



Architectures avancées des transistors FinFETs: Réalisation, caractérisation et modélisation

Romain Ritzenthaler

► **To cite this version:**

Romain Ritzenthaler. Architectures avancées des transistors FinFETs: Réalisation, caractérisation et modélisation . Micro et nanotechnologies/Microélectronique. Institut National Polytechnique Grenoble (INPG), 2006. Français. <tel-01353709>

HAL Id: tel-01353709

<https://tel.archives-ouvertes.fr/tel-01353709>

Submitted on 12 Aug 2016

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.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|

THESE

pour obtenir le grade de

DOCTEUR DE L'INP Grenoble

Spécialité : "Microélectronique"

préparée au laboratoire Commissariat à l'énergie atomique (CEA)

dans le cadre de l'Ecole Doctorale "*Electronique, Electrotechnique, Automatique
Télécommunications, Signal (EEATS)*"

présentée et soutenue publiquement

par

Romain RITZENTHALER

le 17 novembre 2006

TITRE:

***Architectures avancées des transistors FinFETs:
Réalisation, caractérisation et modélisation***

DIRECTEUR DE THESE: Sorin CRISTOLOVEANU, Directeur de recherche CNRS

CO-ENCADRANT: Olivier FAYNOT, Dr. et Ing. CEA-LETI

JURY

M. J.-P. COLINGE	, Président et Rapporteur
M. C. LALLEMENT	, Rapporteur
M. G. PANANAKAKIS	, Représentant INPG
M. S. CRISTOLOVEANU	, Directeur de thèse
M. O. FAYNOT	, Co-encadrant
Mme. A. VANDOOREN	, Examinatrice
M. T. ERNST	, Examineur

Remerciements

Je tiens tout d'abord à remercier M. Benoit Giffard, chef de l'ex-laboratoire Silicium sur Isolant (LSOI), ainsi que M. Simon Deleonibus, chef du laboratoire des Nano-Dispositifs Electroniques (LNDE), pour m'avoir accueilli dans leurs laboratoires respectifs. Je tiens également à remercier M. Pierre Bichon, qui a été chef du Service Recherche Dispositif (SRD) lorsque j'ai débuté ma thèse, ainsi que M. Olivier Demolliens et M. Michel Wolny, chefs du Département NaNoTech300 (D2NT) du LETI. Merci également à M. Pierre Gentil, directeur de l'école doctorale EEATS.

J'exprime ma reconnaissance à M. Jean-Pierre Colinge pour avoir accepté de présider ma soutenance et d'être rapporteur de ce travail de thèse. Mes plus vifs remerciements vont à M. Christophe Lallement qui a également accepté d'être rapporteur. Je remercie chaleureusement Mme. Anne Vandooren, M. Georges Pananakakis et M. Thomas Ernst pour leur participation au jury.

Je tiens à remercier M. Sorin Cristoloveanu, admirable directeur de thèse apprécié de tous ses thésards. Aussi bien au niveau scientifique que relationnel, il a toujours été une source de motivation positive et attentionnée. Merci également à Olivier Faynot, mon encadrant au CEA, pour sa clarté d'esprit et sa très grande compétence.

Un grand merci aussi à Carine Jahan, la reine du process et des carnets de lots. Un remerciement spécial à Xavier Mescot de l'IMEP pour son aide lors des expériences à basses températures, et dieu sait qu'elles lui ont demandé du temps et beaucoup d'attention. Merci également à Thomas Ernst pour son approche physique et créative des phénomènes. Merci à Loan, relectrice exhaustive de ce manuscrit. Je tiens également à remercier les gens du laboratoire, aussi bien pour la vie de tous les jours que pour le travail réalisé en commun. Un grand merci donc à Cécilia Dupré et Emilie Bernard. Merci à François Andrieu, Sylvain Barraud, Thierry Poiroux et Bernard Guillaumot pour les corrections et/ou les très intéressantes discussions que j'ai pu avoir avec eux.

Mention spéciale pour Julien, mon acolyte pendant ces trois (presque quatre!) années. Merci à tous les potes qui font de la vie au travail un endroit accueillant. Si j'en oublie, qu'ils me pardonnent ou exigent une nouvelle version de ces remerciements de thèse. Plus ou moins dans l'ordre chronologique, merci donc à Manu, Damien, Gabriel El Molas, Jérôme, Luca, Julie, Olivier W., Stéphanie, Atsushi-sama (mon grand pourvoyeur de matériel informatique), Vince,

Fred, Tiziana, Juliano, Christian, Maud, Georges, Mike, Manu, Estelle, JP et Marc le petit nouveau qui court vite.

Pour les autres laboratoires, petite dédicace à Marianne (pardon Marie-anne, je sais), Erwan, Mike, Vincent, Florent (encore pardon pour ce tacle assassin), Xavier et Jérôme du LSCDP (j'espère que votre labo n'aura pas encore changé de nom d'ici là). De l'IMEP, labo sympa s'il en est, je remercie également Wipa, Samir, Kerem (guide touristique à Istanbul en free-lance), Marc, Leily, les deux Brice, Marlène, Kruno, Nico Bresson et Marilyn.

Merci également aux autres amis, de Grenoble, de Chambéry ou d'ailleurs. Matthias, même si tu t'es exilé dans les froidures alsaciennes je ne désespère pas d'être à nouveau dans la même ville que toi (à l'autre bout du monde !). Merci également à Manu, Alex, Max, Diego y Ricardo los gauchos, Amandine, Eva, Charlotte, et tous les autres potes présents ou passés.

A ma mère, mon père, ainsi que toute ma famille.

"Carac la nuit, bonjour les ennuis"

Stéphanie J.

"Tout a une fin, sauf la saucisse qui en a deux"

Frédéric M., arguant qu'il s'agit d'une vieille chanson traditionnelle allemande

Table des matières

Remerciements.....	3
Table des matières	7
Introduction générale.....	11
Lexique des grandeurs et notations utilisées	14
Abréviations et acronymes.....	16
Chapitre 1	17
Transistor MOS et technologie SOI: principes de fonctionnement	17
1 Fonctionnement général d'un transistor à semiconducteur.....	21
1.1 Structure de bandes d'un semiconducteur.....	21
1.2 Dopage d'un semiconducteur.....	21
1.3 Les différents régimes d'une structure MOS	22
1.4 Le transistor MOSFET	24
2 Les effets de canaux courts.....	28
2.1 Les enjeux de la miniaturisation.....	28
2.2 Les effets de canaux courts.....	29
2.3 Les matériaux innovants	31
3 La technologie 'silicium-sur-isolant' (SOI)	32
3.1 Introduction - principe de fonctionnement	32
3.2 Fabrication des plaques SOI.....	34
3.3 Architectures partiellement et complètement désertées.....	35
4 Conclusions	38
Chapitre 2	41
Les composants multigrilles	41
1 Introduction	45
2 Dispositifs double-grille planaires.....	46
3 Dispositifs multigrilles non planaires	46
3.1 Transistors FinFET	46
3.2 Transistors Triple-grille.....	48
3.3 Transistors Π -FET et Ω -FET.....	49
3.4 <i>Body-Tied</i> FinFETs.....	50
3.5 Mémoires FinFlash	51
3.6 Optimisation de la densité d'intégration dans les transistors verticaux.....	51
4 Dispositifs à grille enrobante - GAAFET	54

5	Technologie SON (Silicon-On-Nothing).....	54
6	Dispositifs multicanaux et nanofils	55
	6.1 Dispositifs à multicanaux.....	55
	6.2 Dispositifs à nanofils	57
7	Autres dispositifs multigrilles	58
	7.1 Structure en 'T inversé' - ITFET	58
	7.2 Dispositifs à grilles indépendantes	59
8	Conclusions	60
	Chapitre 3	71
	Fabrication et performances électriques des transistors réalisés.....	71
1	Fabrication d'un lot de transistors Ω FETs	75
	1.1 Description schématique des étapes de fabrication	75
	1.2 Description des étapes critiques.....	76
	1.3 Morphologie des transistors	78
2	Performances électriques du lot réalisé.....	79
	2.1 Courant de drain en fonction de la tension de drain $I_D(V_D)$	79
	2.2 Courant de sortie (I_{ON}) et courant à l'état bloqué (I_{OFF}).....	81
	2.3 Tensions de seuil.....	83
	2.4 Pente sous le seuil et <i>DIBL</i>	85
3	Conclusions	88
	Chapitre 4 :	91
	Analyse des mécanismes physiques spécifiques aux transistors multigrilles: aspects dimensionnels et capacitifs, modélisation des effets de couplage.....	91
1	Introduction	95
2	Modélisation du fonctionnement d'un transistor	95
	2.1 Equations classiques	95
	2.2 Effets quantiques	96
3	Influence de la structure des transistors sur leurs performances électriques.....	99
	3.1 Règles d'échelle.....	99
	3.2 Impact du nombre de grilles.....	102
	3.3 Impact de la non-verticalité des flancs.....	104
	3.4 Effets des résistance d'accès et des contacts.....	105
	3.5 Conclusions sur le contrôle des effets de canaux courts.....	106
4	Effets tridimensionnels dans les structures Triple-grille: les effets de coins.....	107
	4.1 Aspects classiques	107
	4.2 Impact de la prise en compte des effets quantiques	111
5	Couplages d'interfaces dans les Ω FETs	113
	5.1 Cas des transistors larges.....	113
	5.2 Cas des transistors étroits	119
6	DIVSB (<i>Drain Induced Virtual Substrate Biasing</i>)	128
	6.1 Effet DIVSB dans les structures Triple-grille.....	128
	6.2 Modélisation analytique du DIVSB.....	133
	6.3 Discussion	139

7	Conclusions	141
	Chapitre 5	145
	Mesures de mobilité dans les transistors ΩFETs – influence des plans cristallins.....	145
1	Introduction du chapitre	149
2	Définition de la mobilité	149
	2.1 Concept de masse effective	149
	2.2 Relation entre vitesse des porteurs et champ électrique.....	150
	2.3 Mécanismes de dégradation de la mobilité.....	151
3	Méthodes d'extraction de la mobilité	154
	3.1 Coefficients d'atténuation de la mobilité	154
	3.2 Méthode du maximum de transconductance $g_{m,max}$	154
	3.3 Méthode de la fonction Y	155
	3.4 Méthode de la dérivée seconde de l'inverse du courant de drain.....	156
	3.5 Méthode Split C-V	157
4	Mesures de mobilité sur les transistors Ω FET'	160
	4.1 Effet de la longueur de grille	160
	4.2 Effet de la largeur du canal	160
5	Mesures en température	173
	5.1 Variation des caractéristiques avec la température.....	173
	5.2 Mesures sur des transistors longs et larges.....	174
	5.3 Mesures sur des transistors longs et étroits.....	175
6	Mesure du taux de balisticité	176
7	Conclusions	178
	Conclusions et perspectives du manuscrit.....	183
	Bibliographie de l'auteur.....	187

Introduction générale

Contexte de la thèse :

En 1930, L. Lilienfeld de l'Université de Leipzig dépose un brevet dans lequel il décrit le fonctionnement d'un élément qui ressemble au transistor MOS (Métal Oxyde Semiconducteur) actuel. La réalisation pratique d'un composant à effet transistor a été faite en 1947 par les américains John Bardeen, William Shockley et Walter Brattain, chercheurs de la compagnie *Bell*. Ils ont reçu le prix Nobel de physique en 1956 pour cette découverte. Le transistor a constitué une invention déterminante sans laquelle l'électronique et l'informatique ne possèderaient pas leurs formes actuelles. Le transistor fut considéré comme un énorme progrès face au tube électronique : il est plus robuste, il fonctionne avec des tensions faibles, il peut donc être alimenté par des piles et il fonctionne instantanément une fois mis sous tension [Lilen'03].

Vers 1959, Jack Kilby des laboratoires *Texas Instruments* allait donner naissance au premier circuit intégré, regroupant sur une même surface de silicium un ensemble de composants actifs et passifs. Quelques mois plus tard, c'est l'équipe de Robert Noyce qui allait déposer à son tour une demande de brevet pour un circuit du même genre. *Texas Instruments* et d'autres fabricants comme *Fairchild Semiconductor* se lancèrent alors dans la production de circuits intégrés.

En 1971, *Intel* conçoit et commercialise le premier microprocesseur au monde, l'*Intel 4004*, fonctionnant sur 4 bits et composé de 2300 transistors. La complexité des microprocesseurs n'a depuis pas cessée d'augmenter. La croissance rapide de l'industrie de l'électronique [Dauvin'95] est basée sur l'évolution de la technologie des circuits intégrés en termes de performances et de coût par fonction. Les avancées technologiques ont été obtenues dans les quarante dernières années principalement en réduisant les dimensions des circuits dans le but d'améliorer leur taille (pour un coût de fabrication moindre) et leur fréquence de fonctionnement. Un des cofondateurs d'*Intel*, Gordon Moore, a d'ailleurs énoncé en 1965 la 'loi' empirique portant depuis son nom : le nombre de composants par circuit double approximativement tous les 18 mois [Mann'00]. Cette loi s'est imposée comme un modèle économique décrivant l'évolution technologique des circuits intégrés (Fig. 1).

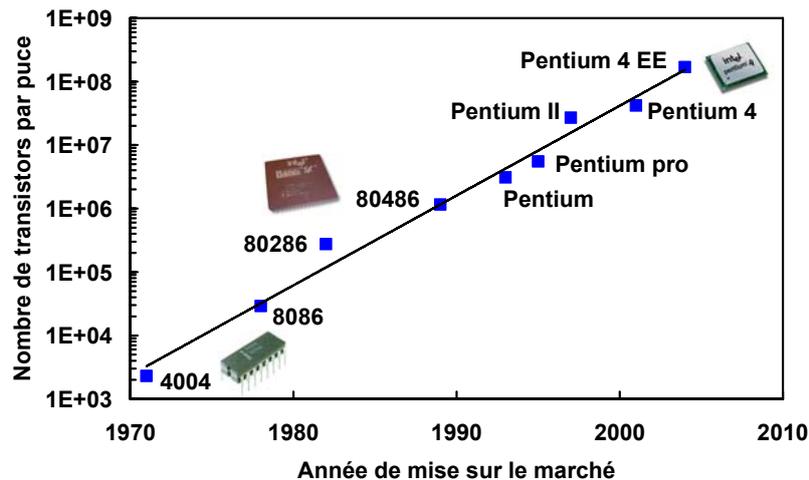


Fig. 1: Nombre de transistors par puce en fonction de leur date de mise sur le marché pour les microprocesseurs commercialisés par *Intel* (loi de Moore).

La brique de base des circuits VLSI (*Very Large Scale Integration*, i.e. Intégration à Très Grande Echelle) modernes est le transistor MOS à effet de champ (MOSFET – *Metal Oxide Semiconductor Field Effect Transistor*, i.e. Transistor à Effet de Champ Métal-Oxyde-Semiconducteur) qui contribue à lui tout seul à environ 75 % du marché mondial des semiconducteurs. Pour les applications digitales, les défis induits par la réduction des dimensions sont multiples. Il faut ainsi limiter les courants de fuites et les effets de canaux courts, augmenter le courant de saturation tout en diminuant la tension d'alimentation, contrôler les paramètres électriques du dispositif (tension de seuil, fuites) et les effets parasites.

Avec la réduction extrême des dimensions, de nombreux problèmes sévères sont apparus. La poursuite de la réduction des dimensions des transistors va de pair avec l'introduction de nouveaux matériaux dans l'empilement de grille (grilles métalliques, oxydes haute permittivité), dans le canal de conduction afin d'améliorer les propriétés de transport des porteurs ou encore dans les zones de source et drain pour diminuer les résistances d'accès ou encore améliorer l'injection des porteurs dans le canal [ITRS'03][ITRS'04].

Il est également envisagé de modifier la structure de base d'un transistor afin d'augmenter le contrôle électrostatique de la grille. En plaçant par exemple d'autres grilles à l'opposé du dispositif, la capacité de grille est augmentée et le potentiel est mieux contrôlé par l'électrode, limitant ainsi les courants de fuite. En augmentant le nombre de canaux de conduction, on augmente également le courant de saturation.

Parmi les nouveaux dispositifs étudiés dans les laboratoires de recherche, les transistors à grilles verticales FinFET s'annoncent comme de prometteurs candidats. C'est ce type d'architecture qui a fait l'objet de cette thèse.

Présentation du manuscrit :

Ce travail de thèse s'intègre dans ce contexte d'évolution rapide de la technologie des transistors MOS. Les objectifs de cette étude consistent en la réalisation, la caractérisation et la compréhension des phénomènes physiques intervenant dans les architectures FinFETs.

Le premier chapitre présente brièvement les concepts de base de la physique des semiconducteurs et des transistors MOS. Il a pour but de positionner le travail effectué dans le cadre de cette thèse dans le contexte global de la microélectronique.

Le deuxième chapitre est plus spécifiquement consacré aux architectures multigrilles. Il a pour objectif de présenter les différents transistors à grilles multiples existants en comparant les techniques de fabrication et les performances.

Le troisième chapitre présente les transistors FinFETs réalisés dans le cadre de cette thèse et leurs performances électriques.

Le but des quatrième et cinquième chapitres est de dégager les phénomènes spécifiques apparaissant dans les structures FinFETs. A travers la caractérisation électrique et la simulation numérique (simulations réalisées avec les logiciels DEVISE et DESSIS de la société ISE®), le quatrième chapitre est consacré aux phénomènes électrostatiques intervenant dans les architectures FinFETs. Les couplages de coin et latéraux sont mis en lumière, ainsi que l'influence de la face arrière et du drain à travers l'oxyde enterré.

Le cinquième chapitre est quant à lui dédié aux phénomènes de transport. A travers la caractérisation électrique, la conduction dans les flancs verticaux permettant de moduler plans et orientations cristallines est étudiée.

Références

- [Dauvin'95] J.-P. Dauvin, J. Olliver, D. Coulon, "Les composants électroniques et leur industrie", ISBN 2-13-047303-2, éditions PUF, collection Que sais-je ?, 1^{ère} édition, 1995.
- [ITRS'03] Site de *l'International Technology Roadmap for Semiconductors (ITRS)*, édition 2003, <http://public.itrs.net/Files/2003ITRS/Home2003.htm>, 2003.
- [ITRS'04] Site de *l'International Technology Roadmap for Semiconductors (ITRS)*, mise à jour 2004, <http://www.itrs.net/Common/2004Update/2004Update.htm>, 2004.
- [Lilen'03] H. Lilen, "Une brève histoire de l'électronique", ISBN 2-7117-5336-0, éditions Vuibert, 2003.
- [Mann'00] C. C. Mann, "The End of Moore's Law?", www.technologyreview.com, mai 2000.

