1.10 \times 10^{15} s$	$4.82 \times 10^{14} s$	$2.82 \times 10^{16} s$	$2.40 \times 10^{05} s$	-2.49702
NMOS : $W/L = 10/0.13\mu m$ - Stress IB					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS} - I_{BS}/I_{DS}$	C	m
a	$3.45 \times 10^{11} s$	$5.30 \times 10^{11} s$	$5.00 \times 10^{10} s$	$1.98 \times 10^{07} s$	-2.55315
b	$4.20 \times 10^{11} s$	$6.49 \times 10^{11} s$	$6.29 \times 10^{10} s$	$3.90 \times 10^{07} s$	-2.51745
c	$2.34 \times 10^{11} s$	$3.52 \times 10^{11} s$	$3.39 \times 10^{10} s$	$1.64 \times 10^{07} s$	-2.53727
d	$8.42 \times 10^{13} s$	$1.27 \times 10^{14} s$	$9.12 \times 10^{12} s$	$5.18 \times 10^{07} s$	-2.89165
e	$4.32 \times 10^{12} s$	$6.51 \times 10^{12} s$	$5.92 \times 10^{11} s$	$1.25 \times 10^{06} s$	-2.60324
NMOS : $W/L = 10/0.1\mu m$ - Stress IB					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS} - I_{BS}/I_{DS}$	C	m
a	$3.52 \times 10^{11} s$	$2.29 \times 10^{10} s$	$1.64 \times 10^{11} s$	$2.66 \times 10^{08} s$	-2.73316
b	$4.53 \times 10^{11} s$	$3.10 \times 10^{10} s$	$2.14 \times 10^{11} s$	$6.43 \times 10^{08} s$	-2.68597
c	$3.80 \times 10^{11} s$	$2.42 \times 10^{10} s$	$1.75 \times 10^{11} s$	$2.21 \times 10^{08} s$	-2.75276
d	$8.99 \times 10^{13} s$	$5.06 \times 10^{12} s$	$4.06 \times 10^{13} s$	$5.54 \times 10^{07} s$	-2.92359
e	$2.39 \times 10^{13} s$	$1.38 \times 10^{12} s$	$1.08 \times 10^{13} s$	$2.55 \times 10^{07} s$	-2.88156

TAB. IV.7 – *Durée de vie des transistors NMOS de longueur $L = 0.20 - 0.13 - 0.1\mu m$ de la technologie T3 soumis au stress IB. Signification des polarisations : a - $V_{GS} = 0.6V$, $V_{DS} = 25mV$; b - $V_{GS} = V_{DS} = 0.6V$ FWD; c - $V_{GS} = V_{DS} = 0.6V$ REV; d - $V_{GS} = V_{DS} = 1.2V$ FWD; e - $V_{GS} = V_{DS} = 1.2V$ REV.*

Pour ce troisième tableau nous avons cette fois rassemblé les calculs de durée de vie des transistors NMOS soumis aux injections de porteurs chauds dans la condition du maximum du courant substrat. Les résultats sont globalement homogènes sur l'ensemble des longueurs et des conditions de mesure. La durée de vie décroît logiquement avec la longueur, sauf pour la représentation $\tau - I_{BS}/I_{DS}$ qui donnent aux dispositifs de $0.1\mu m$ une longévité supérieure à ceux dont la longueur de Grille est de $0.13\mu m$. Dans le mesure où IB est le pire cas de stress et que $L_{eff} = 0.1\mu m$ est une longueur critique pour la technologie T3-GO1, on peut supposer que les résultats expérimentaux présentent une marge d'erreur plus importante pour cette géométrie.

NMOS : $W/L = 10/0.2\mu m$ - Stress HE					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS}/W - I_{GS}/I_{DS}$	C	m
a	$1.60 \times 10^{17} s$	$7.37 \times 10^{15} s$	$1.04 \times 10^{16} s$	$2.63 \times 10^{08} s$	-2.78116
b	$3.18 \times 10^{17} s$	$1.41 \times 10^{16} s$	$1.99 \times 10^{16} s$	$3.03 \times 10^{08} s$	-2.81154
c	$4.36 \times 10^{16} s$	$2.23 \times 10^{15} s$	$3.12 \times 10^{15} s$	$2.24 \times 10^{08} s$	-2.71889
d	$2.47 \times 10^{21} s$	$6.31 \times 10^{19} s$	$9.41 \times 10^{19} s$	$3.64 \times 10^{08} s$	-3.30494
e	$1.99 \times 10^{19} s$	$6.55 \times 10^{17} s$	$9.50 \times 10^{17} s$	$4.71 \times 10^{08} s$	-3.01561
NMOS : $W/L = 10/0.13\mu m$ - Stress HE					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS}/W - I_{GS}/I_{DS}$	C	m
a	$7.74 \times 10^{11} s$	$1.68 \times 10^{13} s$	$2.60 \times 10^{13} s$	$1.33 \times 10^{08} s$	-2.66782
b	$9.19 \times 10^{11} s$	$1.91 \times 10^{13} s$	$2.94 \times 10^{13} s$	$2.62 \times 10^{08} s$	-2.63222
c	$3.24 \times 10^{11} s$	$6.63 \times 10^{12} s$	$1.02 \times 10^{13} s$	$1.22 \times 10^{08} s$	-2.61319
d	$1.13 \times 10^{13} s$	$2.33 \times 10^{14} s$	$3.58 \times 10^{14} s$	$3.98 \times 10^{07} s$	-2.61811
e	$5.71 \times 10^{12} s$	$1.22 \times 10^{14} s$	$1.89 \times 10^{14} s$	$1.27 \times 10^{07} s$	-2.65008
NMOS : $W/L = 10/0.1\mu m$ - Stress HE					
Polarisation	$\tau - 1/V_{DS}$	$\tau - I_{BS}$	$\tau I_{DS}/W - I_{GS}/I_{DS}$	C	m
a	$3.72 \times 10^{08} s$	$1.09 \times 10^{09} s$	$9.98 \times 10^{08} s$	$5.19 \times 10^{08} s$	-2.38029
b	$9.24 \times 10^{08} s$	$2.67 \times 10^{09} s$	$2.45 \times 10^{09} s$	$5.61 \times 10^{08} s$	-2.44423
c	$5.76 \times 10^{08} s$	$1.65 \times 10^{09} s$	$1.51 \times 10^{09} s$	$1.99 \times 10^{08} s$	-2.48739
d	$9.23 \times 10^{09} s$	$2.83 \times 10^{10} s$	$2.60 \times 10^{10} s$	$6.96 \times 10^{07} s$	-2.43222
e	$6.78 \times 10^{09} s$	$1.91 \times 10^{10} s$	$1.76 \times 10^{10} s$	$2.05 \times 10^{07} s$	-2.49679

TAB. IV.8 – Durée de vie des transistors NMOS de longueur $L = 0.20 - 0.13 - 0.1\mu m$ de la technologie T3 soumis au stress HE. Signification des polarisations : **a** - $V_{GS} = 0.6V$, $V_{DS} = 25mV$; **b** - $V_{GS} = V_{DS} = 0.6V$ FWD; **c** - $V_{GS} = V_{DS} = 0.6V$ REV; **d** - $V_{GS} = V_{DS} = 1.2V$ FWD; **e** - $V_{GS} = V_{DS} = 1.2V$ REV.