Lexique des grandeurs et notations utilisées

Constantes physiques

Notation	Commentaire	Unités (S. I.)	Valeur approximative
ϵ_0	Permittivité diélectrique du vide	F.m ⁻¹	8.85 10 ⁻¹²
e	Charge de l'électron	C	1.60 10 ⁻¹⁹
h	Constante de Planck	J.s	6.63 10 ⁻³⁴
\hbar	Quantum de moment cinétique	J.s	$h/2\pi = 1.05 10^{-34}$
k _B (ou k)	Constante de Boltzmann	J.K ⁻¹	1.38 10 ⁻²³
m _e	Masse de l'électron 'au repos'	kg	9.11 10 ⁻³¹

Notations

Notation	Signification	Unités usuelles
α	Paramètre d'échelle lié à la géométrie	
β	Paramètre d'échelle lié au champ électrique	
χ	Affinité électronique du semiconducteur	eV
C _{DEP}	Capacité de désertion	F.m ⁻²
C _{INV}	Capacité d'inversion	F.m ⁻²
C _{OX}	Capacité d'oxyde	F.m ⁻²
D	Coefficient de diffusion	m ² .s ⁻¹
D _{it}	Densité d'états d'interface	J ⁻¹ .m ⁻²
DIBL	<i>Drain Induced Barrier Lowering</i>	mV/V
ϵ_{OX}	Permittivité diélectrique de l'oxyde de grille	F.m ⁻¹
ϵ_{Si}	Permittivité diélectrique du Si	F.m ⁻¹
ϵ_r	Permittivité relative	
E _A	Energie des niveaux accepteurs	eV
E _C	Energie de la bande de conduction	eV
E _D	Energie des niveaux donneurs	eV
E _{EFF}	Champ effectif transverse	MV.cm ⁻¹
E _F	Energie du niveau de Fermi	eV
E _{FI} (ou E _I)	Energie du niveau de Fermi intrinsèque	eV
E _{FN}	Energie du quasi-niveau de Fermi des électrons	eV
E _{FP}	Energie du quasi-niveau de Fermi des trous	eV
E _G	Energie du <i>gap</i>	eV
EOT	<i>Equivalent Oxide Thickness</i>	m

E_V	Energie de la bande de valence	eV
G	Coefficient de génération	$m^{-3}.s^{-1}$
g_m	Transconductance	$A.V^{-1}$ (ou S)
H_{FIN}	Hauteur d'aile d'un FinFET	m
I_{DS}	Courant circulant entre la source et le drain	A
I_{OFF}	Courant à l'état bloqué normalisé	$A.m^{-1}$
I_{ON}	Courant à l'état passant normalisé	$A.m^{-1}$
J	Densité de courant	$A.m^{-2}$
k	Permittivité (ou ϵ) – notation anglo-saxonne	$F.m^{-1}$
L_G	Longueur de grille	m
μ_0	Mobilité à faible champ transverse	$cm^2.V^{-1}.s^{-1}$
μ_{EFF}	Mobilité effective	$cm^2.V^{-1}.s^{-1}$
m^*	Masse de confinement des porteurs dans la direction transverse	kg
$N_{A,D}$	Dopage du canal	cm^{-3}
n_i	Concentration intrinsèque des porteurs dans le silicium	cm^{-3}
N_S	Dopage du substrat	cm^{-3}
Φ_F	Différence entre les potentiels de Fermi intrinsèque et extrinsèque du substrat	eV
Φ_S	Potentiel de surface	eV
Φ_M	Travail de sortie	eV
P_S	Puissance statique dissipée	W
Q_C	Charge de surface dans le canal	$C.m^{-2}$
Q_{DEP}	Charge de désertion dans le canal	$C.m^{-2}$
Q_{INV}	Charge d'inversion dans le canal	$C.m^{-2}$
Q_F	Densité de charges fixes dans l'oxyde	$C.m^{-2}$
Q_{IT}	Densité d'états d'interface	$C.m^{-2}$
R	Coefficient de recombinaison	$m^{-3}.s^{-1}$
R_{SD}	Résistance source/drain	$\Omega.m$
σ	Conductivité électrique	$\Omega^{-1}.cm^{-1}$
S	Pente sous le seuil	mV/décade
τ	Délai intrinsèque du transistor	s
τ_C	Valeur moyenne du temps entre chaque collision	s
T	Température	K (ou °C)
t_{OX}	Epaisseur électrique de l'oxyde de grille	m
t_{Si}	Epaisseur d'un film de silicium	m
v	Vitesse de dérive des porteurs	$m.s^{-1}$
V_D	Tension aux bornes du drain	V
V_{DD}	Tension nominale d'alimentation	V
V_{DSAT}	Tension de drain en saturation	V

Abréviations et acronymes

Nom	Signification	Traduction française
ALD	<i>Atomic Layer Deposition</i>	Depôt de couches atomiques
BOX	<i>Buried OXide</i>	Oxyde enterré
CMP	<i>Chemical Mechanical Polishing</i>	Polissage mécano-chimique
CVD	<i>Chemical Vapor Deposition</i>	Dépôt en phase vapeur
DIBL	<i>Drain Induced Barrier Lowering</i>	Abaissement de barrière induite par le drain
DIVSB	<i>Drain Induced Virtual Substrate Biasing</i>	Création d'une polarisation de face arrière virtuelle induite par le drain
DUV	<i>Deep Ultra Violet</i>	Ultra violet profond
e-beam	<i>Electronic beam</i>	Faisceau électronique
EOT	<i>Equivalent Oxide Thickness</i>	Epaisseur d'oxyde équivalente
FDSOI	<i>Fully Depleted SOI</i>	SOI complètement déserté
FET	<i>Field Effect Transistor</i>	Transistor à effet de champ
GAA	<i>Gate All Around</i>	Grille enrobante
HDD	<i>High Doped Drain</i>	Extensions fortement implantées
HTO	<i>High Thermal Oxide</i>	Oxyde déposé à haute température
LDD	<i>Low Doped Drain</i>	Extensions faiblement implantées
LOP	<i>Low Operating Power</i>	Faible puissance en fonctionnement
LSTP	<i>Low Standby Power</i>	Faible consommation en veille
MEB	<i>Scanning Electron Microscope</i>	Microscope Electronique à Balayage
MOS	<i>Metal Oxide Semiconductor</i>	Metal Oxyde Semiconducteur
PDSOI	<i>Partially Depleted SOI</i>	SOI partiellement déserté
RIE	<i>Reactive Ion Etching</i>	Gravure ionique réactive
SCE	<i>Short Channel Effect</i>	Effet de canal court
SEG	<i>Selective Epitaxial Growth</i>	Croissance épitaxiale sélective
SOI	<i>Silicon On Insulator</i>	Silicium sur isolant
SON	<i>Silicon On Nothing</i>	Silicium suspendu
SOS	<i>Silicon On Sapphire</i>	Silicium sur saphir
SRAM	<i>Static Random Access Memory</i>	Mémoire statique à accès aléatoire
STI	<i>Shallow Trench Isolation</i>	Isolation par tranchée étroite
TEM	<i>Transmission Electron Microscope</i>	Microscope électronique en transmission
VLSI	<i>Very Large Scale Integrated</i>	Intégré à très grande échelle

Chapitre 1

Transistor MOS et technologie SOI: principes de fonctionnement

Ce premier chapitre est une présentation des concepts de base de la physique des dispositifs à semiconducteurs et de la microélectronique. L'objectif de ce chapitre est de positionner le contexte de ce travail de thèse. Les structures MOS sont introduites, ainsi que les grandeurs typiques utilisées pour caractériser un transistor. Le contrôle du canal par la grille est primordial surtout quand la longueur de celle-ci devient petite, et nécessite de constantes améliorations. Brièvement, les défis successifs induits par la miniaturisation de ces dispositifs (effets de canaux courts) et les solutions classiques appliquées sont présentés. L'accent est notamment mis sur la technologie SOI, et sur les types de dispositifs découlant de cette architecture (dispositifs partiellement et complètement désertés).

Sommaire du chapitre 1

1	Fonctionnement général d'un transistor à semiconducteur.....	21
1.1	Structure de bandes d'un semiconducteur.....	21
1.2	Dopage d'un semiconducteur.....	21
1.3	Les différents régimes d'une structure MOS	22
1.4	Le transistor MOSFET.....	24
1.4.1	Régime linéaire – traitement très simplifié:	25
1.4.2	Régime de saturation.....	25
1.4.3	Régime sous le seuil.....	26
1.4.4	Le MOS extrinsèque	27
2	Les effets de canaux courts	28
2.1	Les enjeux de la miniaturisation	28
2.2	Les effets de canaux courts	29
2.2.1	L'effet de canal court (<i>roll-off</i>).....	29
2.2.2	L'effet de canal court inverse (<i>roll-up</i>).....	30
2.2.3	L'effet DIBL.....	30
2.2.4	Le perçage (<i>punch-through</i>).....	31
2.3	Les matériaux innovants.....	31
2.3.1	Les oxydes haute permittivité (<i>high-k</i>)	31
2.3.2	Les grilles métalliques	32
3	La technologie 'silicium-sur-isolant' (SOI)	32
3.1	Introduction - principe de fonctionnement	32
3.2	Fabrication des plaques SOI.....	34
3.3	Architectures partiellement et complètement désertées.....	35
3.3.1	Définition.....	35
3.3.2	Transistors SOI 'Partiellement Désertés'	35
3.3.3	Transistors SOI 'Complètement Désertés'	37
4	Conclusions	38

1 Fonctionnement général d'un transistor à semiconducteur

1.1 Structure de bandes d'un semiconducteur

Lorsque des atomes forment un réseau cristallin, les niveaux d'énergie correspondant aux orbitales des atomes tendent à s'étaler autour d'un niveau moyen et à former un continuum: on parle alors de bandes d'énergie (Fig. 2) [Ashcroft'76][Kittel'96]. Le plus haut niveau d'énergie occupé à 0 K s'appelle la bande de valence, le premier niveau d'énergie non occupé la bande de conduction. De par les propriétés des orbitales, les niveaux d'énergie situés entre bande de valence et bande de conduction ne sont pas permis: c'est la bande interdite. De plus, la différence d'énergie entre la bande de conduction et la bande de valence est appelée *gap* du matériau. L'énergie de Fermi E_F correspond à l'énergie limite qui sépare, à 0 K, les niveaux occupés des niveaux vides.

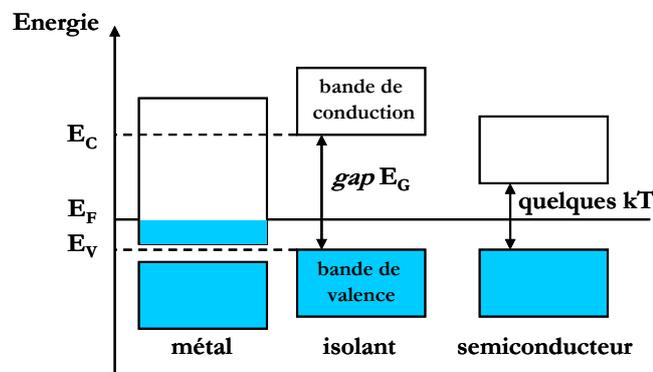


Fig. 2: Différences de structure de bandes d'énergie pour les métaux, isolants et semiconducteurs. E_F est l'énergie du niveau de Fermi, E_V l'énergie de la bande de valence et E_C l'énergie de la bande de conduction.

Si, le *gap* est de l'ordre de quelques eV, la bande de valence contient quelques électrons thermiques qui suffisent à assurer une conduction minimale: le matériau est dit semiconducteur.

Un semiconducteur est un matériau présentant une conductivité électrique intermédiaire entre un métal et un isolant ($10^{-10} \Omega^{-1}.cm^{-1} < \sigma < 10^4 \Omega^{-1}.cm^{-1}$). Ils sont primordiaux en microélectronique car ils permettent de contrôler la quantité de courant susceptible de les traverser. Les matériaux satisfaisant à cette caractéristique sont situés en haut à droite du tableau périodique de Mendeleïev, les plus connus étant le silicium et le germanium (*gaps* E_G valant respectivement 1.12 et 0.66 eV à 300 K [Sze'81]).

1.2 Dopage d'un semiconducteur

Un atome situé dans la quatrième colonne du tableau de Mendeleïev possède 4 électrons de valence [Mathieu'96]. Si dans le réseau cristallin d'un tel élément on substitue un atome de la cinquième colonne (à 5 électrons de valence), l'atome de substitution présentera 4 liaisons covalentes et un électron libre. Cet électron, qui n'est pas un électron de liaison, n'est que

faiblement lié à l'atome et peut être facilement excité vers la bande de conduction. Comme l'excitation de ces électrons ne conduit pas à la formation de trous dans ce genre de matériau, le nombre d'électrons dépasse de loin le nombre de trous. Les électrons sont des porteurs majoritaires et les trous des porteurs minoritaires. Et parce que les atomes à cinq électrons ont un électron supplémentaire à « donner », ils sont appelés atomes donneurs. Les matériaux ainsi formés sont appelés semiconducteurs de type N.

De manière symétrique pour un dopage de type P on introduit un déficit d'électron dans le matériau. On peut donc considérer que le dopant a donné une charge virtuelle positive au matériau.

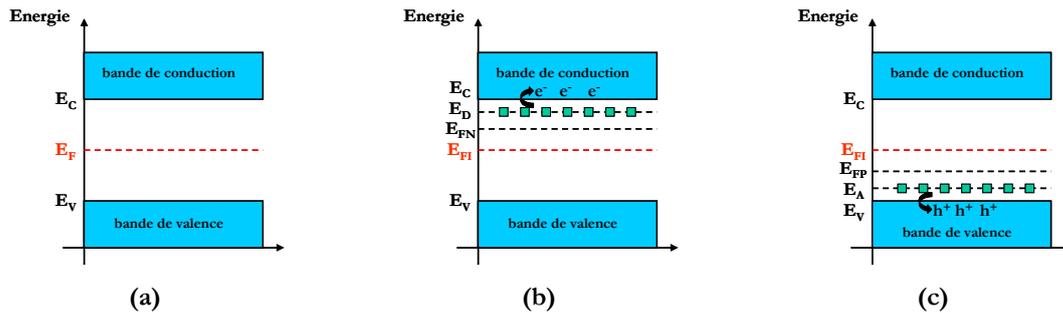


Fig. 3: Bandes d'énergie dans les cas d'un semiconducteur intrinsèque (a), dopé N (b) et dopé P (c). E_F : énergie du niveau de Fermi; E_{FP} : énergie du niveau de Fermi intrinsèque; E_V : énergie de la bande de valence; E_C : énergie de la bande de conduction; $E_{D,A}$: énergie des niveaux donneurs (accepteurs); $E_{FN,FP}$: énergie du quasi-niveau de Fermi des électrons (des trous).

On modifie ainsi les bandes d'énergie (Fig. 3), et donc considérablement les propriétés électriques d'un semiconducteur en le dopant de manière contrôlée. Pour des raisons de compatibilité technologique, c'est le Bore qui est utilisé principalement pour doper positivement un substrat Silicium alors que le Phosphore et l'Arсениc sont utilisés habituellement pour le doper négativement.

1.3 Les différents régimes d'une structure MOS

A la base des circuits intégrés numériques se trouve la structure MOS (Metal-Oxide Semiconducteur). La plus simple des structures MOS est la capacité MOS, qui consiste en un empilement substrat-diélectrique-grille.



Fig. 4: Définition du travail de sortie d'un métal (a) et de l'affinité électronique d'un semi-conducteur (b). Φ_M (ou Φ): travail de sortie; χ : affinité électronique du semiconducteur. Les autres grandeurs sont définies dans la Fig. 3.

Le travail de sortie est l'énergie à fournir pour extraire un électron du matériau (Fig. 4.a) ; l'affinité électronique d'un semiconducteur est quant à elle définie comme l'énergie à fournir pour passer de la bande de conduction au niveau du vide (Fig. 4.b).

On peut définir également Φ_F la différence entre les potentiels de Fermi intrinsèque et extrinsèque du substrat et Φ_S le potentiel de surface, correspondant à la différence entre le potentiel à l'interface et sa valeur en volume. Muni de ces grandeurs, on peut représenter la structure de bandes d'une capacité MOS à l'équilibre (Fig. 5).

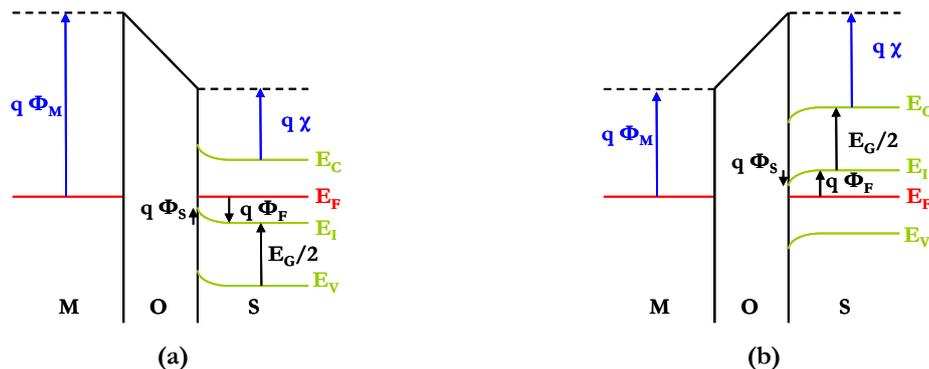


Fig. 5: Structure de bandes pour une capacité NMOS (a) et PMOS (b) à l'équilibre. E_I : énergie de Fermi du matériau intrinsèque; E_G : énergie du gap; Φ_F : différence entre les potentiels de Fermi intrinsèque et extrinsèque du substrat; Φ_M : travail de sortie du métal; Φ_S : potentiel de surface. Les autres grandeurs sont définies dans la Fig. 3.

En jouant sur la tension de commande ($V_G - V_{FB}$), on modifie l'état de la capacité. Considérons un substrat dopé P (Fig. 5.b). On distingue principalement trois zones de fonctionnement dans une capacité MOS (Fig. 6):

- ✓ L'**accumulation** apparaît typiquement pour des tensions négatives, quand les charges négatives sur la grille attirent les trous vers l'interface oxyde-semiconducteur.
- ✓ La **désertion** quant à elle apparaît pour des tensions positives. La tension positive aux bornes de la grille repousse les trous vers l'intérieur du substrat. Alors, le semiconducteur est vidé de porteurs mobiles à l'interface et une charge négative, due aux ions accepteurs ionisés, apparaît dans la zone de charge espace. La tension séparant ces régimes d'accumulation et de désertion est nommée tension de bandes plates V_{FB} . Le terme 'bandes plates' fait référence au fait qu'à cette tension la structure de bandes dans le semiconducteur est horizontale. Ceci implique qu'aucune charge n'existe dans le semiconducteur.
- ✓ L'**inversion** apparaît à des tensions se trouvant au-delà de la tension de seuil V_T . En inversion, il existe une zone chargée négativement à l'interface oxyde-semiconducteur en plus de la zone de désertion qui a atteint son extension maximale. La couche d'inversion est composée de porteurs minoritaires, qui sont attirés à l'interface oxyde-semiconducteur par les tensions de grille positives. Aux hautes fréquences, il peut arriver que la couche d'inversion soit longue à se former

à cause d'un taux de génération de porteurs minoritaires trop faible. On parle alors de **désertion profonde** car la zone de désertion est alors beaucoup plus grande qu'à l'équilibre.

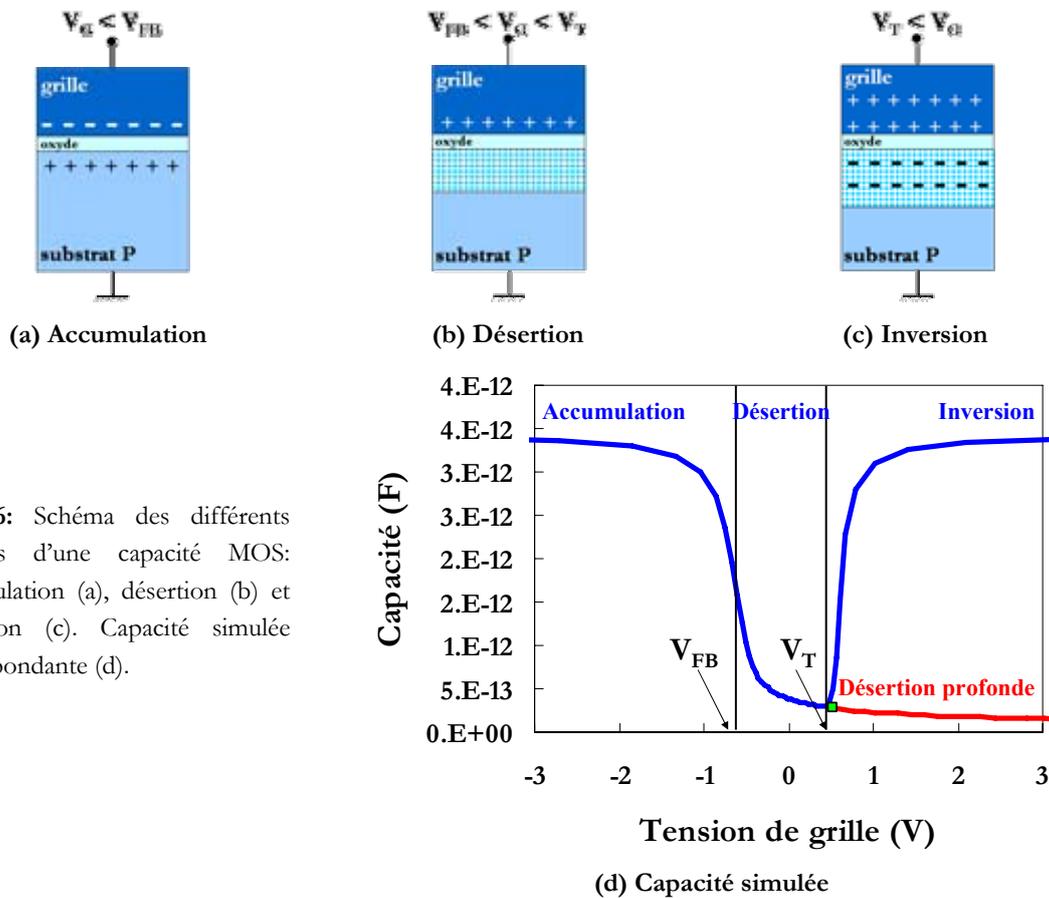


Fig. 6: Schéma des différents régimes d'une capacité MOS: accumulation (a), désertion (b) et inversion (c). Capacité simulée correspondante (d).

1.4 Le transistor MOSFET

Un transistor MOS est une source de courant commandée en tension. Par effet de champ, la tension appliquée entre la grille et la source permet de créer la couche d'inversion, et la tension entre le drain et la source déplace les porteurs présents dans cette couche d'inversion (Fig. 7).

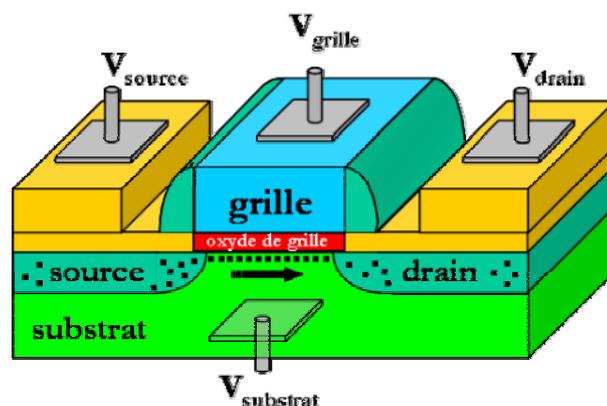


Fig. 7: Schéma de principe d'un transistor MOS classique.

1.4.1 Régime linéaire – traitement très simplifié:

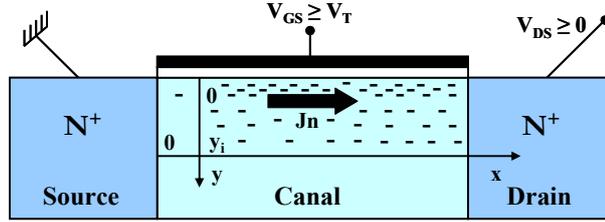


Fig. 8: Coupe longitudinale simplifiée d'une structure MOS.

On considère uniquement le transport des électrons dans une structure MOS. Le champ électrique est supposé uniquement longitudinal, et donc:

$$\vec{E} = -\vec{\nabla}(V) = -\frac{\partial V}{\partial x} \vec{u}_x \quad \text{Eq. (1)}$$

La mobilité μ d'un porteur de charge relie la vitesse de dérive d'un porteur et le champ électrique appliqué [Kittel'96]:

$$\vec{v}_n = -\mu_n \vec{E} \quad \text{Eq. (2)}$$

Au déplacement des charges correspond un courant dont la densité est définie comme la quantité de charge traversant une unité de surface pendant une unité de temps, soit:

$$J_n = -en\mu_n \frac{\partial V}{\partial x} \quad \text{Eq. (3)}$$

Le courant de drain local i_{DS} s'exprimera donc sous la forme:

$$\begin{aligned} i_{DS} &= -\int_y \int_z J_n = W \int_0^{y_i} en\mu_n \frac{\partial V}{\partial x} dy = We\mu_n \frac{\partial V}{\partial x} \int_0^{y_i} ndy \\ &= W\mu_n \frac{\partial V}{\partial x} C_{ox} (V_{GS} - V - V_T) \end{aligned} \quad \text{Eq. (4)}$$

Le courant de drain total I_{DS} est donc donné par la relation:

$$I_{DS} = \mu_n C_{ox} \frac{W}{L} \int_0^{V_{DS}} (V_{GS} - V - V_T) dV = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \quad \text{Eq. (5)}$$

1.4.2 Régime de saturation

Quand la tension drain/source augmente, le potentiel le long du canal augmente et réduit l'épaisseur de la couche d'inversion. Pour une tension de polarisation $V_D = V_{DSAT}$, la capacité MOS n'est plus en inversion coté drain. On dit alors que le canal est en régime de pincement. Le transistor continue de conduire, mais le courant de sortie n'augmente plus avec la tension de drain: on est en régime de saturation (Fig. 9). Le régime de saturation s'obtient pour un transistor peu dopé avec un oxyde mince en considérant la densité de porteurs nulle à l'interface. On obtient alors l'expression de la tension et du courant de saturation [Mathieu'96]:

$$I_{DSAT} = C_{OX} \frac{W}{2L} V_{DSAT}^2 = C_{OX} \frac{W}{2L} (V_G - V_T)^2 \quad \text{Eq. (6)}$$

avec $V_{DSAT} = V_G - V_T$.

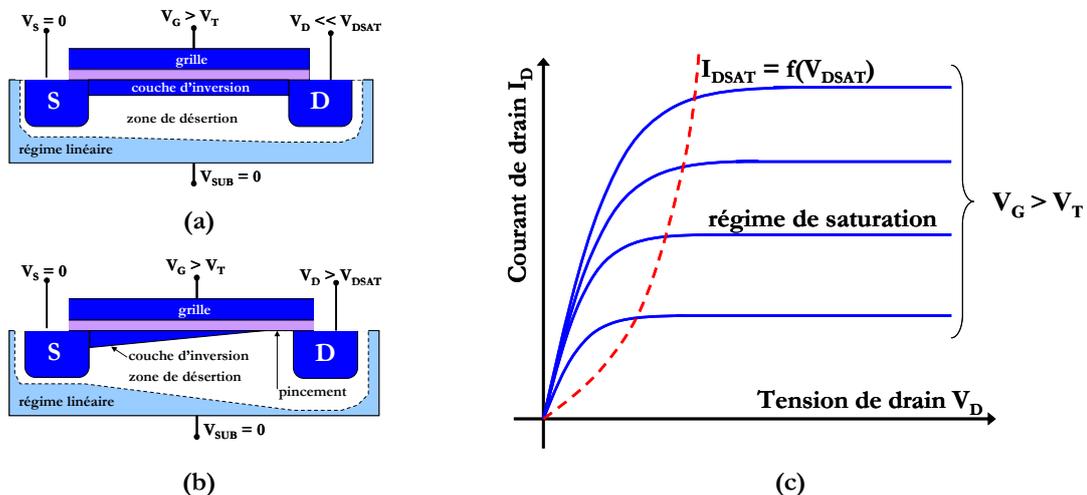


Fig. 9: Schéma de principe d'un MOSFET en régime linéaire (a) et en saturation (b). Réseau de caractéristiques $I_D(V_D)$ dans un MOSFET long (c).

1.4.3 Régime sous le seuil

Quand la tension de grille est inférieure à la tension de seuil et que la surface du semi-conducteur est en inversion faible, on parle de régime 'sous le seuil'. Sous le seuil, le courant de drain est dominé par le courant de diffusion et dépend exponentiellement de la tension de grille et de la tension de drain.

Par définition, la pente sous le seuil S (*stricto sensu*, il s'agit de l'inverse de la pente sous le seuil) est définie par la relation:

$$S = \frac{1}{\frac{\partial(\log_{10} I_D)}{\partial V_G}} = 2.3 \frac{kT}{q} \frac{\partial V_G}{\partial \phi_s} \quad \text{Eq. (7)}$$

avec ϕ_s le potentiel de surface du semiconducteur.

La pente sous le seuil pour un transistor long sur silicium massif peut se mettre sous la forme [Sze'81]:

$$S = \frac{kT}{q} \ln(10) \left(1 + \frac{C_{DEP}}{C_{OX}}\right) \quad \text{Eq. (8)}$$

avec $C_{DEP} = \frac{\epsilon_{Si}}{W}$ et $C_{OX} = \frac{\epsilon_{OX}}{t_{OX}}$ les capacités de désertion et d'oxyde, ϵ_{Si} et ϵ_{OX} les permittivités diélectriques du silicium et de l'oxyde de grille, W la profondeur de la zone de désertion et t_{OX} l'épaisseur de l'oxyde de grille.

On voit donc que l'on a tout intérêt à minimiser la capacité de la zone de désertion C_{DEP} par rapport à la capacité d'oxyde C_{OX} . On peut ainsi utiliser un oxyde le plus fin possible ou un

canal faiblement dopé. La limite théorique de la pente sous le seuil est fixée par le phénomène de diffusion et vaut $2.3 (kT/q) \approx 60 \text{ mV/décade}$ à 300 K.

Il est intéressant de noter que de nouveaux dispositifs où le courant à l'état passant est généré par le phénomène d'ionisation par impact (dispositifs IMOS) permettent de s'affranchir de cette limite théorique. Les porteurs dans ce dispositif ne sont plus soumis au phénomène de diffusion et des pentes sous le seuil de l'ordre de 5 mV/décade ont été obtenues [Gopalakrishnan'02][Mayer'06].

La Fig. 10 montre l'allure typique d'une caractéristique de transistor MOS avec les grandeurs que nous venons de définir.

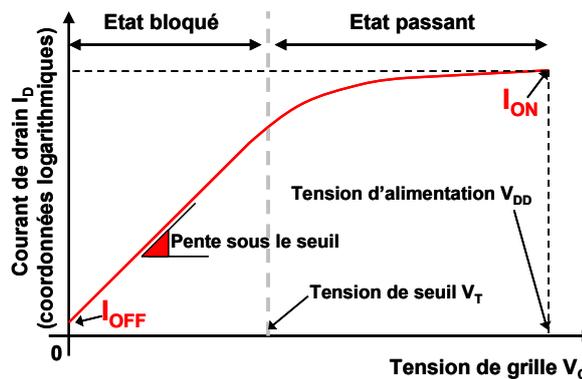


Fig. 10: Allure typique du courant de drain en fonction de la tension de grille $I_D(V_G)$ dans un transistor MOS.

1.4.4 Le MOS extrinsèque

Une structure MOS n'est jamais parfaite à cause des éléments parasites (résistifs, capacitifs et inductifs) apportés entre autres par la grille, les jonctions source/drain et les contacts. Pour créer l'effet de champ, on doit ainsi déposer une couche diélectrique (l'oxyde de grille) et une grille sur le semiconducteur. Pour créer le canal et y accéder, on rajoute de même les zones de diffusion de source et drain de chaque côté de celui-ci. Pour contacter les quatre terminaux de source, drain, grille et substrat on dépose des contacts métalliques et des interconnexions. Tous ces éléments rajoutés sont à l'origine d'effets parasites pour la plupart capacitifs et résistifs qui dégradent les performances intrinsèques des dispositifs :

- ✓ Résistance d'accès à la grille (due aux contacts, au type de grille utilisé).
- ✓ Résistances d'accès au drain et à la source (dues aux contacts, aux extensions et aux source/drain siliciurées ou non).
- ✓ Résistances d'accès au substrat.
- ✓ Capacités grille/drain et grille/source: capacités de bord (*fringe*) et de recouvrement (*overlap*).
- ✓ Capacités drain/substrat et source/substrat (capacités métalliques et de jonction).
- ✓ Capacité drain/source (capacité métallique et de proximité).
- ✓ Capacités grille/substrat.

On nomme alors MOS extrinsèque la structure composée de tous ces éléments parasites, par opposition au MOS intrinsèque qui joue le rôle d'une source idéale de courant [Haslett'69] (Fig. 11).

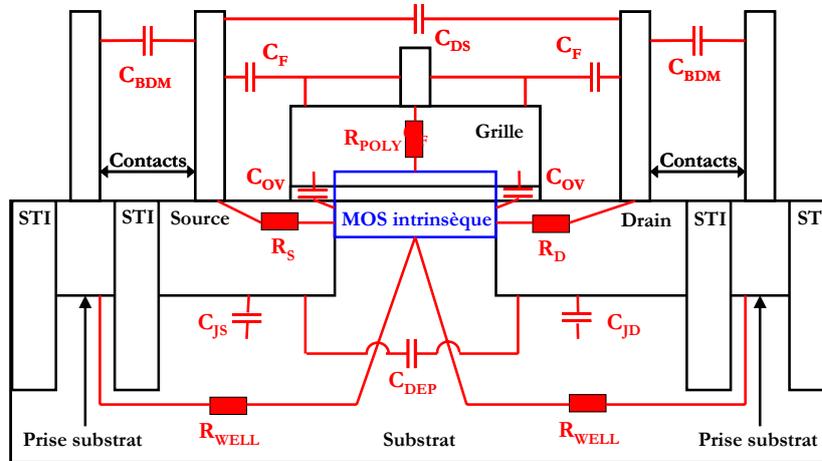


Fig. 11: Schéma d'une structure MOS extrinsèque.

2 Les effets de canaux courts

2.1 Les enjeux de la miniaturisation

Dans un circuit électronique, les deux paramètres essentiels à contrôler sont la consommation des circuits et leur vitesse de fonctionnement. La puissance statique dissipée P_S s'exprime sous la forme:

$$P_S = I_{OFF} V_{DD} \quad \text{Eq. (9)}$$

avec V_{DD} la tension d'alimentation et I_{OFF} le courant à V_{DD} nul.

Le temps de commutation τ du transistor (ou délai intrinsèque du transistor) est donné par :

$$\tau = \frac{C_{OX} V_{DD}}{I_{ON}} \quad \text{Eq. (10)}$$

avec C_{OX} la capacité d'oxyde, V_{DD} la tension d'alimentation et I_{ON} le courant de sortie à la tension d'alimentation.

Afin de minimiser la puissance statique dissipée, le courant I_{OFF} doit être le plus faible possible. Il est également nécessaire d'obtenir un courant I_{ON} maximal afin d'avoir un temps de commutation très faible et donc, une fréquence de commutation la plus élevée possible. I_{ON} peut être optimisé de plusieurs manières (voir Eq. (5)):

- ✓ Soit par une augmentation de C_{OX} , c'est-à-dire une diminution de l'épaisseur effective d'oxyde.
- ✓ Soit par une augmentation de la tension d'alimentation V_{DD} .
- ✓ Soit par une diminution de la longueur de grille L_G .

Cependant, le choix d'une des deux premières solutions implique forcément une augmentation du délai intrinsèque. La diminution de la longueur de grille (L_G) s'avère donc être le paramètre pour à la fois améliorer I_{ON} et τ .

Conjointement à la longueur de la grille, toutes les autres dimensions du transistor (comme la tension d'alimentation V_{DD} , l'épaisseur de l'oxyde de grille t_{OX} , la largeur du canal W , les profondeurs de jonctions X_j , le dopage du canal) doivent être modifiées: ce sont les règles d'échelle. Selon une règle proposée dans les années 1970 s'appuyant sur l'invariance du champ électrique suivant les générations technologiques, un coefficient de réduction d'échelle α peut être défini et piloter les dimensions des différentes grandeurs technologiques (Tab. 1, [Dennard'74]). Cette règle a donné de bons résultats jusqu'à $V_{DD} = 1$ V, où les courants de fuite deviennent trop importants. Une autre règle dite 'généralisée' a été proposée (Tab. 1, [Davari'95]) Elle définit deux paramètres d'échelle, un lié à la géométrie α , l'autre au champ électrique β .

$\alpha < 1 ; \beta > 1$	règle d'échelle à champ constant	règle d'échelle généralisée
L_G, W, t_{OX}, X_j	α	α
C_{OX}	α	α
Champ électrique	1	β
I_{ON}/W	1	β^2
V_{DD}	α	$\alpha\beta$
Dopage du film	$1/\alpha$	β/α
Temps de propagation intrinsèque τ	α	α/β

Tab. 1: Règles d'échelle à champ constant et généralisée [Dennard'74][Davari'95].

2.2 Les effets de canaux courts

Avec la réduction des dimensions, des effets parasites qui étaient négligeables pour de grandes longueurs de grille peuvent devenir importants. Pour les générations de transistors ayant une longueur de grille de 50 nm ou moins, la suppression du courant de fuite devient un défi technologique de plus en plus difficile à résoudre. Ce problème, qui porte le nom générique d'effets de canal court (*short channel effects*, SCE), limite à terme la réduction d'échelle des transistors MOSFETs conventionnels.

2.2.1 L'effet de canal court (*roll-off*)

Dans un canal court, l'étendue des zones désertées du côté de la source et du drain devient non négligeable devant la longueur du canal; par conséquent, on ne peut plus supposer le champ électrique uniquement transverse mais prendre en compte également sa composante longitudinale. Il s'en suit que la grille ne contrôle qu'une charge globalement moins importante dans le canal, d'où une chute de la tension de seuil.