Le dernier tableau regroupe les extrapolations des durées de vie des transistors NMOS que nous avons soumis à la condition d'injection d'électrons HE. Pour chaque représentation, la technique d'extraction montre que la durée de vie est décroissante avec la longueur de la Grille, mais suivant la représentation, la variation de τ suivant L_{eff} n'est pas la même.

A la vue de tous ces résultats croisés, il semble peu aisé de déterminer le moniteur le plus adéquat. Suivant le type de stress appliqué, il existe une fourchette dans laquelle s'inscrit la durée de vie, sa valeur dépendant alors du modèle de représentation choisi. Toutefois les moniteurs en $1/V_{DS}$ et I_{BS} semblent universels vis à vis de plusieurs technologies ou de plusieurs dimensions au sein d'une même technologie. Ainsi pour étudier l'évolution de la durée de vie avec l'épaisseur de l'isolant de Grille, on trace (Fig. IV.16) les durées de vie des NMOS des technologies T1, T2 et T3 suivant V_{DS}^{-1} . On observe que la durée de vie décroît avec la diminution de l'épaisseur de l'oxyde. Du plus épais au plus mince on trouve : $2 \times 10^{11}s$ et 3.8×10^9s et 6.5×10^7s . Le même type de comparaison est faisable avec le courant Substrat comme moniteur, mais l'étude met en évidence une universalité des tracés de durée de vie vis-à-vis de I_{BS} (Fig. IV.2). Par contre au sein d'une même génération, pour des longueurs différentes, les courants Substrat sont assez proches (Fig. IV.2). La dernière figure compare les durées de vie des transistors NMOS de la technologie T3 de longueur $L = 0.1 - 0.13 - 0.2\mu m$ soumis aux stress IB (Fig. IV.17). Comme sur la Fig. IV.2, on observe que les transistors avec les Grilles les plus longues (au sein d'une technologie) ont les durées de vie les plus longues.

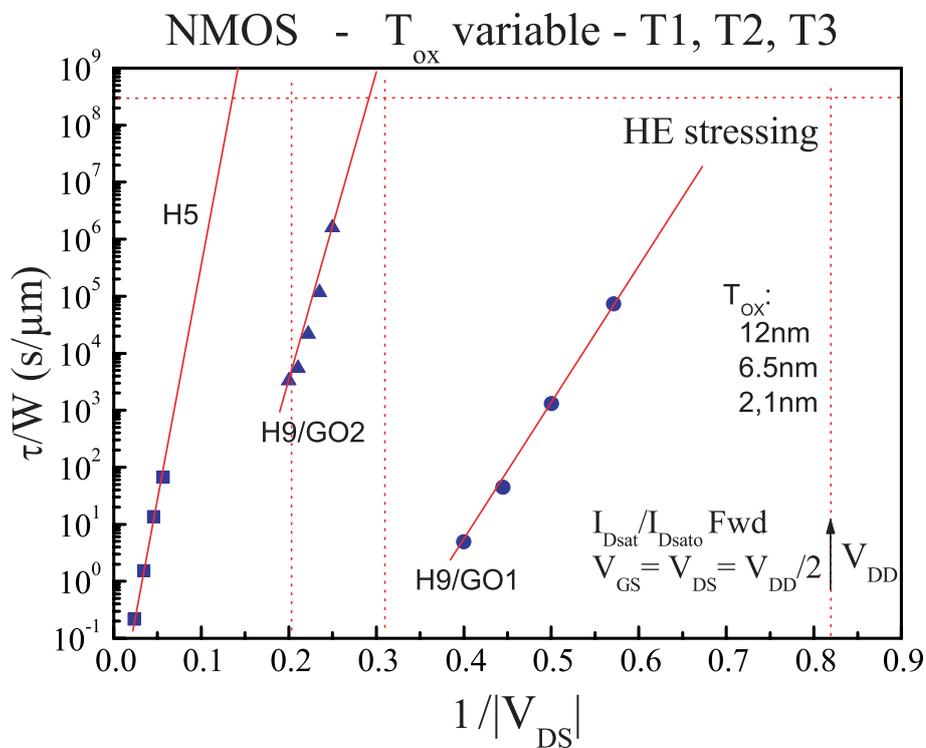


FIG. IV.16 – Comparaison des durées de vie en fonction de la tension de Drain des transistors NMOS des technologies T1, T2 et T3 soumis au stress HE.