Afin de minimiser cet effet de "partage de charge" (*charge sharing*), on doit rendre les lignes de potentiel les plus parallèles au canal possible en renforçant le dopage du canal ou en réduisant l'épaisseur de l'oxyde de grille. La solution classique pour juguler le *roll-off* consiste à

augmenter localement le dopage du substrat (poches ou *pockets*) afin d'empêcher les lignes de champ électrique venant de la région du drain de pénétrer dans le canal.

2.2.2 L'effet de canal court inverse (*roll-up*)

Ce phénomène est le résultat d'une non-uniformité du dopage de substrat ainsi que de la variation de l'épaisseur d'oxyde de grille. Ainsi, lorsque L_G diminue, le dopage moyen vu par la grille est plus important et la tension de seuil nécessaire pour amener la surface en inversion forte s'en trouve augmentée. Ce phénomène peut être minimisé par des recuits rapides avec des pentes très raides et un palier de quelques secondes à une température très élevée (typiquement 1000°C afin d'activer les dopants).

2.2.3 L'effet DIBL

L'effet DIBL (*Drain Induced Barrier Lowering*) est attribué à l'influence électrostatique du drain sur la hauteur de barrière de la jonction source-substrat ou barrière d'injection (Fig. 12). En augmentant la tension de drain, il y a extension de la zone de charge d'espace au niveau du drain. Cette zone de charge d'espace peut aller jusqu'à diminuer la hauteur de la barrière d'injection.

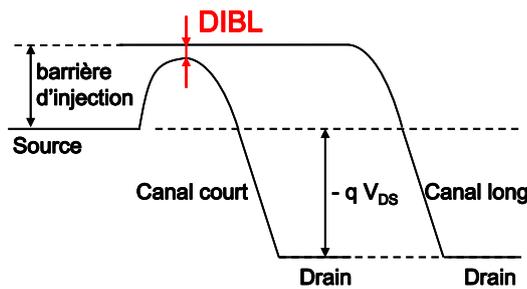


Fig. 12: Schéma énergétique le long du canal dans le cas d'un canal long et d'un canal court montrant l'effet DIBL.

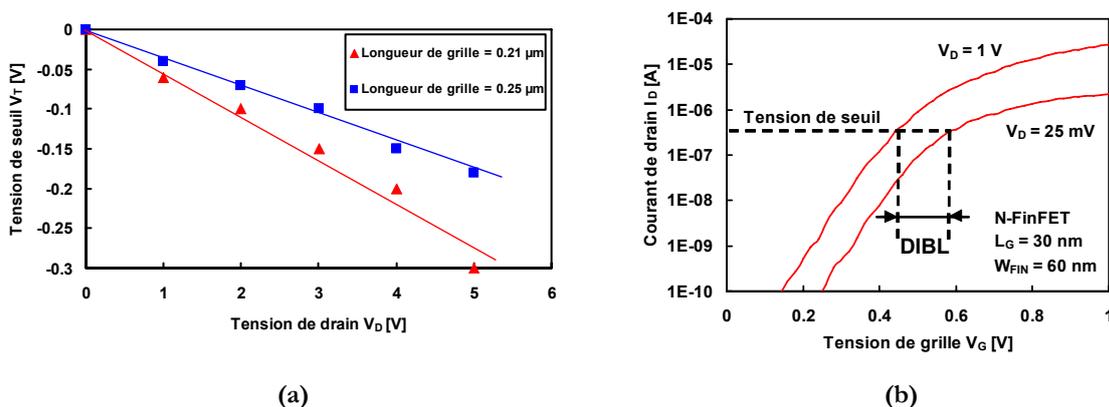


Fig. 13: Courbe $V_T(V_{DS})$ illustrant la chute de tension de seuil [Fjeldly'98] (a). $I_D(V_G)$ (échelle logarithmique) dans le cas d'un transistor court pour des tensions de drain en régime ohmique et en saturation (b)

La conséquence de cet effet est une diminution de la tension de seuil du transistor avec V_D d'autant plus importante que la distance source/drain (la longueur de canal) est faible (Fig.

13.a). On augmente par là même les courants sous le seuil. Habituellement, on mesure l'effet DIBL par le décalage de la courbe de transfert $I_D(V_G)$ entre une mesure en régime linéaire (V_D faible) et une mesure en saturation (V_D fort) (Fig. 13.b). L'unité de mesure du DIBL est donc le mV/V si on normalise par la différence des deux tensions de drain utilisées.

2.2.4 Le perçage (*punch-through*)

Nous avons vu que la pente sous le seuil est améliorée en utilisant de faibles dopages. Cependant, avec de faibles niveaux de dopages les zones de désertion issues du drain et de la source s'étendent dans le film jusqu'à pouvoir fusionner (Fig. 14). La barrière de potentiel à la source s'abaisse alors et des porteurs sont injectés de la source vers le drain. Ce passage se fait en profondeur (on parle également de '*sub-surface DIBL*'). Le phénomène de perçage dépend fortement de la tension de drain appliquée et de l'épaisseur des jonctions. Afin de limiter le perçage, on peut augmenter le dopage du substrat afin de diminuer l'extension de la zone de désertion du drain. Cependant, la pente sous le seuil est alors dégradée. Une solution consiste à implanter des zones dopées localisées là où se fait le perçage, c'est-à-dire à la base des jonctions (technologie Halo). On diminue ainsi le perçage sans dégrader la pente sous le seuil.

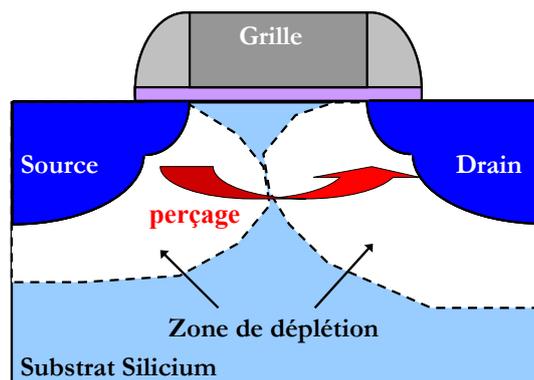


Fig. 14: Schéma du phénomène de perçage.

2.3 Les matériaux innovants

2.3.1 Les oxydes haute permittivité (*high-k*)

La réduction des dimensions des transistors nécessite d'utiliser des épaisseurs d'oxyde de grille de plus en plus fines. Ceci conduit malheureusement à des courants de fuite à travers les grilles plus importants. Un moyen de conserver une capacité équivalente sans trop diminuer l'épaisseur de l'oxyde de grille est alors d'utiliser un matériau à haute permittivité (*high-k*). De nombreux diélectriques de grille ont été étudiés comme *high-k* potentiels (AlO_3 , HfO_2 , ZrO_2 , etc.) ; l'oxyde d'hafnium HfO_2 semble être un bon candidat pour les diélectriques de permittivité moyenne ($\epsilon_r \approx 20$) de par son grand gap (5.65 eV) et sa relative stabilité thermodynamique sur le silicium.

2.3.2 Les grilles métalliques

Le développement des dispositifs est axé sur le couple silicium/silice. L'utilisation du silicium polycristallin a permis la mise au point des technologies auto-alignées: on dépose la grille avant implantation des sources et drains des transistors. L'implantation dans la grille permet d'obtenir des grilles Si n^+ ou p^+ et d'améliorer la conductivité du silicium polycristallin. Cependant, le silicium polycristallin même dégénéré reste un semiconducteur. Par conséquent, en condition d'inversion les porteurs dans la grille sont repoussés de l'interface grille-oxyde. Il se crée donc une capacité parasite dans la grille (phénomène de *polydepletion*), qui diminue la capacité totale de la structure MOS et la rapidité des dispositifs. L'avantage majeur d'utiliser un métal en tant que matériau de grille est donc l'élimination des phénomènes de désertion de grille. De plus, on supprime l'effet de diffusion du bore propre au polysilicium dopé. Les métaux sont cependant très sensibles aux recuits, et les étapes suivant le dépôt de grille peuvent décaler la tension de bandes plates, changer la densité de charges fixes dans l'oxyde et affecter le travail de sortie de la grille.

3 La technologie 'silicium-sur-isolant' (SOI)

On le voit, les effets de canaux courts deviennent très préoccupants aux faibles longueurs de grille. Les méthodes pour les limiter sont de plus en plus difficiles à mettre en œuvre et à contrôler. De nouvelles voies ont donc tout naturellement été explorées, comme les technologies SOI.

3.1 Introduction - principe de fonctionnement

L'idée à la base de l'élaboration de la technologie SOI (*Silicon On Insulator* – Silicium Sur Isolant) est assez simple. Dans une plaque classique, en regard de l'épaisseur totale de la plaque (entre 400 et 800 μm) la zone utilisée pour la fabrication et le fonctionnement des transistors est très faible (de 100 à 200 nm, soit moins de 1 % de l'épaisseur totale!). Toute cette épaisseur utilisée uniquement comme support mécanique conduit à des phénomènes parasites comme le *latch-up* (formation d'un thyristor parasite entre les zones actives via le substrat). D'où l'idée d'isoler la zone électriquement active du transistor du reste de la plaque avec un diélectrique. On obtient ainsi une isolation totale entre les dispositifs et les capacités des régions de source et drain sont réduites par rapport au silicium massif.

Historiquement, la technologie SOI a été développée dans les années 1960-1970 pour des applications 'niches' en raison de performances particulières [Celler'03]. On peut notamment citer les applications militaires et spatiales, pour les propriétés de résistance aux radiations ionisantes. Le premier matériau, le silicium sur saphir (SOS), a été suivi par toute une variété de structures SOI. Leur dénominateur commun est d'offrir, grâce à un oxyde enterré (le BOX : *Buried OXide*), une parfaite isolation diélectrique entre la couche active des circuits et le substrat de silicium massif (Fig. 15).

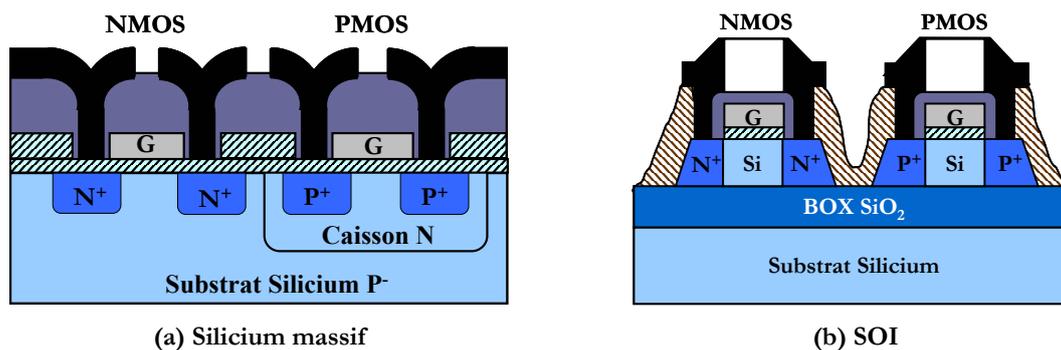


Fig. 15: Comparaison de l'architecture d'un MOSFET en technologie silicium massif (a) ou SOI (b).

Depuis le début des années 1990, la mise au point de nouveaux matériaux SOI ainsi que l'explosion des appareils électroniques portables a promu le SOI comme une technologie de choix pour la fabrication de composants à basse consommation et à haute fréquence. La technologie SOI est actuellement considérée comme une solution avantageuse pour remplacer la technologie silicium classique et atteindre des dimensions déca-nanométriques (Tab. 2). Nous assistons actuellement au développement rapide du SOI grâce notamment à l'apparition de nouvelles méthodes de fabrication qui ont permis d'accroître la qualité du matériau tout en diminuant les coûts.

	SOI	Silicium massif
Vitesse des circuits	Amélioration d'environ 30 %	
Opérations à faibles tensions	Pour une même vitesse, réduction d'environ 80 % de la puissance nécessaire	
Coût	En principe, procédé complet moins cher	Plaques beaucoup moins chères
Design	En principe, design plus simple	Bibliothèques de design existantes
Utilisation en température	Réduction des variations de V_T et des courants de fuite	L'oxyde enterré en SiO_2 du SOI a une très faible conductivité thermique: autoéchauffement
Tolérance aux radiations	Excellente grâce à l'isolation par l'oxyde enterré	

Tab. 2: Tableau comparatif des performances du SOI par rapport au silicium massif.

Les investissements nécessaires en microélectronique sont de plus en plus importants, et une technologie émergente est quasiment toujours beaucoup plus chère qu'une technologie mature. Les circuits SOI ont ainsi été longtemps handicapés par le coût des plaquettes et par le faible nombre de bibliothèques de design existantes (donc par le surcoût occasionné par leur création). De manière classique, les circuits SOI ont donc été utilisés quasi-exclusivement pour des applications niches où ils étaient indispensables tandis que les technologies sur silicium massif convenaient parfaitement à la réduction des dimensions. Cependant, les frontières technologiques du silicium se rapprochent de plus en plus et les effets parasites sur les transistors sur silicium massif deviennent très préoccupants. En même temps, les produits

SOI deviennent de moins en moins chers et leurs performances sont très bonnes. La technologie SOI deviendra peut-être dominante dans l'avenir. Aujourd'hui, les principaux fabricants de composants à travers le monde ont mis en place des programmes SOI, et même six d'entre eux à ce jour ont décidé de passer au SOI au moment de leur migration au diamètre de plaques 300 mm. En 2003, *AMD* a introduit le SOI dans certaines de ces puces très hautes performances (*Athlon 64* et *Opteron*). *AMD* est également le premier fabricant à avoir converti totalement toute sa fabrication de microprocesseurs au SOI.

3.2 Fabrication des plaques SOI

Il existe toute une variété de méthodes de fabrication des plaques SOI (BSOI, SIMOX, FIPOS, ZMR, ELO, etc. [Cristoloveanu'02]). Nous ne décrivons que la technologie *Smart Cut*, qui a servi pour les dispositifs utilisés dans le cadre de cette thèse.

Mise au point au LETI [Bruel'95], la technologie *Smart Cut* (90 % des plaques SOI vendues actuellement dans le monde) est à la base du développement de la société SOITEC. Elle met en jeu l'adhérence moléculaire (*wafers bonding*) et le transfert de couche mince par détachement suite à une implantation ionique (Fig. 16).

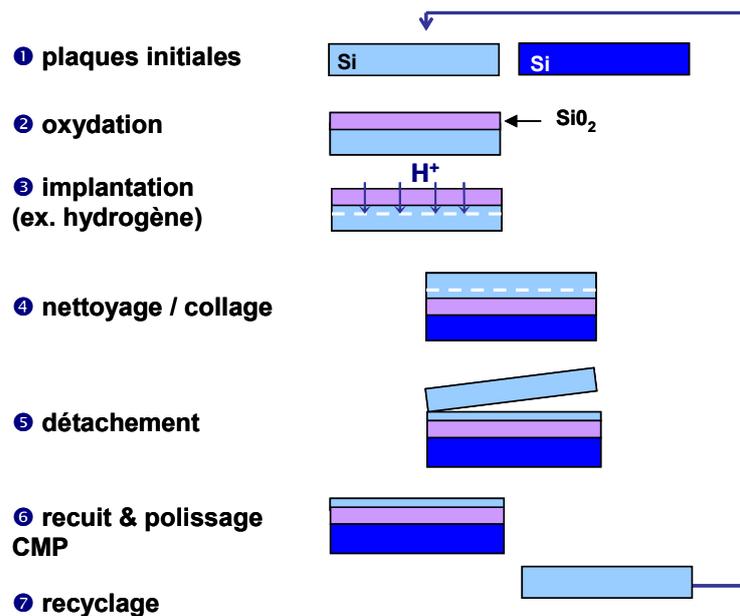


Fig. 16: Principales étapes de fabrication d'une plaque SOI *Smart-cut Unibond*.

L'adhérence moléculaire est la propriété grâce à laquelle deux surfaces planes, lisses et propres peuvent par simple mise en contact donner lieu à un "collage moléculaire", c'est-à-dire créer des liaisons entre elles de telle manière qu'il faut appliquer une force pour les séparer. Il s'agit à température ambiante de liaisons d'interaction à courte distance, de faible énergie (type Van der Waals). Lors d'un recuit, ces liaisons pourront être remplacées par des liaisons de forte énergie (type covalentes).

Le transfert par détachement se fait quant à lui moyennant un apport d'énergie (thermique, mécanique) qui provoque la formation et la croissance de cavités issus de défauts

spécifiques créés lors de l'implantation d'hydrogène. Ces cavités évoluent en microfissures qui aboutissent à la fracture le long de cette zone fragilisée. L'énergie et la dose d'implantation définissent l'épaisseur de la couche active (de 50nm à 1,5µm). Une épaisseur plus importante peut être réalisée par une étape ultérieure d'épitaxie de silicium.

3.3 Architectures partiellement et complètement désertées

3.3.1 Définition

En supposant le dopage homogène, on peut exprimer la profondeur W de la zone de désertion dans le film en résolvant l'équation de Poisson [Cristoloveanu'95]:

$$W = \sqrt{\frac{2\epsilon_{Si}\psi_s}{eN_A}} = \sqrt{\frac{4\epsilon_{Si}\psi_F}{eN_A}} \quad \text{Eq. (11)}$$

avec ψ_F le potentiel de Fermi et N_A le dopage du film.

La charge d'espace dans la zone de désertion est quant à elle de la forme:

$$Q_{dep} = -eN_A W = -\sqrt{2eN_A\epsilon_{Si}\psi_s} \quad \text{Eq. (12)}$$

Selon l'extension de la couche de désertion dans le film de silicium, on distingue deux familles de transistors SOI. On parle ainsi de transistors 'complètement désertés' (FDSOI pour *Fully Depleted SOI*) si la profondeur de la zone de désertion W de l'Eq. (11) est supérieure à l'épaisseur du film (Fig. 17.a). Dans le cas contraire, on parle de transistors 'partiellement désertés' (PD pour *Partially Depleted SOI*) (Fig. 17.a).

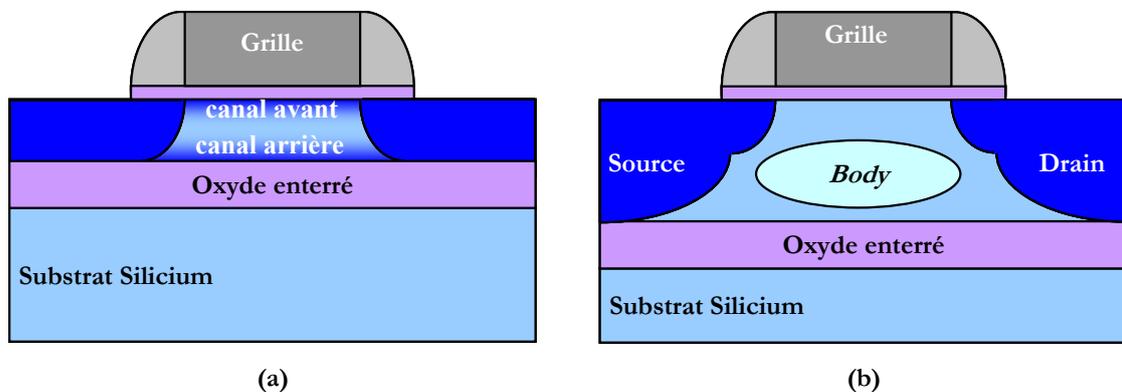


Fig. 17: Représentation schématique d'un transistor SOI 'complètement déserté' FDSOI (a) et 'partiellement déserté' PDSOI (b).

3.3.2 Transistors SOI 'Partiellement Désertés'

Dans un transistor partiellement déserté, une région neutre subsiste dans le film et par conséquent isole de tout effet de couplage interface avant/interface arrière. Quand la couche de silicium est connectée à la masse (via des contacts indépendants sur le film de silicium ou des connexions directes source/film), les composants SOI partiellement désertés se comportent de façon similaire à ceux fabriqués sur silicium massif. Si les contacts sur le film

ne sont pas disponibles, le potentiel du substrat interne n'est pas contrôlé. Des effets appelés de 'substrat flottant' se produisent alors, entraînant plusieurs conséquences:

- ✓ **Effet Kink.** L'effet Kink est l'un des principaux effets de substrat flottant qui sont déclenchés par l'accumulation de charges produites par ionisation par impact dans le film de silicium (Fig. 18.a). L'ionisation par impact dans la région de fort champ électrique latéral, près du drain, mène à la création de paires électrons-trous. Dans les transistors à canal N, les électrons vont vers le drain avec le courant de drain et les trous constituent le courant de substrat. Ce courant de porteurs majoritaires provoque, outre la polarisation de la diode source/substrat, une augmentation du potentiel du film de silicium induisant une réduction de la tension de seuil du transistor à l'origine de l'effet Kink. Il se traduit dans les transistors SOI partiellement désertés par une augmentation du courant de drain (Fig. 18.b) et par du bruit électrique en saturation.
- ✓ **Effet de bipolaire parasite.** A cause du courant d'ionisation (Fig. 18.a), un transistor bipolaire parasite Source-Film-Drain peut se déclencher.
- ✓ **Effets transitoires.** La charge en excès ou en déficit dans le silicium retrouve l'équilibre par effet de génération-recombinaison. Ceci conduit à des courants transitoires [Munteanu'99].
- ✓ **Effet d'hystérésis** suivant le sens de balayage de V_G [Wei'97].

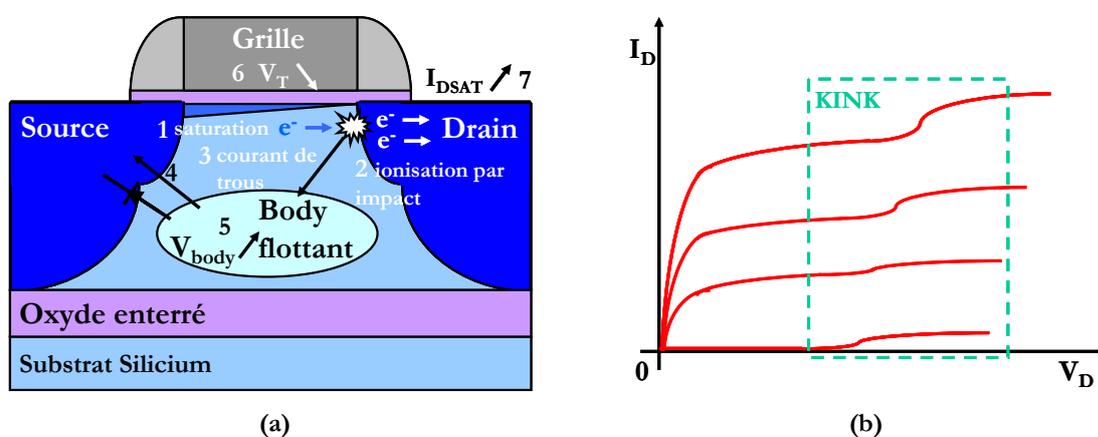


Fig. 18: Représentation schématique de l'effet Kink (a) et de l'augmentation du courant de drain résultante (b).

Dans l'état présent de la technologie SOI, les transistors partiellement désertés apparaissent comme plus séduisants que les transistors complètement désertés. Ils sont plus tolérants aux fluctuations d'épaisseur du film et leur dopage peut être augmenté (ajustement de la tension de seuil) ou profilé (maîtrise des effets de canal court). De façon générale, les transistors 'partiellement déserté' rappellent mieux le fonctionnement et la conception des MOSFETs sur silicium massif. Le principal défi les concernant est la réduction des effets de substrat flottant, réalisable par l'ingénierie des jonctions (siliciuration, utilisation de SiGe, formations de poches) ou l'utilisation de contacts sur le film qui compliquent la conception.

3.3.3 Transistors SOI 'Complètement Désertés'

Dans un transistor complètement déserté, le contrôle électrostatique de la grille avant sur le canal est bien meilleur que dans les technologies partiellement désertées ou sur silicium massif. On obtient donc par rapport aux autres technologies:

- ✓ Une amélioration de la mobilité du film puisque le champ effectif vertical est plus faible et le dopage non nécessaire.
- ✓ Pour de grandes longueurs de canal et des oxydes enterrés relativement fins, le couplage des deux canaux (Fig. 17.a) induit des caractéristiques quasi idéales. La pente sous le seuil est proche de 60 mV/décade à cause de la suppression de la capacité différentielle de désertion.
- ✓ Les effets de substrat flottant sont quasiment supprimés puisque la désertion est totale.
- ✓ L'épaisseur réduite du film permet de faibles profondeurs de jonctions et donc des capacités de jonctions minimales. Sur des dispositifs classiques, il faut pour réduire les effets de canaux courts augmenter le dopage. Malheureusement, utiliser de telles concentrations dégrade la mobilité, augmente les capacités de jonction et les fuites dans les jonctions. Vu la faible épaisseur de la zone active du silicium, la profondeur des jonctions de source et drain est limitée par l'oxyde enterré.

Cependant, de nouveaux phénomènes problématiques apparaissent également:

- ✓ L'uniformité de l'épaisseur du film de silicium est particulièrement importante pour les transistors MOS complètement désertés car elle gouverne les fluctuations de la tension de seuil. On doit alors utiliser des dispositifs non dopés dont l'avantage réside principalement dans la faible variation de la tension de seuil due aux fluctuations aléatoires de dopage dans le canal. Cependant, il faut maîtriser la valeur du travail de sortie de la grille afin d'ajuster la tension de seuil du dispositif. En modifiant le matériau de grille, on peut diminuer les valeurs des tensions de seuil et ainsi ramener les tensions d'alimentation vers des valeurs acceptables ($\pm 0.3 - 0.5$ V). Ceci conduit à l'introduction de matériaux '*mip-gap*' (dont le travail de sortie est égal à celui du silicium intrinsèque) comme les grilles métalliques.
- ✓ Augmentation drastique des résistances d'accès au canal avec la réduction de l'épaisseur de film pouvant fortement dégrader le courant de sortie. Les zones de sources et drain surélevées permettent de juguler ce problème.
- ✓ Pour de petites longueurs de grille les caractéristiques peuvent être dégradées à cause de la pénétration dans le BOX du champ électrique issu du drain (voir l'effet *DIVSB* dans le chap. 4).
- ✓ L'autoéchauffement a toutes les raisons d'être plus fort dans ces structures que dans des structures sur silicium massif à cause de la faible conductivité thermique de l'oxyde enterré en SiO_2 .

4 Conclusions

Dans ce premier chapitre, les concepts de base et les grandeurs caractéristiques de la microélectronique ont été présentés. Les effets apparaissant aux faibles longueurs de grille ont été introduits. L'accent a été en particulier mis sur la technologie SOI, point de départ de nombreux dispositifs novateurs et pressentis pour les prochaines générations de transistors.

Les architectures SOI complètement désertées sont à la base des dispositifs à grilles multiples, dont les FinFETs font partie. Le prochain chapitre est consacré aux différentes architectures multigrilles existantes.

Références du chapitre 1

- [Ashcroft'76] N. W. Ashcroft, D. M. Mermin, "Solid State Physics", *Ed. Brooks Cole*, ISBN 2868835775, 1976.
- [Bruel'95] M. Bruel, "Silicon on insulator material technology", *Electronics Letters*, vol. 31, no. 14, 6th july 1995, pp. 1201-1202.
- [Celler'03] G.K. Celler, S. Cristoloveanu, "Frontiers of silicon-on-insulator", *Journal of Applied Physics*, vol. 93 no. 9, pp. 4955-4978, 2003.
- [Cristoloveanu'95] S. Cristoloveanu and S.S. Li, "electrical characterization of silicon-on-insulator materials and devices", *Kluwer Academic Publishers*, 1995.
- [Cristoloveanu'02] S. Cristoloveanu, F. Balestra, "Technologie silicium sur isolant (SOI)", *Techniques de l'ingénieur*, ref. E 2 380, 2002.
- [Davari'95] B. Davari, R.H. Dennard, G.G. Shahidi, "CMOS Scaling for High Performance and Low Power – The Next Ten Years", *proceedings of the IEEE*, vol. 83, no. 4, pp. 595-606, 1995.
- [Dennard'74] R. H. Dennard, F.H. Gaensslen, H.-N. Yu, V.L. Rideout, E. Bassous, A.R. Leblanc, "Design of Ion-Implanted MOSFETs with very small Physical Dimensions", *IEEE of Solid State Circuits*, vol. 9, no. 5, pp. 256-268, 1974.
- [Fjeldly'98] T. A. Fjeldly, T. Ytterdal and M. S. Shur, "Introduction to device modeling and circuits simulation", *Ed. Wiley*, New York, 1998.
- [Gopalakrishnan'02] K. Gopalakrishnan, P. B. Griffin and J. Plummer, "I-MOS: A Novel Semiconductor Device with a Subthreshold Slope lower than kT/q ", *IEDM'02 Technical Digest*, pp. 289-292, 2002.
- [Haslett'69] J.W Haslett and F.N Trofimenkoff, "Small-signal, high frequency equivalent circuit for the metal-oxide-semiconductor field-effect transistor", *proc. IEEE*, vol. 116, no. 5, May 1969.
- [Kittel'96] C. Kittel, "Introduction to Solid State Physics", 7th edition, *John Wiley and Sons (WIE)*, ISBN 0471111813, 1996.

- [Mathieu'96] H. Mathieu, “Physique des semiconducteurs et des composants électroniques”, 3^{ème} édition, *Ed. Masson*, ISBN 2-225-85124-7, 1996.
- [Mayer'06] F. Mayer, C. Le Royer, G. Le Carval, L. Clavelier and S. Deleonibus, “Static and Dynamic TCAD Analysis of IMOS Performance: From the Single Device to the Circuit”, *Electron Devices, IEEE Transactions on*, vol. 53 no. 8, pp. 1852-1857, 2006.
- [Munteanu'99] D. Munteanu, "Modélisation et caractérisation des transistors SOI: du pseudo-MOSFET au MOSFET submicronique ultra-mince", *thèse de doctorat*, INPG, 1999.
- [Sze'81] S. M. Sze, “Physics of Semiconductor Devices (2nd ed.)”, *John Wiley and Sons (WIE)*. ISBN 0-47-105661-8, 1981.
- [Wei'97] A. Wei, D. A. Antoniadis, "Design methodology for minimizing hysteretic V_T variations in partially depleted SOI CMOS", *IEDM'97 Technical Digest*, pp. 411-414, 1997.

Chapitre 2

Les composants multigrilles

Nous avons vu dans le chapitre précédent que la réduction des dimensions induisait un certain nombre de complications que les architectures sur silicium massif avaient de plus en plus de mal à contrôler. En plus de l'utilisation des technologies SOI, ceci a conduit à repenser la géométrie des transistors. L'adjonction d'une ou de plusieurs autres grilles dans le transistor permet d'améliorer le contrôle électrostatique du transistor, donc sa capacité à être miniaturisé. C'est ainsi qu'un certain nombre d'architectures sont proposées pour être intégrées à moyen terme dans les circuits. Dans ce chapitre, ces architectures multigrilles sont passées en revue. L'accent est particulièrement mis sur les transistors FinFETs, qui font l'objet de cette thèse.

Sommaire du chapitre 2

1	Introduction	45
2	Dispositifs double-grille planaires.....	46
3	Dispositifs multigrilles non planaires	46
	3.1 Transistors FinFET.....	46
	3.2 Transistors Triple-grille	48
	3.3 Transistors Π -FET et Ω -FET.....	49
	3.4 <i>Body-Tied</i> FinFETs	50
	3.5 Mémoires FinFlash.....	51
	3.6 Optimisation de la densité d'intégration dans les transistors verticaux	51
4	Dispositifs à grille enrobante - GAAFET	54
5	Technologie SON (Silicon-On-Nothing)	54
6	Dispositifs multicanaux et nanofils.....	55
	6.1 Dispositifs à multicanaux	55
	6.1.1 McFET.....	55
	6.1.2 MBCFET	56
	6.2 Dispositifs à nanofils.....	57
7	Autres dispositifs multigrilles.....	58
	7.1 Structure en 'T inversé' - ITFET.....	58
	7.2 Dispositifs à grilles indépendantes.....	59
	7.2.1 Les MIGFET	59
	7.2.2 Le G ⁴ -FET.....	59
8	Conclusions	60

1 Introduction

La réduction d'échelle dans les transistors CMOS en dessous du nœud 50 nm est difficile. Le contrôle des courants de fuite nécessite des oxydes si fins et des canaux tellement dopés qu'un procédé pourrait ne pas être pas trouvé. Outre les transistors en SOI partiellement et complètement désertés présentés dans le chapitre précédent, de nouvelles architectures dites 'multigrilles' sont proposées pour les prochains nœuds technologiques. En 2006, les records mondiaux en terme de miniaturisation ($L_G = 5$ nm pour NMOS [Lee'06] et PMOS [Yang'04]) sont détenus par des transistors à grilles multiples. Les structures à grilles multiples ont été introduites dans les années 80 (Fig. 19) et combinent un certain nombre d'avantages. Le couplage électrostatique entre les grilles permet ainsi de mieux contrôler le canal et donc de réduire les effets de canaux courts et les fuites dans le transistor à l'état bloqué. En diminuant le champ électrique transverse, la mobilité des porteurs est également améliorée. De plus, le courant de sortie sera augmenté si la surface des canaux de conduction est supérieure à celle d'un transistor planaire. Au niveau du procédé, une faible variation par rapport aux étapes d'une filière CMOS classique est recherchée afin de minimiser les coûts technologiques liés à ce changement d'architecture du transistor.

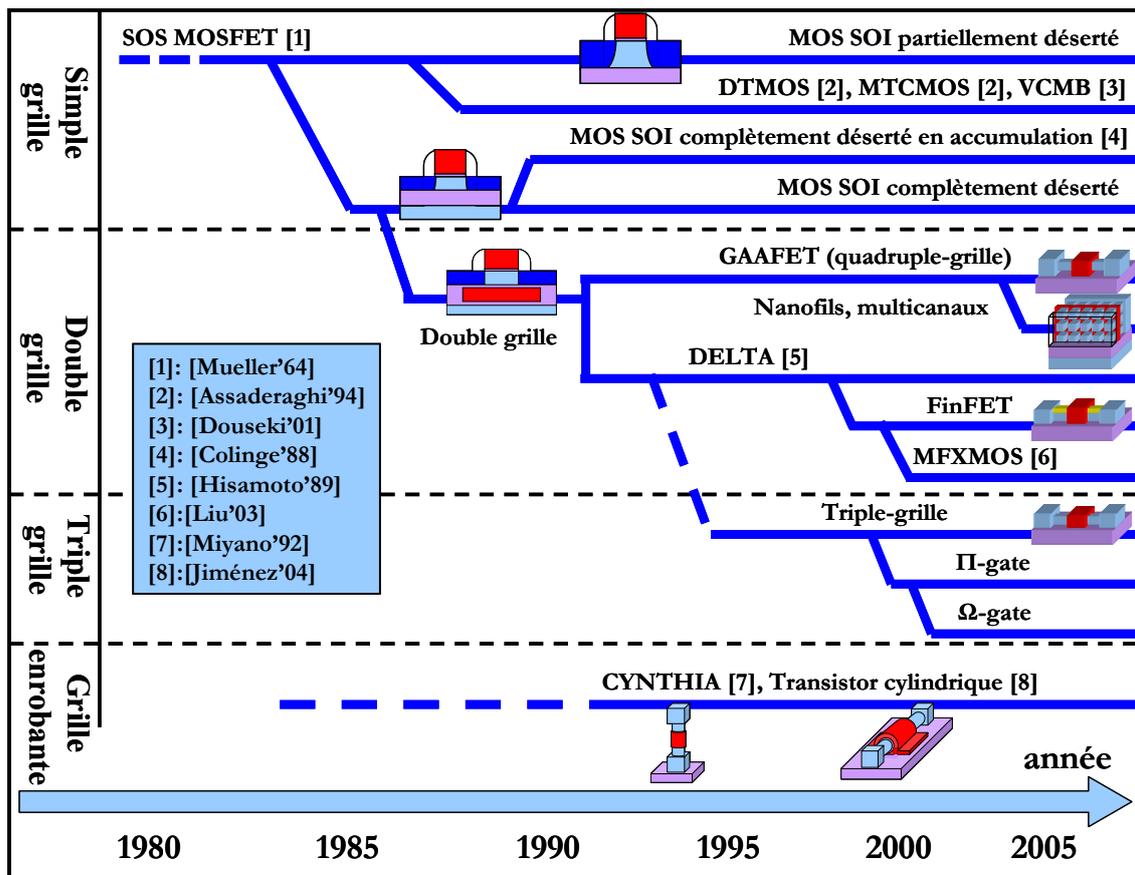


Fig. 19: Evolution historique des transistors à grilles multiples. Les dispositifs dont nous ne parlerons pas dans ce manuscrit sont référencés (adapté d'après [Colinge'04]).

2 Dispositifs double-grille planaires

De manière simple, un transistor double-grille est un transistor planaire où une deuxième grille est ajoutée sous le film de silicium (Fig. 20.a). Le principal avantage d'un composant double-grille par rapport à un dispositif simple-grille planaire vient de la capacité additionnelle apportée par la grille arrière. Les grilles contrôlent le potentiel dans le canal efficacement et la distribution de potentiel est relativement plate entre les grilles. C'est un candidat extrêmement prometteur pour les composants déca-nanométriques à cause de ses excellentes performances électriques : faibles effets de canaux courts, pente sous le seuil quasi parfaite et très bonne transconductance [Widiez'05].

A cause de son caractère planaire et des deux canaux de conduction, les composants double-grille permettent de cumuler excellent contrôle électrostatique et forte densité d'intégration. Cependant, les problèmes d'alignement des grilles (Fig. 20.b) et de connexion de la grille arrière entravent le développement de ces dispositifs. En effet, il a été montré que tout désalignement des grilles induisait une dégradation des paramètres électriques comme la pente sous le seuil ou le courant de sortie [Widiez'04][Widiez'05-2].