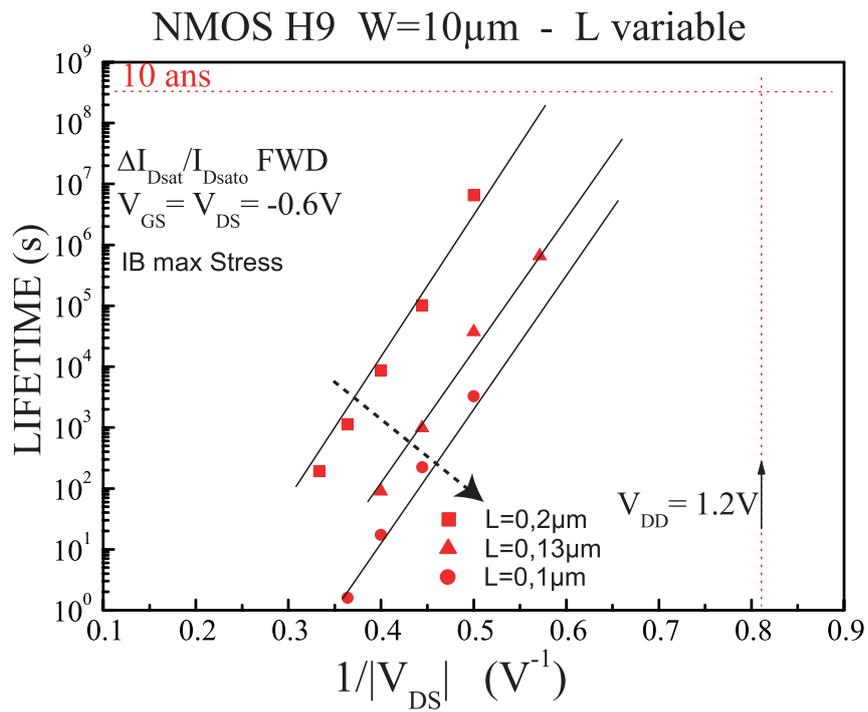


FIG. IV.17 – Comparaison des durées de vie exprimées en fonction de la tension de Drain des transistors PMOS de la technologie T3 ($L_{eff} = 0.1 - 0.13 - 0.2\mu m$) soumis au stress IB.

Conclusion

Dans ce chapitre nous avons fait l'étude des modèles qui décrivent les mécanismes d'injections de porteurs chauds dans les isolants de Grille des transistors NMOS. Basé sur une étude probabiliste, le modèle décrivant ces mécanismes, appelé modèle des électrons chanceux (LEM), permet d'évaluer les courants de porteurs à la Grille et au Substrat. Les observations empiriques des cinétiques de génération des défauts ont permis d'établir des expressions pour extrapoler la durée de vie en fonction de ces courants de porteurs chauds. Pour les transistors PMOS, un modèle équivalent est apparu pour expliquer les dégradations occasionnées par l'injection d'électrons dans l'oxyde, dont la conséquence est le piégeage de charges négatives près du Drain, entraînant une réduction de la longueur effective du canal. A l'aide de ce modèle, les courants de porteurs chauds dans la structure ont été établis, permettant d'obtenir les différentes expressions de durée de vie pour chaque mécanisme de dégradation.

Les deux modèles en électrons et en trous chauds sont valables pour les oxydes épais et se sont montrés en accord avec les résultats expérimentaux relatifs aux stress subis par les dispositifs des technologies T1 (LDD - $12nm$ - $5V$) et T2 (MDD - $6.5nm$ - $3.3V$). En revanche, nous avons mis en évidence que ces expressions pour la durée de vie ne s'appliquent plus aux oxydes ultra minces de technologie T3/T4 (MDD - $2.1nm$ - $1.2V$). Le courant tunnel très important dans ces dispositifs, n'est en effet pas pris en compte et s'avèrent dégradant. Nous avons proposé un modèle permettant de déduire l'influence, d'une part des trous chauds, et d'autre part du courant tunnel, sur la durée de vie du PMOS, où les trous se sont montrés particulièrement plus dégradants, limitant la durée de vie pour les conditions de fonctionnement à haut $V_{GS}(=V_{DS})$.

Références du Chapitre IV

- [1] T. C. Ong, P. K. Ko, and C. Hu, “Hot-carrier modeling and device degradation in surface-channel p-MOSFET’s,” vol. 37, no. 7, pp. 1658–1666, 1990. **IV.1, a), b), b), IV.1.5**
- [2] B. S. Doyle, M. Bourcerie, J.-C. Marchetaux, and A. Boudoux, “Dynamic channel hot carrier degradation in MOS transistors by enhanced-hole injection into oxide,” *IEEE Trans. Electron Devices*, vol. EDL-8, p. 237, 1987. **IV.1**
- [3] E. Takeda and N. Suzuki, “An empirical model for device degradation due to hot carrier injection,” *IEEE Trans. Electron Devices*, vol. EDL-4, no. 4, pp. 111–113, 1983. **IV.1.1, IV.1.1, IV.1.1, IV.1.1**
- [4] E. Takeda, H. Kume, T. Toyabe, and S. Asai, “Submicron MOSFET structures for minimizing channel hot-electron injection,” *IEEE Trans. Electron Devices*, vol. EDL-29, pp. 612–618, 1982. **IV.1.1**
- [5] S. Tam, P. K. Ko, and C. Hu, “Lucky electron model of hot electron injections in MOSFET’s,” *IEEE Trans. Electron Devices*, vol. ED-31, p. 1116, 1994. **a), a), b)**
- [6] Y. A. E. Mansy and D. M. Caugheyd, “Modeling weak avalanche multiplication in IGFETs and SOS transistor for CAD,” in *IEDM ech Dig.*, 1975, p. 31, 1975. **a)**
- [7] C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, and K. W. Terrill, “Hot-electron-induced mosfet degradation - model, monitor, and improvement,” *IEEE Trans. Electron Devices*, vol. ED-32, pp. 375–385, 1985. **a), a), a), a), a), b), b), b), IV.1.5**
- [8] R. Bellens, P. Heremans, G. Groeseneken, and H. E. Maes, “On the channel-length dependence of the hot-carrier degradation of n-channel MOSFET’s,” *IEEE Electron Device Lett.*, vol. 10, no. 12, pp. 553–555, 1989. **a)**
- [9] B. Doyle, M. Bourcerie, J.-C. Marchetaux, and A. Boudou, “Interface state creation and charge trapping in the medium-to-high gate voltage range ($vd/2 \leq vg \leq vd$) during hot-carrier stressing of n-MOS transistors,” *IEEE Trans. Electron Devices*, vol. EDL-37, pp. 744–754, 1990. **b)**
- [10] R. K. Mistry and B. S. Doyle, “Ac versus dc hot-carrier degradation in n-channel mosfet’s,” vol. 40, no. 1, pp. 66–104, 1993. **b), b), IV.1.5**
- [11] J. J. Tzou, C. C. Yao, R. Cheung, and H. W. K. Chan, “Hot-carrier-induced degradation in p-channel LDD MOSFET’s,” *IEEE Electron Device Lett.*, vol. 7, p. 5, 1989. **a), a), b), b)**