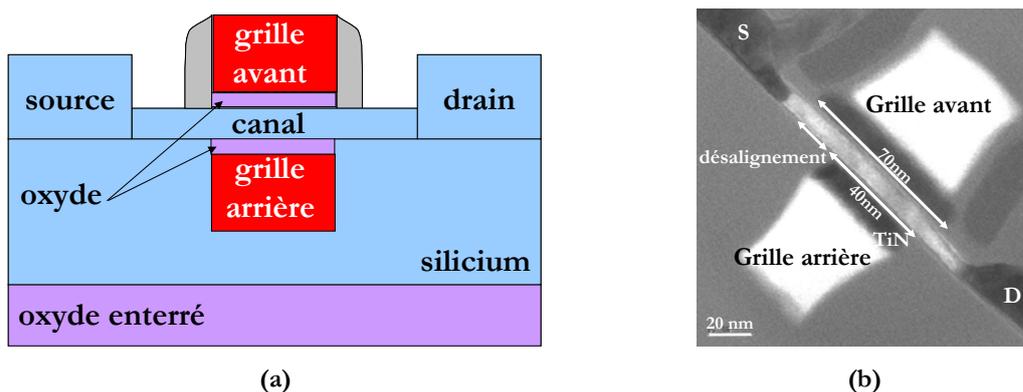


Fig. 20: (a): Schéma de principe d'un transistor double-grille. (b): Coupe TEM d'un transistor de 50 nm de longueur de grille avec un film de 10 nm d'épaisseur et un empilement de grille TiN/PolySi (d'après [Vinet'05]).

Il faut donc dans la mesure du possible disposer d'un procédé auto-aligné. A l'heure actuelle, des dispositifs double grille de faibles longueurs (jusqu'à $L_G = 10$ nm) ont été fabriqués [Vinet'05] mais ne sont pas auto-alignés. Des procédés auto-alignés [Guarini'01][Zhang'03] existent mais ne permettent pas pour l'instant de descendre à des longueurs de grille agressives. L'idée de disposer les deux grilles verticalement et non plus horizontalement conduit au concept de FinFET.

3 Dispositifs multigrilles non planaires

3.1 Transistors FinFET

Dans un FinFET, la zone active est un barreau de silicium (Fig. 21.a). L'oxyde de grille et la grille sont ensuite réalisés perpendiculairement à l'axe source/drain (Fig. 21.b). Avec la présence du masque dur nitruré au sommet du film, les canaux de conduction sont verticaux

et la direction de transport est horizontale. Dans ce manuscrit, nous nommerons toujours W_{FIN} la largeur de l'aileron et H_{FIN} sa hauteur. Initialement nommé DELTA (pour *fully DEpleted Lean-channel TrAnsistor*) [Hisamoto'89], ce transistor a pris le nom de FinFET (pour '*fin*', c'est-à-dire aileron en anglais) à la fin des années 90 [Hisamoto'98][Hisamoto'00]. Le procédé utilisé pour les transistors réalisés dans le cadre de cette thèse sera plus amplement évoqué dans le prochain chapitre.

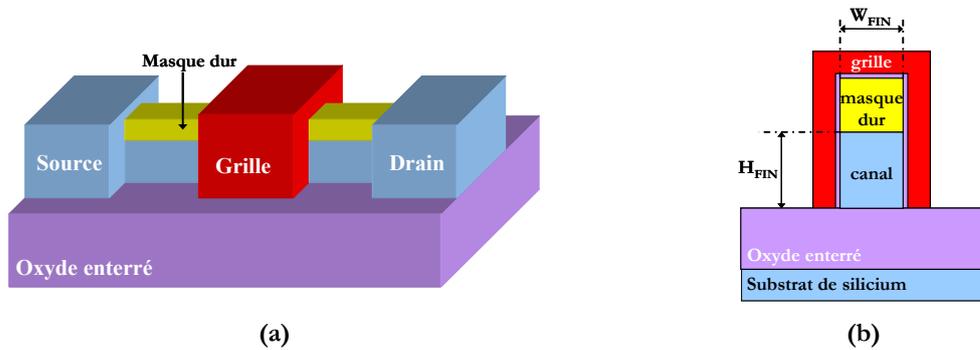


Fig. 21: Schéma de principe global (a) et en coupe transversale (b) d'un transistor FinFET.

Le transistor FinFET offre les avantages propres aux dispositifs double-grille avec un procédé relativement proche de celui d'un CMOS classique.

Pour un dispositif planaire complètement déserté, l'épaisseur du film doit être environ quatre fois plus petite que la longueur de grille pour conserver de bonnes performances électriques [Colinge'04]. Pour un transistor FinFET, cette contrainte dimensionnelle est relâchée puisque la largeur de l'aileron doit être d'environ les deux tiers de la longueur de grille [Colinge'04]. Concernant la hauteur de l'aileron, la limitation est due au procédé de fabrication puisque des ailerons fins et hauts (i.e. avec un fort 'rapport d'aspect' H_{FIN}/W_{FIN}) optimisent la densité d'intégration (le courant de sortie par surface occupée) sur la plaque. La Fig. 22 présente deux exemples d'ailerons obtenus dans la bibliographie.

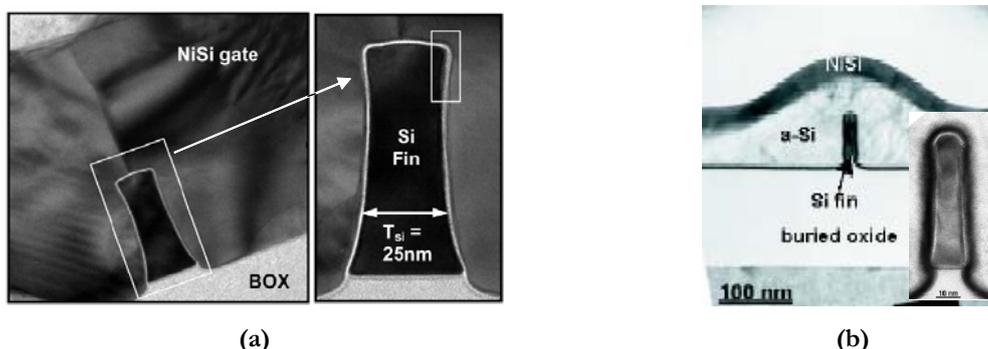


Fig. 22: Deux exemples d'ailerons obtenus pour des transistors multigrilles de type FinFET (d'après [Kedzierski'04] (a) et [Collaert'05] pour (b)).

Par rapport à un dispositif double-grille planaire, l'avantage d'un transistor FinFET est l'auto-alignement intrinsèque des deux grilles.

Cependant, la structure tridimensionnelle d'un FinFET soulève un certain nombre de questions. Pouvoir contrôler des canaux si fins avec des rapports d'aspect les plus élevés possibles n'est pas facile. La gravure classique RIE (*Reactive Ion Etching*, gravure ionique réactive) de l'aileron de silicium peut se révéler problématique. Il existe néanmoins une solution possible avec la gravure humide chimique utilisant le TMAH (tetramethylammonium hydroxide); ce composé grave très lentement les plans (111) du silicium par rapport aux autres plans cristallins. En partant d'un substrat (110) (voir chap. 5 Fig. 125), on peut ainsi obtenir des plans verticaux (111) spectaculaires [Liu'03][Liu'03-2].

La faible surface de canal disponible pour un FinFET rend également indispensable l'utilisation de réseaux de transistors. Les contraintes sur l'espacement entre chaque aileron seront abordées dans la section 3.6.

On retrouve la même exigence dans le contrôle de l'épaisseur de silicium que pour les transistors complètement désertés et les transistors double-grille. Il faut donc bien contrôler la largeur de l'aileron, qui pilote la tension de seuil. La réduction des effets de canaux courts avec la miniaturisation impose classiquement une augmentation du dopage du canal. Cependant, plus le canal du transistor est dopé et plus la variation de la tension de seuil en fonction de la largeur de l'aileron sera forte. Il faudrait donc contrôler parfaitement la largeur de l'aileron. Il faudrait aussi pouvoir doper le canal de manière très uniforme pour éviter les fluctuations de dopage qui changent les paramètres électriques, ce qui pose des problèmes d'implantation insolubles sur des transistors verticaux. C'est un argument de poids pour l'utilisation d'ailerons non dopés, mais qui nécessite l'emploi d'une grille métallique afin d'ajuster les tensions de seuil à des valeurs raisonnables. La question de la modulation du travail de sortie de la grille pour optimiser les CMOS [Wakabayashi'01][Kuriyama'06] reste cependant encore en débat, et l'aspect tridimensionnel du transistor complique singulièrement le problème.

Les étapes d'implantation des zones de source et drain sont également problématiques à cause de la géométrie tridimensionnelle du transistor. Pour réduire les résistances parasites, des source et drain surélevés sont aussi nécessaires.

L'utilisation d'un oxyde à haute permittivité se justifie quant à lui de la même manière que pour les structures planaires. Ces choix technologiques de grille métallique sur *high-k* sont ceux qui ont été fait pour les transistors réalisés dans notre étude.

3.2 Transistors Triple-grille

Dans un dispositif FinFET, le canal supérieur ne conduit pas à cause de l'épaisseur du masque dur (Fig. 21.b). Un transistor Triple-grille est un transistor pour lequel le masque dur a été gravé avant le dépôt de l'empilement de grille (Fig. 23). C'est donc un transistor avec deux canaux de conduction verticaux et un canal horizontal. Ceci soulève néanmoins la question du comportement électrostatique des zones de coins, qui sera étudiée dans le chapitre 4.

L'amélioration du couplage de la grille par rapport à un FinFET permet d'alléger les contraintes technologiques (les règles d'échelle, voir chap. 4). Ainsi, un transistor Triple-grille peut se contenter d'avoir une largeur et une hauteur de l'ordre de grandeur de la longueur de grille, ce qui est beaucoup moins contraignant du point de vue du procédé.

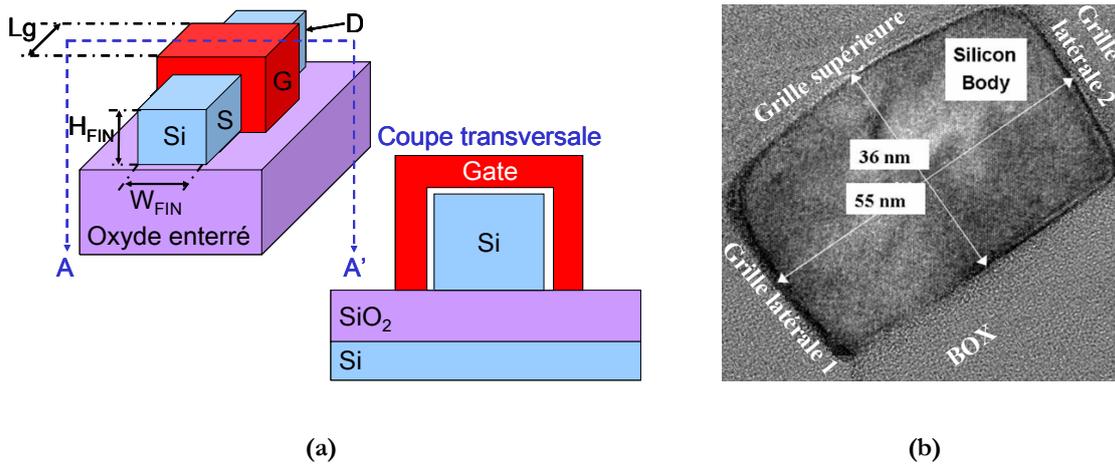


Fig. 23: Schéma de principe d'un transistor Triple-grille (a) et photo TEM de la coupe transversale d'un transistor Triple-grille (b) (d'après [Doyle'03]).

3.3 Transistors Π -FET et Ω -FET

En surgravant l'oxyde enterré lors de l'étape de définition des zones actives, on réalise des extensions des grilles latérales d'un transistor Triple-grille. On parle ainsi de transistor Π FET (ou Π FET) [Park'01] si les grilles latérales s'étendent verticalement dans l'oxyde enterré et de Ω FET [Yang'02] si les grilles s'étendent également sous le canal (Fig. 24.a). Ces extensions sont très faciles à réaliser et permettent d'améliorer considérablement le contrôle électrostatique de la grille (Fig. 25). La géométrie réalisée dans le cadre de cette thèse est de type Ω FET (Fig. 24.b). Nous verrons dans les prochains chapitres que le contrôle des effets de canaux courts et l'isolation des effets dus à l'oxyde enterré sont excellents avec ce type d'architectures.

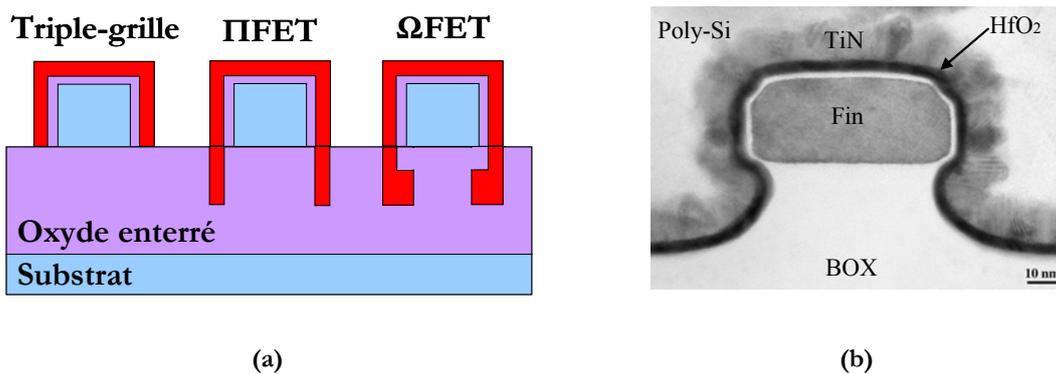


Fig. 24: Schéma de principe des transistors Π FET et Ω FET (a) et photo TEM de la coupe transversale d'un transistor Ω FET du LETI (b).

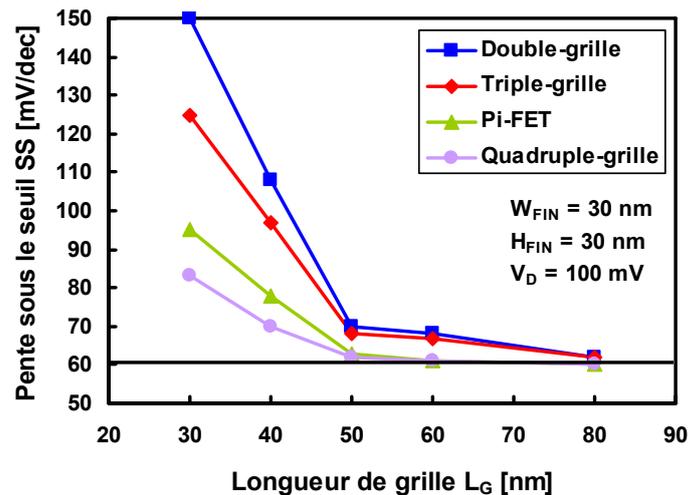


Fig. 25: Comparaison de la pente sous le seuil simulée en fonction de la longueur de grille pour des transistors double-grille, triple-grille, PiFET et quadruple-grille (d'après [Park'01-2]).

3.4 *Body-Tied* FinFETs

Initialement nommé Ω FET puis *Body-tied* FinFET (le nom Ω FET est désormais adopté pour les FinFETs SOI surgravés [Yang'02]), cette structure n'est rien d'autre qu'un FinFET réalisé sur silicium massif (Fig. 26).

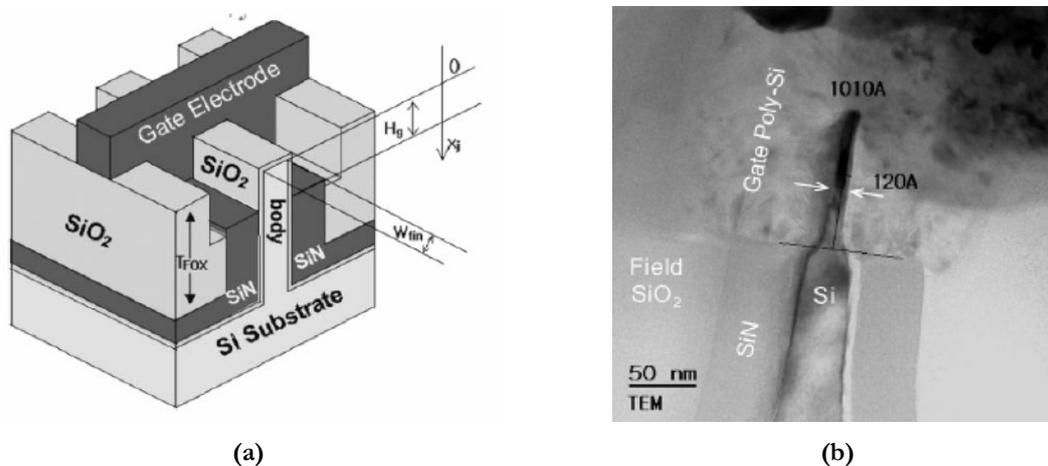


Fig. 26: Schéma (a) [Park'06] et coupe TEM transversale (b) [Kim'06-2] d'un transistor FinFET *body-tied*.

Dans le cas d'un FinFET réalisé sur structure en silicium massif, on garde bien sûr une très bonne compatibilité par rapport aux procédés CMOS planaires. On élimine les problèmes liés au SOI tels le coût des plaques, les défauts aux interfaces oxyde-silicium, l'effet de *body* flottant (si les dispositifs sont partiellement désertés). L'auto-échauffement potentiel du canal, dû à la très mauvaise conductivité thermique de l'oxyde, se trouve également naturellement éliminé. Le rôle du contact aileron/substrat est discutable. Par ailleurs, les transistors FinFETs sur SOI sont beaucoup plus simples à réaliser puisque l'oxyde enterré sert de couche d'arrêt lors de la gravure de la zone active.

3.5 Mémoires FinFlash

Une architecture FinFlash est l'adaptation du transistor FinFET (sur SOI ou silicium massif) aux mémoires *flash*. Les mémoires *flash* sont des mémoires non volatiles avec un faible temps d'effacement dont le stockage de l'information est assuré par une grille flottante. Elles sont très utilisées pour toutes les applications embarquées. Cependant, la réduction de la dimension des oxydes tunnel et de contrôle (Fig. 27) est limitée par la nécessité de conserver les données stockées. Des épaisseurs trop épaisses conduisent par contre à un faible couplage électrostatique induisant de fortes tensions d'alimentation. De plus, la dégradation de la pente sous le seuil et le perçage lors de la réduction de la longueur de grille de la mémoire diminuent considérablement la marge d'utilisation (la plage des tensions de seuil exploitables).

L'amélioration du couplage électrostatique et de la résistance aux effets de canaux courts dans les structures FinFET permet donc aux mémoires FinFlash d'avoir de meilleures perspectives de miniaturisation que les mémoires planaires [Jacob'05]. De plus, les dispositifs FinFlash offrent de plus forts courants de sortie, améliorant ainsi les temps d'accès.