-), **IV.1.5**
- [12] M. Brox, A. Schwerin, Q. Wang, and W. Weber, “A model for the time- and bias-dependance of p-MOSFET degradation,” vol. 41, no. 7, pp. 392–401, 1984. **a)**
- [13] R. Woltjer, G. Paulzman, H. G. Henk, H. Lifka, and P. H. Woerlee, “Three hot-carrier degradation mechanisms in deep-submicron PMOSFET’s,” vol. 42, no. 1, pp. 392–401, 1995. **a)**
- [14] R. Woltjer, A. Hamada, and E. Takeda, “Time dependence of p-MOSFET hot-carrier degradation measured and interpreted consistently over ten orders of magnitude,” *IEEE Trans. Electron Devices*, vol. ED-40, p. 1993, 392-401. **a)**
- [15] C. Crowell and S. M. Sze, “Temperature dependence of avalanche multiplication in semiconductors,” *Appl. Phys. Lett.*, vol. 9, no. 6, pp. 242–244, 1966. **a)**
- [16] G. Van Den Bosch, G. Groeseneken, and E. Maes, “Critical analysis of the substrate hot-hole injection technique,” *Solid-State Electronics*, vol. 37, no. 3, pp. 3393–3399, 1994. **a)**
- [17] A. Bravaix, “Etudes des dégradations du transistor PMOS soumis aux injections de porteurs chauds,” Ph.D. dissertation, Univ. PARIS VII, Paris, 1991. **b)**
- [18] T. H. Ning, C. M. Osburn, and H. N. Hu, “Emission probability of hot electron from silicon into silicon dioxide,” *J. App. Phys*, vol. 48, pp. 286–293, 1977. **III.1.1, b)**
- [19] C. T. Sah, “Models and experiments on degradation of oxidized silicon,” *Solid State Electronics*, vol. 3, p. 147, 1990. **b)**
- [20] D. L. Griscom, *Defect and impurities in α -quartz and fuse Silica, dans The physics of Si-O₂ and its interfaces*. New York: Ed. S. T. Pantelides (Pergamon), 1978. **b)**

Conclusion générale

Dans cette étude nous avons montré les effets de l'intégration continue des technologies CMOS sur leurs performances et leur fiabilité lors d'injections de porteurs chauds. Dans cette optique nous avons présenté dans ce manuscrit quatre chapitres, dans lesquels nous avons montré l'influence de la réduction des dimensions des dispositifs MOSFETs sur leurs performances, sur les méthodes qui permettent d'en caractériser le vieillissement, sur les mécanismes de dégradation dans leur structure, et enfin sur les méthodes qui permettent de prédire leur durée de vie.

Ainsi dans le premier chapitre, nous avons établi les équations de base qui régissent le comportement des structures MOS, puis des transistors MOS. L'influence des petites dimensions sur le comportement des transistors a été décrite au travers des effets parasites (SCE, RSCE, NCE, DIBL, réduction de la mobilité, CLM), qui ont été modélisés. Ces effets tendent à réduire les performances électriques des dispositifs.

La seconde grande partie rassemble l'ensemble des techniques qui permettent la détection et l'analyse des dégradations des transistors MOSFETs. Tout d'abord nous avons décrit les techniques standards de suivi des caractéristiques I-V, qui permettent de localiser les dégradations à l'interface, ou une charge piégée dans l'oxyde. Ces techniques doivent être complétées par des méthodes plus poussées telles que le pompage de charges à un ou deux niveaux variables. Nous avons mis en évidence les difficultés rencontrées dans les dispositifs à oxydes ultra-minces, et nous avons proposé plusieurs solutions pour s'en affranchir dans certaines conditions de mesures.

Le troisième chapitre s'est consacré aux résultats des stress que nous avons fait subir à trois technologies avec les épaisseurs d'oxydes 12nm, 6.5nm et 2.1nm. La caractérisation des modes porteurs chauds a permis de mettre en avant les différences entre oxydes épais et fins. En effet pour les premiers, la corrélation entre les hautes et les basses tensions existent vis-à-vis de la génération de porteurs chauds. En revanche pour les oxydes fins, les conditions d'injections à la tension nominale ne correspondent pas aux conditions de vieillissement accélérés. Les deux

premières épaisseurs d'oxyde ont donné des résultats conformes à la littérature. L'oxyde le plus fin a logiquement mis en évidence un courant tunnel très important à travers l'oxyde de Grille. Ces courants se sont montrés dégradants, avec une similitude temporelle (mais avec une efficacité moindre) avec les injections de porteurs chauds, auxquelles nous avons prêté une attention particulière. C'est dans le PMOS que la cohabitation des effets des courants tunnel et porteurs chauds est la plus gênante, en raison de la plus forte efficacité des trous dans les mécanismes de dégradation. Pour les distinguer nous avons proposé un modèle simple de séparation basé sur les surfaces d'injections respectives des mécanismes : une zone d'injection de porteurs froids côté Source et une zone d'injection de porteurs chauds côté Drain. Enfin, des expérimentations d'injections alternées de trous et d'électrons chauds, ont mis en évidence la capacité des trous à générer des pièges neutres dans les oxydes épais, alors que dans les oxydes minces, ce type de pièges a disparu en raison de la nitruration de l'oxyde de Grille et de l'effet de la distance tunnel.

C'est dans le dernier chapitre que nous avons décrit en détail les modèles des électrons chanceux (LEM) et des trous chanceux (LHM), sur lesquels reposent les expressions les plus courantes de détermination de la durée de vie suivant un critère et un moniteur donné. Nous avons également présenté des méthodes très simples pour calculer la durée de vie des transistors MOSFETs dans le cas d'injections uniformes de porteurs faiblement énergétiques. Les modèles LEM et LHM étant basés sur des courants de porteurs chauds exclusivement, nous avons utilisé notre modèle de séparation développé au chapitre III pour établir une nouvelle technique de détermination de la durée de vie. Il est possible à l'aide de ce modèle d'exprimer séparément l'influence du courant tunnel et des porteurs chauds sur la durée de vie : bien qu'en nombre inférieur, les porteurs chauds injectés sont plus dégradants que les porteurs froids.

L'ensemble de ces travaux montre donc que les dégradations des transistors MOSFETs sur silicium restent un enjeu important. D'autant plus que la structure telle que nous la connaissons aujourd'hui, sera encore utilisée pour les prochaines générations technologiques, puisque certaines équipes savent fabriquer en laboratoire les dispositifs de 15 – 20nm de longueur de canal. On peut toutefois se demander si de telles dimensions autoriseront la génération de porteurs suffisamment énergétiques pour générer des défauts. Les préoccupations actuelles de la communauté, semblent plutôt se tourner vers les dégradations de type NBTI^{1, 2} (*Negative Bias Temperature Instability*), dont les effets sont de plus en plus présents avec l'intégration. Il convient néanmoins de souligner que l'ensemble des techniques et méthodes présentées dans ce manuscrit restent aujourd'hui le moyen le plus répandu pour déterminer la fiabilité porteurs chauds d'une technologie en cours de développement.