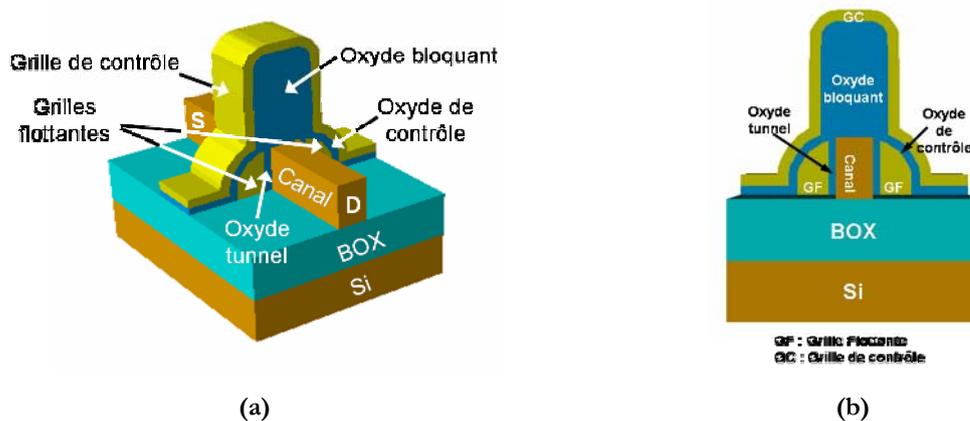


Fig. 27: Cellule mémoire FinFlash complète (a) et coupe transversale (b). Le stockage de l'information est assuré par les grilles flottantes (d'après [Jacob'05]).

3.6 Optimisation de la densité d'intégration dans les transistors verticaux

Pour avoir des courants de sortie acceptables, les dispositifs multigrilles doivent être intégrés en réseaux (Fig. 28.a et b). L'espacement (*pitch*) entre deux structures multigrilles verticales est un paramètre critique dans la mesure où il détermine la densité d'intégration, c'est-à-dire le courant de sortie d'une structure par unité de surface occupée sur une plaque (Fig. 28.c).

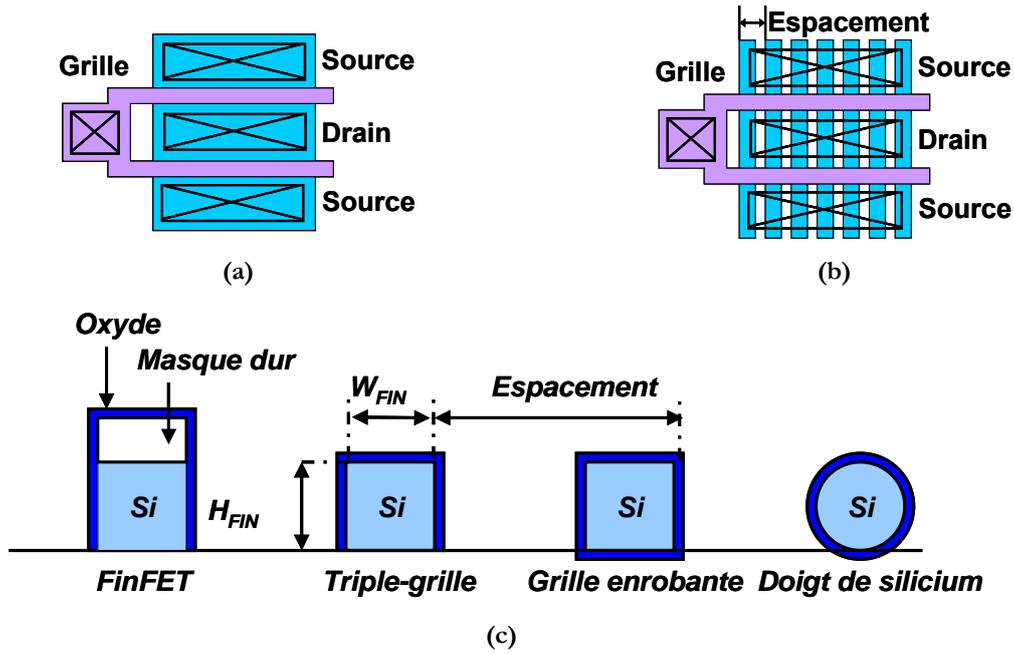


Fig. 28: Schémas d'un transistor planaire (a) et à grilles verticales (b) tels que dessinés sur le masque. Grandeurs géométriques caractéristiques pour les dispositifs FinFET, Triple-grille, à grille enrobante et en doigt (c).

	FinFET	Triple-grille	Grille enrobante	Doigt de silicium
Espacement maximal toléré	$2 H_{FIN}$	$W_{FIN} + 2 H_{FIN}$	$2 (W_{FIN} + H_{FIN})$	πW_{FIN}

Tab. 3: Valeur des espacements maximaux tolérés pour avoir une densité d'intégration supérieure à un transistor planaire de même surface.

Pour différents types de transistors multigrilles, on peut ainsi définir géométriquement l'espacement maximal toléré afin d'avoir un courant de sortie au moins égal à celui d'une structure planaire de même surface (Tab. 3). Sur la Fig. 29, on représente le rapport entre la surface de canal de conduction pour un dispositif multigrille et un dispositif planaire de même surface sur la plaque en fonction de l'espacement pour une épaisseur de film H_{FIN} donnée et pour différents rapports d'aspect (rapport H_{FIN} / W_{FIN}). Les courants de sortie sont bien sûr proportionnels au nombre de canaux de conduction. Ainsi à espacement égal un transistor à grille enrobante a un plus fort courant de sortie qu'un Triple-grille, lui-même étant supérieur à un FinFET. Le cas du transistor cylindrique est un peu à part, dans la mesure il perd son avantage en terme de courant de sortie par rapport aux autres structures quand on réduit les dimensions.

Dans le cas des FinFETs, l'espacement maximal est seulement proportionnel à la hauteur des dispositifs: l'espacement maximal reste le même quelque soit la largeur du transistor. Les règles d'échelle conseillent d'avoir un rapport d'aspect proche de 1 (Fig. 29.b) dans le cas d'un Triple-grille, donc un espacement faible si on veut avoir une faible largeur de canal W_{FIN} . Les transistors de type Π - ou Ω - FETs se situent quant à eux entre la courbe du Triple-grille et celle du transistor à grille enrobante. En réduisant le rapport d'aspect, on

augmente les performances en courant de sortie mais les espacements minimaux à respecter deviennent alors beaucoup plus contraignants.

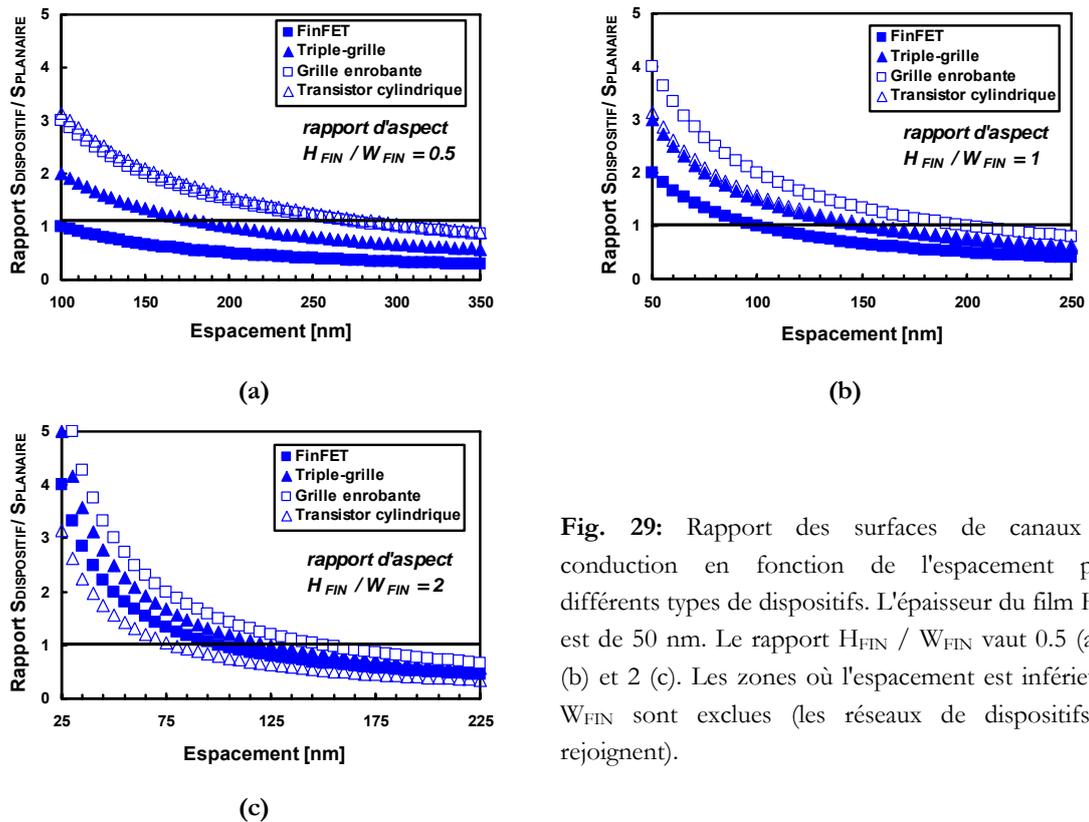


Fig. 29: Rapport des surfaces de canaux de conduction en fonction de l'espacement pour différents types de dispositifs. L'épaisseur du film H_{FIN} est de 50 nm. Le rapport $H_{\text{FIN}} / W_{\text{FIN}}$ vaut 0.5 (a), 1 (b) et 2 (c). Les zones où l'espacement est inférieur à W_{FIN} sont exclues (les réseaux de dispositifs se rejoignent).

Pour conserver un bon courant de sortie, il faut donc respecter un espacement minimal assez strict pour ne pas perdre en performance par rapport à un dispositif simple-grille ou double-grille planaire. Les cotes lithographiques doivent donc être améliorées. La lithographie dite 'par espaceurs' [Choi'02] permet de réduire l'espacement entre chaque aileron autrement qu'en réduisant la longueur d'onde utilisée lors de l'insolation de la résine (Fig. 30.a). Cette technique s'appuie sur la sélectivité de gravure du SiGe par rapport au silicium. On fabrique tout d'abord un réseau d'ailerons en SiGe en utilisant une lithographie classique (Fig. 30.b). Ensuite, on fait croître des espaceurs en oxyde. Le SiGe est ensuite gravé jusqu'au silicium. En utilisant l'oxyde restant comme masque dur, on obtient ainsi un réseau plus dense et plus étroit.

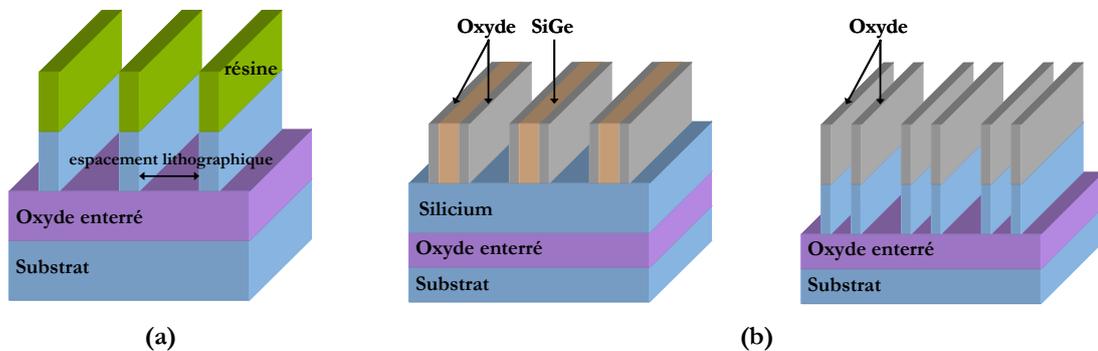


Fig. 30: Schéma de la lithographie classique (a) et par espaceurs (b).

4 Dispositifs à grille enrobante - GAAFET

Une structure quasi-circulaire ou carrée où la grille enrobe le canal permet d'avoir le meilleur couplage électrostatique, donc en théorie les meilleures perspectives de réduction des dimensions. Ces structures sont appelées 'quadruple-grille' ou GAAFET (pour '*gate-all-around*', i.e. 'à grille enrobante'). Proposée au début des années 90 [Colinge'90], le GAAFET est réalisé à partir d'un substrat SOI sur lequel le film de silicium est gravé de manière à ne laisser qu'une membrane de silicium. L'oxyde enterré sous la membrane est alors retiré par gravure humide (au FH) (Fig. 31.a) puis la cavité ainsi créée est remplie avec l'empilement de grille (Fig. 31.b). L'inconvénient de ce procédé est la difficulté de réaliser des transistors de très petites dimensions.

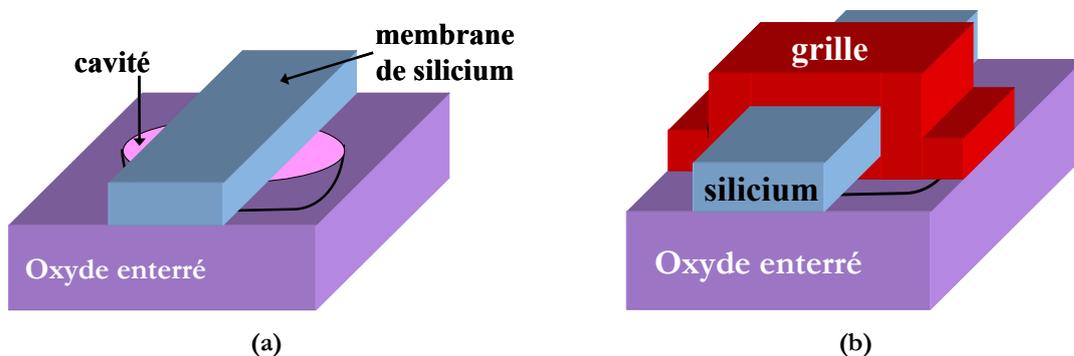


Fig. 31: Schéma de la cavité creusée sous l'oxyde enterré (a) et transistor GAAFET après dépôt de la grille (b).

5 Technologie SON (Silicon-On-Nothing)

Le procédé SON (*Silicon-On-Nothing*, brevet *ST Microelectronics*) a été proposé comme alternative pour réaliser des transistors à grille enrobante à partir de substrats en silicium massif [Jurczak'00][Monfray'01]. Cette méthode opère en quelque sorte en 'mode SOI localisé'; elle consiste à définir la zone active grâce à une isolation STI (*Shallow Trench Isolation*) et à réaliser une double épitaxie : une épitaxie sélective d'une couche de SiGe et une épitaxie non sélective de Si (Fig. 32.a). Après une série d'étapes conventionnelles jusqu'au dépôt des espaceurs (Fig. 32.b), la couche de SiGe est retirée sélectivement par gravure isotrope (dite gravure tunnel), laissant la membrane de silicium suspendue (Fig. 32.c). Le tunnel peut ensuite être rempli avec l'oxyde et la grille arrière (Fig. 32.d) afin de former une structure double-grille ou à grille enrobante. Pour terminer, les zones de source et drain sont créées par épitaxie (Fig. 32.e). Par contre, la grille est débordante (Fig. 32.f) et pourrait induire de fortes capacités de recouvrement. Ce procédé a l'avantage d'être pragmatique et d'offrir de bonnes perspectives de miniaturisation (les plus petits transistors réalisés sur SON ont une longueur de grille de 25nm [Harrison04-2]).

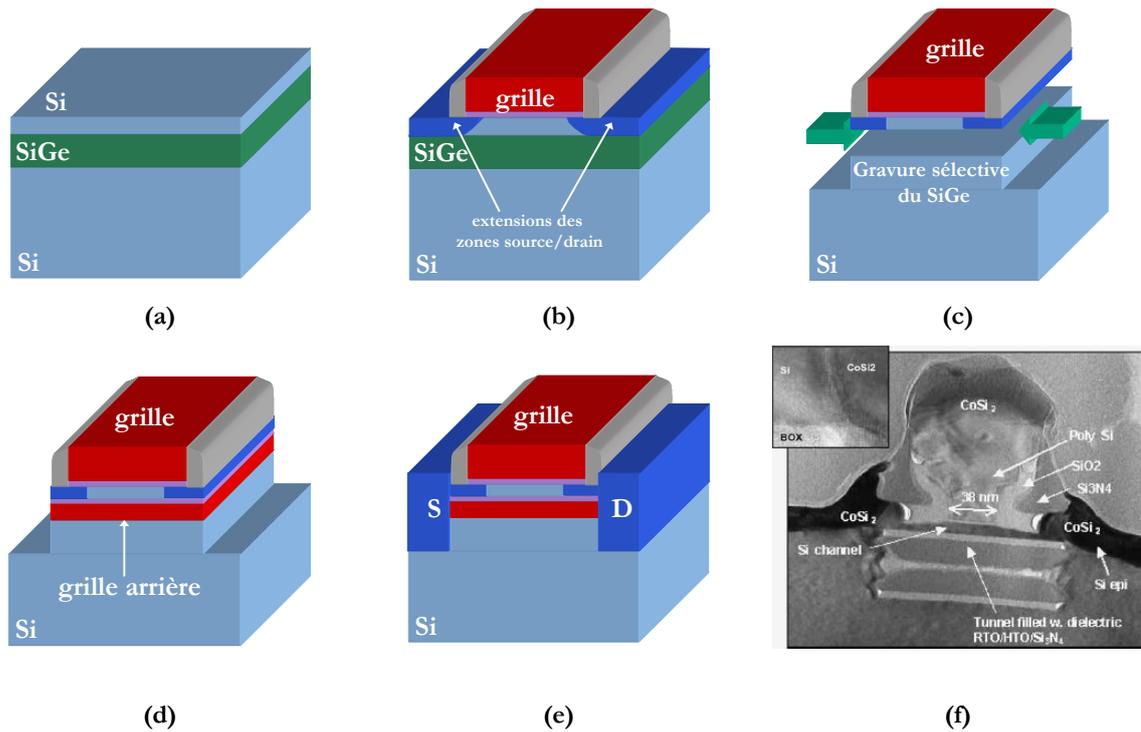


Fig. 32: (a-e): Description des étapes d'un procédé SON double-grille. (f): photo TEM d'un transistor SON de 38 nm de longueur de grille (d'après [Monfray'04-2]).

6 Dispositifs multicanaux et nanofils

6.1 Dispositifs à multicanaux

6.1.1 McFET

Cette méthode développée par *Samsung* [Kim'04][Kim'05][Kim'06] permet de passer outre la limite due à la côte lithographique.