1. S. Ogawa, M. Shimaya, and N. Shiono, *Interface-trap generation at ultrathin SiO-Si interfaces during negative-bias temperature aging*, J. Appl. Phys., vol. 77, pp. 1137-1148, 1995.

2. M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, N. Revil, *New hole trapping characterization during NBTI in 65nm node technology with distinct nitridation processing*, Proceedings of the IEEE Integrated Reliability Workshop (IRW'04), p. 121-124, 2004.

ANNEXE A

Régime d'inversion faible avant saturation

Dans cette annexe, nous allons établir l'expression du courant de Drain en inversion faible, à partir de laquelle nous avons obtenu la pente sous-seuil au chapitre II (II.1.4). Nous partons de (I.45) du chapitre I, où nous avons établi l'équation du courant I_{DS} :

$$I_{DS} = -\frac{\mu_0 W}{L} \int_0^{V_{DS}} Q_n(V) dV = -\frac{\mu_0 W}{L} \int_0^{V_{DS}} (Q_{SC} - Q_D)(V) dV \quad (\text{A.1})$$

Q_n peut être considéré comme la charge totale du semi-conducteur à laquelle on doit soustraire la charge Q_D de la zone désertée sous la zone d'inversion. Le calcul du courant se ramène donc au calcul de Q_{SC} et Q_D .

A.1 Calcul des charges Q_{SC} , Q_D et Q_n

La charge du semi-conducteur I.15 a été calculée dans (I.2) :

$$Q_{sc} = \pm \sqrt{2\epsilon_{si} kT} \sqrt{n_0 (\exp^{\beta\psi_S} - \beta\psi_S - 1) + p_0 (\exp^{-\beta\psi_S} + \beta\psi_S - 1)} \quad (\text{A.2})$$

La charge de la région désertée se détermine avec la résolution de l'équation de Poisson dans laquelle on ne doit pas tenir compte des électrons (de densité n) de la couche d'inversion. On obtient ainsi la densité:

$$\rho = qp_0 \left[\exp(-\beta(\psi(y))) - 1 + \frac{n_0}{p_0} \right] \quad (\text{A.3})$$

ce qui donne dans l'équation de Poisson:

$$\frac{d^2\psi}{dy^2} = -\frac{qp_0}{\epsilon_{Si}} \left[\exp(-\beta\psi(y)) - 1 + \frac{n_0}{p_0} \right] \quad (\text{A.4})$$

On utilise la même technique qu'au chapitre I et la résolution de (A.4) donne:

$$\left(\frac{d\psi}{dy} \right)^2 = \frac{2qp_0}{\beta\epsilon_{Si}} \left[\exp(-\beta\psi(y)) - \beta\psi + \frac{n_0}{p_0}\beta\psi \right]_{\psi}^0 \quad (\text{A.5})$$

pour arriver finalement à:

$$\left(\frac{d\psi}{dy} \right) = \sqrt{\frac{2qp_0}{\beta\epsilon_{Si}}} \left[\exp(-\beta\psi(y)) - \beta\psi + \frac{n_0}{p_0}\beta\psi - 1 \right]^{1/2} \quad (\text{A.6})$$

Pour obtenir l'expression de la charge de la zone désertée, il suffit alors d'appliquer le théorème de Gauss dans les mêmes conditions que pour l'obtention de la charge Q_{SC} (I.15), soit :

$$Q_D = \sqrt{2kT\epsilon_{Si}} \left[\exp(-\beta\psi_S) - \beta\psi_S + \frac{n_0}{p_0}\beta\psi_S - 1 \right] \quad (\text{A.7})$$

Il est à noter que dans (A.6) le problème du signe devant la racine ne se pose plus puisque dans notre cas d'étude (transistor à canal N), le substrat étant de type P, la zone déplétée n'apparaît que pour $\psi_S > 0$.

La charge d'inversion Q_n est donnée par $Q_n = Q_{SC} - Q_D$. Les expressions que nous avons établies plus haut sont rigoureusement exactes, mais particulièrement lourdes à manipuler. Il convient donc de les récrire plus simplement. En inversion faible, tant que $\psi_S + V_{BS} \ll 2\phi_F - \frac{kT}{q}$, on peut écrire (puisque $\psi_S + V_{BS} - 2\phi_F < 0$): $\exp(\beta(\psi_S - V + V_{BS} - 2\phi_F)) \ll \beta\psi_S - 1$. En développant Q_{SC} au premier ordre et en remarquant que:

- $\exp(-\beta\psi_S) \ll \beta\psi_S - 1$
- $\exp(-\beta\phi_c) \ll \exp(\beta\psi_S - \beta\phi_c)$
- $\beta\psi_S \ll \exp(\beta\psi_S - \beta\phi_c)$
- $\frac{n_0}{p_0} = \exp\left(-\frac{2q\phi_F}{kT}\right)$ avec $\phi_F = V - V_{BS}$

il vient ainsi:

$$Q_{SC} = -\sqrt{2kT\epsilon_{Si}p_0} \left[\exp(\beta(\psi_S - V + V_{BS} - 2\phi_F)) + \beta\psi_S - 1 \right]^{1/2} \quad (\text{A.8})$$

$$\approx -\sqrt{2kT\epsilon_{Si}p_0} \left[1 + \frac{\exp(\beta(\psi_S - V + V_{BS} - \phi_F))}{2(\beta\psi_S - 1)} \right]^{1/2} [\beta\psi_S - 1]^{1/2} \quad (\text{A.9})$$

De même si l'on considère que ψ_S est positif:

- $\exp(-\beta\phi_s) \ll -\beta\psi_S$
- $1 - \frac{n_0}{p_0} = 1 - \frac{n_i^2}{N_A} \approx 1$
- $p_0 \approx N_A$

on arrive à:

$$Q_D = -\sqrt{2kT\epsilon_{Si}N_A} [\beta\psi_S - 1]^{1/2} \quad (\text{A.10})$$

Et finalement, on obtient l'expression de la charge d'inversion faible avant saturation:

$$Q_n = -\frac{1}{2} \sqrt{\frac{2kT\epsilon_{Si}N_A}{\beta\psi_S - 1}} \exp[\beta(\psi_S - V + V_{BS} - 2\phi_F)] \quad (\text{A.11})$$

A.2 Expression du courant I_{DS}

Afin d'établir l'expression du courant de Drain, commençons par déterminer la relation entre ψ_S , $V(x)$, V_{GS} et V_{BS} en utilisant la condition de neutralité de la charge totale du semi-conducteur:

$$Q_G + Q_{ox} + Q_{it} + Q_{SC} = 0 \quad (\text{A.12})$$

Q_{ox} représente les charges de l'isolant de grille équivalentes à l'interface $Si - SiO_2$, Q_{it} est la charge piégée dans les états d'interfaces. On suppose une densité d'états d'interfaces D_{it}

($eV^{-1}m^{-2}$) indépendante de l'énergie et on montre¹ que Q_{it} peut se mettre sous la forme²:

$$Q_{it} = Q_{it0} - qD_{it}(\psi_S - \phi_C) \quad (\text{A.13})$$

où Q_{it0} s'exprime en (Cm^{-2}). Par ailleurs, à travers la structure MOS on vérifie:

$$V_{GS} - V_{BS} = V_{ox} + \psi_S + \phi_{MS} \quad (\text{A.14})$$

Ces trois dernières équations permettent d'établir:

$$V_{GS} = \phi_{MS} - \frac{Q_{ox} + Q_{it0}}{C_{ox}} + \psi_S - \frac{Q_{SC}}{C_{ox}} + \frac{qD_{it}}{C_{ox}}(\psi_S - \phi_C) + V_{BS} \quad (\text{A.15})$$

En régime d'inversion faible la charge totale du semi-conducteur est quasiment égale à la charge de la zone désertée ($Q_{SC} \approx Q_D$). De plus, au niveau de la source, le potentiel de surface ψ_S varie avec la polarisation de la grille de $\phi_F - V_{BS}$ à $2\phi_f - V_{BS}$. On peut donc considérer que le potentiel moyen dans le canal du transistor équivaut à $\psi_{S0} = 1.5\phi_F - V_{BS}$. Ceci va nous permettre de développer Q_{SC} au premier ordre autour de ψ_{S0} :

$$Q_D \approx Q_D(\psi_{S0}) + \frac{dQ_D}{d\psi_S}(\psi_S - \psi_{S0}) = Q_{D0} - C_{D0}(\psi_S - \psi_{S0}) \approx Q_{SC} \quad (\text{A.16})$$

$$(\text{A.17})$$

$$\text{avec} \begin{cases} Q_{D0} = -\sqrt{2kT\epsilon_{Si}N_A}(\beta\psi_{S0} - 1)^{1/2} \\ C_{D0} = \sqrt{\frac{q^2\epsilon_{Si}N_A}{2kT}}(\beta\psi_{S0} - 1)^{1/2} \end{cases} \quad (\text{A.18})$$

1. Dans la bande interdite pour les états donneurs on a:

$$Q_{it} = qD_{it} \int_{E_{vs}}^{E_{cs}} (1 - f(E))dE = qD_{it} \int_{E_{Fn}}^{E_{cs}} dE = qD_{it}(E_{cs} - E_{Fn}) = qD_{it} \left(\frac{E_g}{2} - q\psi_S + q\phi_F + q\phi_C \right)$$

$$Q_{it} = Q_{it0d} - qD_{it}(\psi_S - \phi_C)$$

Pour les états de type accepteurs on intègre $f(E)$ et on arrive à $Q_{it} = Q_{it0a} - qD_{it}(\psi_S - \phi_C)$ avec $Q_{it0a} = -Q_{it0d}$

2. P. Masson, "Etude par pompage de charge et par mesures de bruit basse fréquence de transistors mos a oxynitride de grille ultra-mince," Thèse de doctorat, Lyon, 13 janvier 1999.

Q_{D0} et C_{D0} sont respectivement la charge et la capacité par unité de surface lorsque $\psi_S = \psi_{S0}$. En reportant l'expression (A.16) de Q_{SC} dans (A.14), le potentiel de surface ψ_S se met sous la forme:

$$\psi_S = \left(\frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} \right) (V_{GS} - V_{GS}^*) + \left(\frac{C_{it}}{C_{ox} + C_{it} + C_{D0}} \right) V - V_{BS} \quad (\text{A.19})$$

où $C_{it} = qD_{it}$ donne la capacité associée aux états d'interface et V_{GS}^* est une tension donnée par:

$$V_{GS}^* = \phi_{SM} - \frac{Q_{ox} + Q_{it0}}{C_{ox}} - 1.5 \frac{C_{D0}}{C_{ox}} \phi_F - \frac{Q_{D0}}{C_{ox}} \quad (\text{A.20})$$

Le potentiel de surface étant désormais connu, il est possible de trouver l'expression du courant. Pour cela, on remplace ψ_S par (A.19) dans la formule (A.11) de Q_n et I_D (I.45) devient:

$$I_{DS} = \frac{W}{L} \mu_0 \int_0^{V_{DS}} \frac{1}{2} \sqrt{\frac{2kT\epsilon_{Si}N_A}{\beta\psi_S - 1}} \exp \left[\beta \left(\frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} (V_{GS} - V_{GS}^*) - \frac{C_{ox} + C_{D0}}{C_{ox} + C_{it} + C_{D0}} V - 2\phi_F \right) \right] dV \quad (\text{A.21})$$

Le terme racine en facteur de l'exponentielle varie peu devant celle-ci et vaut à peu près $\frac{1}{2} \sqrt{\frac{2kT\epsilon_{Si}N_A}{\beta\psi_{S0} - 1}}$. On peut le sortir de l'intégrale et remarquer que multiplié par $\frac{kT}{q}$ il est égal à $C_{D0} (kT/q)^2$. (A.21) se transforme donc en³:

$$I_{DS} = \mu_0 \frac{W}{L} \left(\frac{kT}{q} \right)^2 C_{D0} \frac{C_{D0} + C_{it} + C_{ox}}{C_{D0} + C_{ox}} \exp(-2\beta\phi_F) \exp \left(\beta \frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} (V_{GS} - V_{GS}^*) \right) \left[1 - \exp \left(-\beta \frac{C_{ox} + C_{D0}}{C_{ox} + C_{it} + C_{D0}} V_{DS} \right) \right] \quad (\text{A.22})$$

Dans le cas de tensions de Drain très faibles (quelques kT/q), un développement limité du dernier terme exponentiel donne :

$$1 - \exp \left(-\beta \frac{C_{ox} V_{DS}}{C_{ox} + C_{it} + C_{D0}} \right)_{V_{DS} \rightarrow 0} \approx \beta \frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} V_{DS} \quad (\text{A.23})$$

3. P. Masson, "Etude par pompage de charge et par mesures de bruit basse fréquence de transistors mos a oxynitru de grille ultra-mince," Thèse de doctorat, Lyon, 13 janvier 1999.

L'expression du courant pour V_{DS} faible est donc:

$$I_{DS} = \mu_0 \frac{W}{L} \exp\left(\beta \frac{C_{ox}}{C_{ox} + C_{it} + C_{D0}} (V_{GS} - V_{GS}^*)\right) \exp(-2\beta\phi_F) \quad (\text{A.24})$$

ANNEXE B

Courant de fuite à travers l'oxyde de grille mince

Nous avons vu dans les chapitres II, III et IV que le courant de Grille joue un rôle de plus en plus prépondérant avec la réduction de T_{ox} . Nous allons donc décrire plus précisément dans cette annexe les phénomènes tunnels dans les oxydes de Grille des structures MOSFETs.

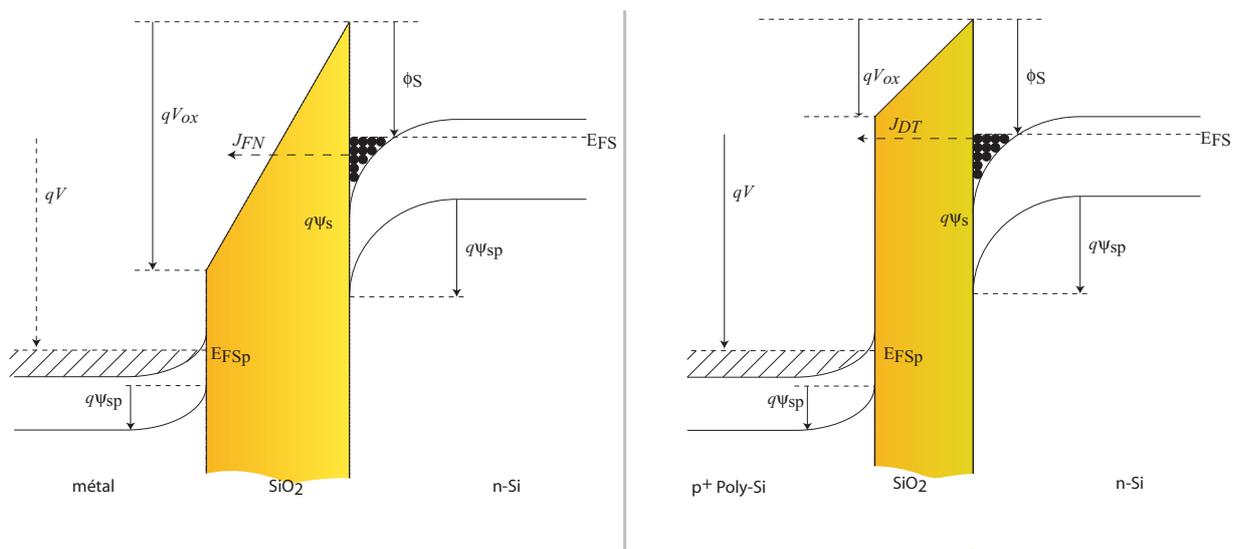


FIG. B.1 – diagramme de bande d'une structure n^+ -poly-Si/SiO₂/n-Si montrant l'effet tunnel pour les électrons de la bande de conduction de la couche d'accumulation dans le cas Fowler-Nordheim J_{FN} (à gauche), direct J_{DT} (à droite)

Le courant de grille est dû au transport des porteurs à travers l'oxyde de grille par effet tunnel. Bien qu'étant isolant, sa faible épaisseur associée à la présence d'un champ vertical F_{ox} entraînent une probabilité non nulle pour les porteurs de franchir cette barrière de potentiel. En

fonction de cette épaisseur, on considère différent types de courant tunnel:

- Courant tunnel de type Fowler-Nordheim (FN), pour des épaisseurs $T_{ox} \geq 3nm$
- Courant tunnel direct (DT) pour $T_{ox} \leq 2 - 3nm$

B.1 Courant tunnel Fowler-Nordheim

Dans le cas d'un oxyde de grille relativement épais, la probabilité de transit à travers la barrière est très faible sauf lorsque la tension de grille est suffisamment importante pour donner une forme triangulaire à la barrière ce qui revient à réduire la longueur tunnel pour les porteurs. La détermination de cette probabilité revient à calculer le coefficient de transmission de la barrière. Sur la Fig.B.1, on considère le cas d'un fort potentiel V_G positif appliqué sur la grille: une couche d'accumulation se forme dans le silicium et une zone de déplétion est présente dans le poly-silicium. La densité de courant d'électrons J_{FN} qui traverse le diélectrique de la couche d'accumulation vers la grille se calcule en supposant que les électrons sont indépendants et que l'effet tunnel est élastique (pas d'interaction avec le SiO_2)¹. On considère également que la composante transverse de l'énergie E_t des électrons (de masse transverse m_t) se conserve durant le transit. J_{FN} est alors donné par¹:

$$J_{FN} = \frac{4\pi q m_t}{h^3} \int_0^{E_{Fs}} dE \int_0^E dE_t T_t(E, E_t) \quad (\text{B.1})$$

où E est l'énergie totale de l'électron, mesurée dans la bande de conduction à l'interface Si-SiO₂, et T_t le coefficient de transmission (ou probabilité de passage). T_t s'obtient en faisant le rapport des fonctions d'onde des électrons incident et transmis. Ces fonctions d'ondes sont déterminées par la résolution de l'équation de Schrödinger en utilisant la méthode d'approximation de Brillouin, Kramers et Wentzel (WKB)²:

$$T_t(E, E_t) = \exp \left[-2 \int_0^{x_1} \kappa_{ox}(E, \bar{k}_t, x) dx \right] \quad (\text{B.2})$$

où κ_{ox} est la partie imaginaire du vecteur d'onde complexe de l'électron qui franchit la barrière et x_1 la largeur de cette barrière. κ_{ox} peut être calculé en considérant que $k_{ox} = k_t + i\kappa_{ox}$. Or dans le cadre WKB, on a :

1. M. Depas, B. Vermeire, P. W. Mertens, R. L. Van Meirhaeghe, "Determination of tunneling parameters in ultra-thin oxide layer poly-Si/SiO₂/Si structures", *Solid-states electronics*, Vol 38, No 8, 1465-1471 pp, 1995.

2. C. Cohen-Tannoudji, B. Diu, F. Laloë, "Mécanique quantique 1 et 2", *Herman éditeurs des sciences et des arts*, 1990.

$$k_{ox}^2(x) = \frac{2m_{ox}}{\hbar} [E - E_{C,ox}(x)] \quad (\text{B.3})$$

Dans cette dernière expression $E_{C,ox}(x)$ est l'énergie dans l'oxyde à la distance x correspondant à E_c et m_{ox} la masse effective de l'électron dans l'isolant. A l'aide de (B.2) et (B.3) on peut écrire³:

$$T_t(E, E_t) = \exp \left[-\frac{2}{\hbar} \int_0^{x_1} \sqrt{2m_t E_t - 2m_{ox} [E - E_{C,ox}(x)]} dx \right] \quad (\text{B.4})$$

D'autre part nous avons:

$$E_{C,ox} = \phi_S + E_{FS} - E - qF_{ox} \quad (\text{B.5})$$

$$x_1 = \frac{\phi_S + E_{FS} - E}{qF_{ox}} \quad (\text{B.6})$$

avec F_{ox} le champ électrique qui règne a travers l'oxyde et ϕ_S la hauteur de la barrière vue par les électrons de la couche d'inversion. Ceci permet d'arriver à³:

$$T_{FN}(E, E_t) = \exp \left[-\frac{4}{3} \frac{(2m_{ox})^{1/2} (\phi_S + E_{FS} - E + m_t/m_{ox} E_t)^{3/2}}{q\hbar F_{ox}} \right] \quad (\text{B.7})$$

Un développement de Taylor au premier ordre de T_{FN} autour de $E = E_{FS}$ avec $E_t = 0$ permet d'intégrer simplement (B.1)³:

$$J_{FN} = AF_{ox}^2 \exp \left[-\frac{B}{F_{ox}} \right] \quad (\text{B.8})$$

$$\text{avec } A = \frac{q^2}{16\pi^3 \hbar \phi_S} \quad (\text{B.9})$$

$$\text{et } B = \frac{4}{3} \frac{(2m_{ox})^{1/2}}{q\hbar} \phi_S \quad (\text{B.10})$$

Les valeurs de ϕ_S et de m_{OX} peuvent être déterminées en traçant sur une échelle logarithmique J_{FN}/F_{OX}^2 en fonction de $1/F_{OX}$: la pente donne B et l'intersection avec les ordonnées donne A .

3. M. Depas, B. Vermeire, P. W. Mertens, R. L. Van Meirhaeghe, "Determination of tunneling parameters in ultra-thin oxide layer poly-Si/SiO₂/Si structures", *Solid-states electronics*, Vol 38, No 8, 1465-1471 pp, 1995.

B.2 Courant tunnel direct

Dans le cas du courant tunnel direct, il faut considérer plusieurs composantes (Fig. III.34) : les électrons de la bande de conduction (ECB), les électrons de la bande de valence (EVB) qui participent au courant substrat et les trous de la bande de valence (HVB)⁴. Les expressions (B.2), (B.3) et (B.6) restent valables, mais cette fois ci, $x_1 = T_{ox}$. On obtient ainsi après intégration³:

$$T_{DT}(E, E_t) = \exp \left[-\frac{4}{3} \frac{(2m_{ox})^{1/2}}{q\hbar} \times \frac{(\phi_S + E_{FS} - E + m_t/m_{ox}E_t)^{3/2} - (\phi_S + E_{FS} - E + m_t/m_{ox}E_t - qV_{ox})^{3/2}}{F_{ox}} \right]$$

et

$$J_{DT} = \frac{AF_{OX}^2}{\left[1 - \left(\frac{\phi_S - qV_{OX}}{\phi_S} \right)^{1/2} \right]^2} \exp \left[-\frac{B}{F_{OX}} \frac{\phi_S^{3/2} - (\phi_S - qV_{OX})^{3/2}}{\phi_S^{3/2}} \right] \quad (\text{B.11})$$

Sur la Fig. B.2 on compare les mesures des densités de courant de Grille pour l'oxyde mince et pour les oxydes épais. La différence de comportement entre les deux modes est nette, le mode DT demande de très faibles tensions V_{GS} pour donner une plus forte densité de courant.

4. W-C. Lee, C. Hu, "modeling CMOS tunneling currents through ultrathin oxide due to conduction- and valence-band electrons and holes tunneling", *IEEE transactions on electronic devices*, vol 48, no. 7, 1366- 1373 pp., 2001.

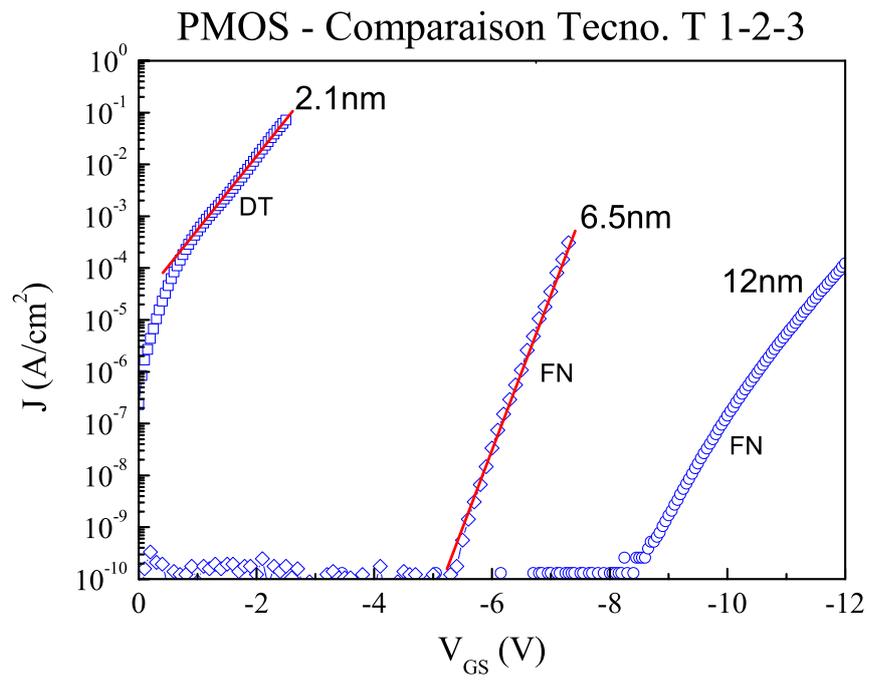


FIG. B.2 – Mesures des densités des courants pour les trois technologies T1, T2 et T3.