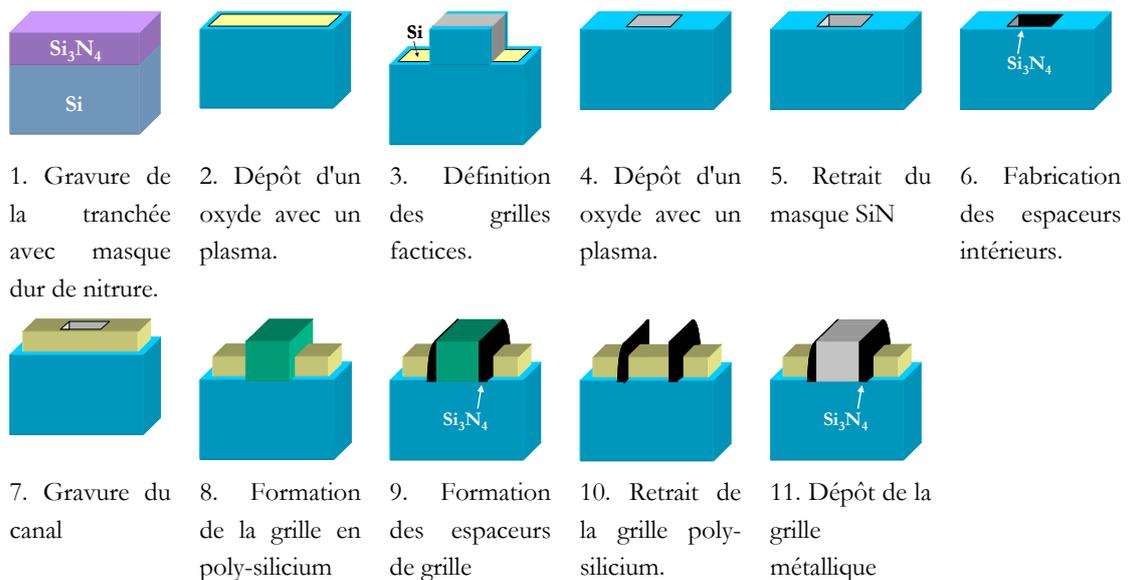


Fig. 33: Procédé de fabrication d'un McFET à grille métallique (d'après [Kim'05]).

Elle consiste grossièrement à creuser un trou dans la zone active puis à le remplir avec le matériau de grille (Fig. 33). On obtient ainsi schématiquement deux Triple-grilles pouvant être assez hauts et surtout très étroits et rapprochés (par exemple deux ailerons de 12 nm de large avec un espacement de 65 nm dans la Fig. 34).

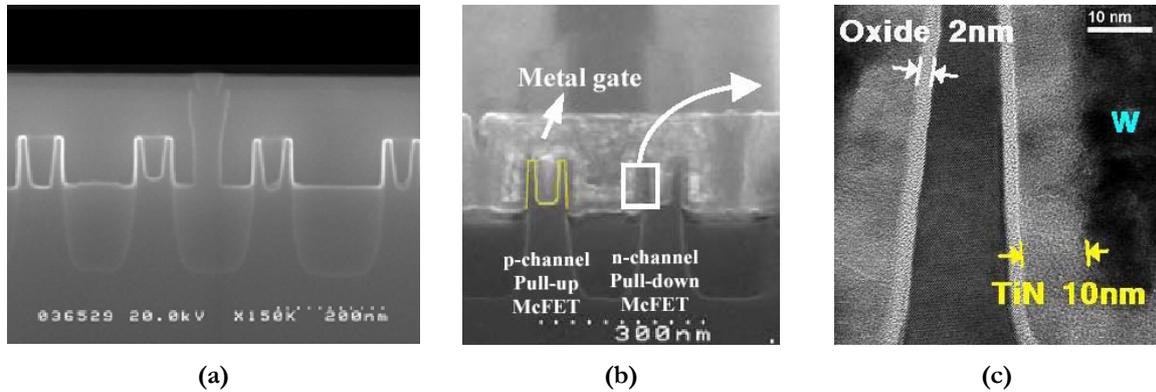


Fig. 34: Photo MEB d'un réseau McFET (a) et d'une cellule SRAM McFET (b). Insert: photo TEM montrant l'aileron gravé, avec l'oxyde et la grille métallique (c) (d'après [Kim'05]).

6.1.2 MBCFET

C'est une variante du procédé SON, s'appuyant sur la possibilité de graver sélectivement le SiGe afin d'obtenir une structure multicanaux de type 'complètement enrobé'. On conjugue ainsi un excellent contrôle des effets de canaux courts avec un courant de sortie dépendant du nombre de canaux que l'on arrive à fabriquer.

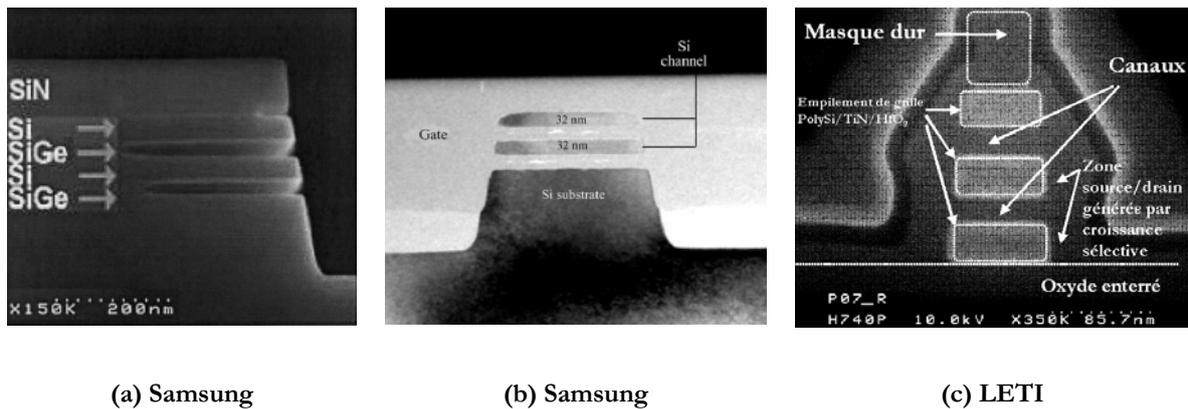


Fig. 35: Gravure sélective du SiGe par rapport au silicium (a), coupe TEM transversale faisant apparaître les canaux complètement entourés par la grille (b) et coupe longitudinale d'une structure MBCFET sur SOI (c). D'après [Lee'03] (a), [Lee'04-2] (b) et photo de E. Bernard (c).

L'empilement étant vertical, on obtient une structure compacte, donc une très bonne intégrabilité. Si tous les canaux ont des performances homogènes (surtout en ce qui concerne les dimensions des canaux et la tension de seuil), on dispose donc en théorie de la structure idéale aussi bien pour le courant de sortie que pour les courants de fuite. Ce type de structure

a été proposé en 2003 par *Samsung* et est depuis régulièrement améliorée (voir Fig. 35) [Lee'03][Lee'04][Lee'04-2][Yoon'04][Park'04]. Une description du procédé est exposée dans la Fig. 36.

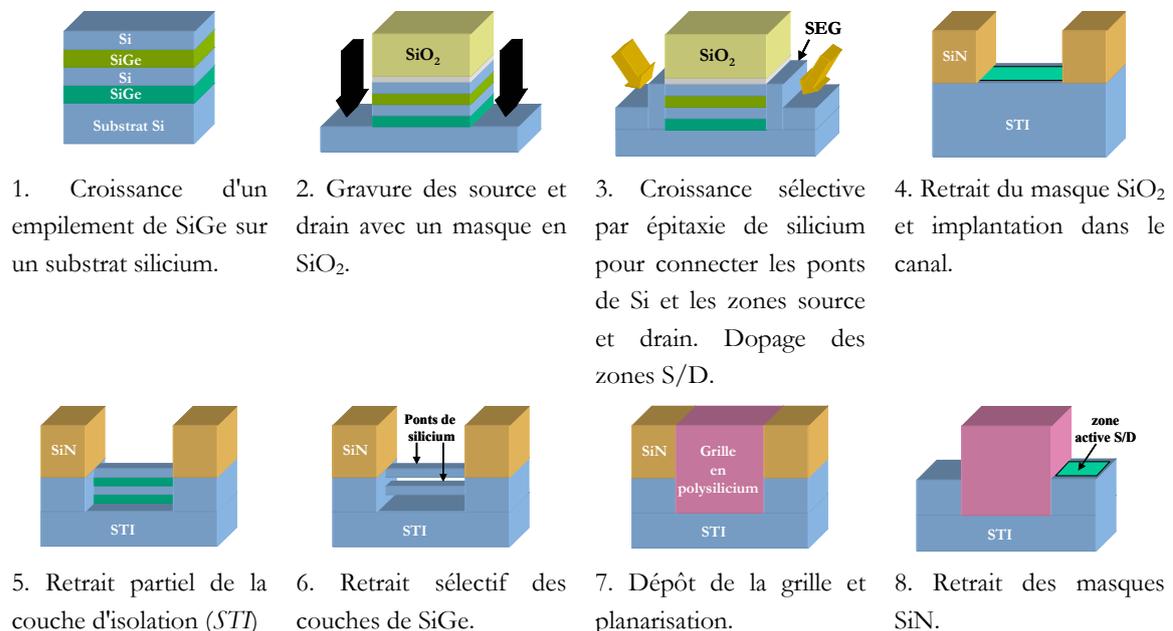


Fig. 36: Schéma des étapes de fabrication d'un transistor à multicanaux MBCFET sur Si massif. (d'après [Lee'04-2]).

6.2 Dispositifs à nanofils

C'est l'extension du concept multicanal à une matrice de transistors (Fig. 37.a). Le principe de la fabrication réside dans l'épitaxie de couches empilées Si/SiGe. On effectue ensuite deux gravures: une gravure latérale qui correspond à la gravure de la zone active puis une gravure horizontale sélective du SiGe par rapport au Si. On obtient ainsi une matrice de transistors à nanofils. Après dépôt de l'oxyde de grille et de la grille, la configuration obtenue est donc du type 'grille enrobante'.

Des dispositifs nanofils ont été récemment présentés par le LETI [Ernst'06]. Des nanofils en réseaux (Fig. 37.b) pleinement fonctionnels avec une longueur de grille descendant jusqu'à $L_G = 80$ nm ont montré d'excellentes pentes sous le seuil et une quasi-absence de DIBL. Le problème avec de telles structures vient du collage intervenant parfois entre les ponts de silicium après la gravure du SiGe (Fig. 37.c) et de la difficulté à faire diffuser les dopants de source et drain uniformément dans la structure. Sur ces dispositifs, le transistor du bas n'est pas 'à grille enrobante' mais de type 'Triple-grille' si une plaque de départ SOI classique est utilisée. Cependant, des simulations numériques montrent (voir chap.4 section 3.2) que l'immunité aux effets de canaux courts n'en est pas considérablement affectée même aux très faibles longueurs de grille. Il existe également la possibilité avec une plaque de base 'SiGe sur BOX' de réaliser des transistors 'à grille enrobante' partout, y compris à l'interface silicium/BOX.

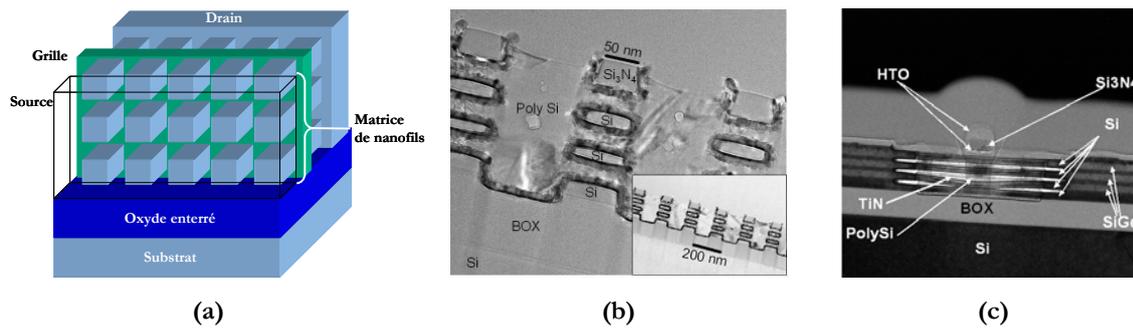


Fig. 37: Schéma de principe d'un transistor à nanofils (a). Vue en coupe transversale (b) et longitudinale (c) des nanofils (d'après [Ernst'06]).

7 Autres dispositifs multigrilles

7.1 Structure en 'T inversé' - ITFET

Dans ce procédé récemment démontré [Mathew'05], on ne grave pas la couche de silicium jusqu'à l'oxyde enterré. On regrave alors la zone active en protégeant la zone non gravée de manière à former une structure 'en T inversé' (Fig. 38.a). On obtient une structure hybride entre un réseau de transistors Triple-grille et un transistor planaire complètement déserté (Fig. 38.b et Fig. 38.c). L'avantage d'une telle structure réside dans la surface de canal disponible: en combinant surfaces planes et verticales on optimise la surface de conduction tout en gardant des canaux fins. On dispose ainsi d'un dispositif alliant un fort I_{ON} et un bon contrôle électrostatique. Cependant, les différences probables d'épaisseurs entre d'une part les deux grilles verticales et d'autre part les surfaces planes peuvent induire des fluctuations de tensions de seuil. Un ajustement avec des dopages différents entre zones planes et verticales sera peut-être nécessaire.

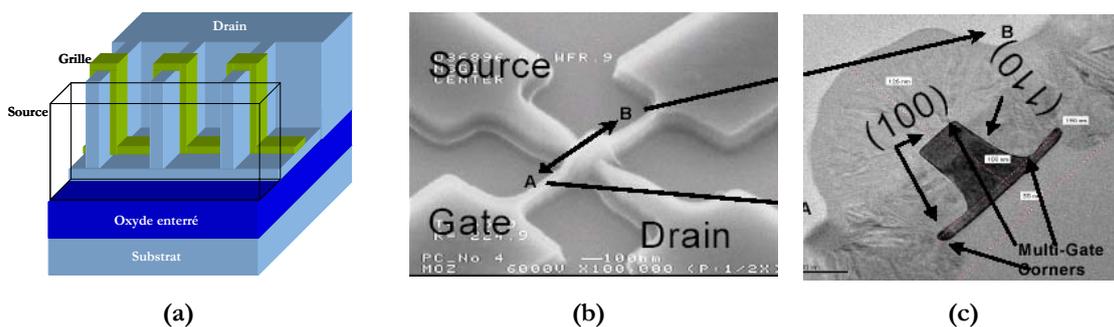


Fig. 38: Schéma de principe d'un ITFET, photo MEB d'un dispositif (b) et coupe TEM transversale montrant la région du canal en 'T inversé' et l'orientation cristalline des différents canaux de conduction (c) (D'après [Mathew'05]).

7.2 Dispositifs à grilles indépendantes

7.2.1 Les MIGFET

Les transistors double-grille sont naturellement des transistors à grilles indépendantes; après la fabrication d'un transistor Triple-grille, une étape supplémentaire de polissage mécano-chimique sur un masque dur épais permet de séparer les deux grilles (Fig. 39.a et Fig. 39.b). L'avantage de disposer de grilles indépendantes est de pouvoir moduler la tension de seuil du transistor simplement en changeant la polarisation d'une des grilles (Fig. 39.c) [Eminente'05] ou de pouvoir mixer deux signaux d'entrée [Mathew'04][Zhang'05].

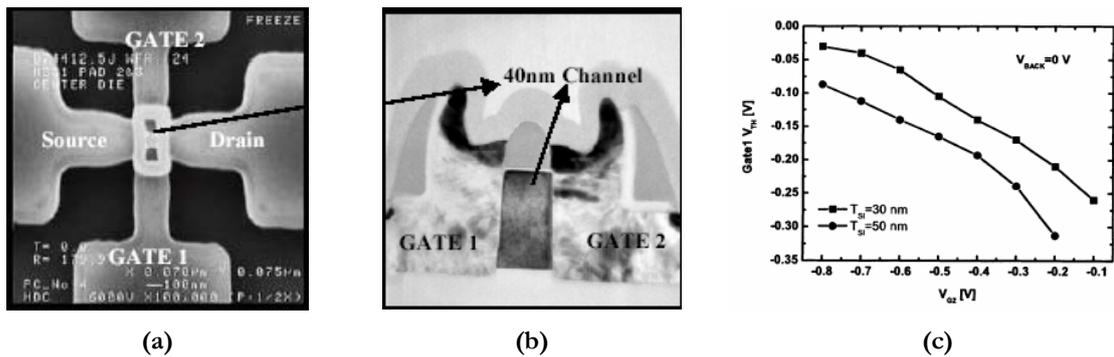


Fig. 39: (a-b): Vue du dessus et en coupe d'un MIGFET montrant les deux grilles rendues indépendantes (d'après [Mathew'05-2]). (c): Tension de seuil d'une grille en fonction de la polarisation de l'autre grille pour différentes largeurs d'aileron (d'après [Eminente'05]).

7.2.2 Le G⁴-FET

Un G⁴-FET est un dispositif à effet de champ avec quatre grilles indépendantes (Fig. 40.a). Il a la même structure qu'un transistor MOS sur SOI avec deux contacts de chaque côté du canal. Il combine ainsi les effets MOS et JFET dans un seul dispositif (possibilité d'avoir jusqu'à quatre signaux d'entrée).

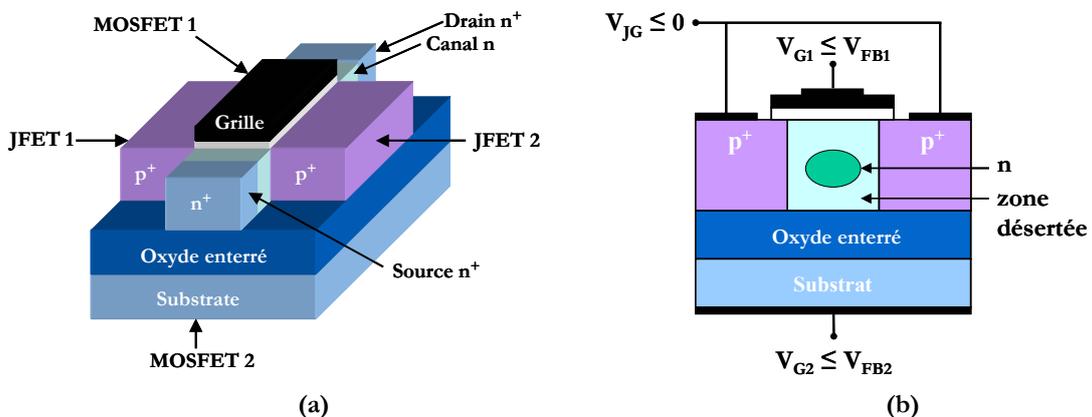


Fig. 40: Schéma d'un dispositif G⁴-FET (a) et mode de fonctionnement en régime 'Depletion-All-Around' (V_{FB}: tension de bandes plates) (b) (d'après [Akarvardar'05]).

En polarisant ses interfaces en désertion ou en inversion, le G^4 -FET fournit un canal volumique entouré des zones de désertion induites par les quatre grilles (*'Depletion-All-Around'*, Fig. 40.b). Ce régime de fonctionnement présente de bonnes performances de bruit basse fréquence et de résistance aux radiations [Akarvardar'05]. Là où le MOS à cause de sa sensibilité aux défauts d'interface est grandement handicapé, le G^4 -FET avec sa conduction volumique peut être intéressant, notamment pour des applications analogiques [Chen'05].

8 Conclusions

Dans ce chapitre, les architectures à grilles multiples ont été passées en revue. Malgré de nombreux défis technologiques liés aux films minces, aux matériaux utilisés (l'utilisation de canaux non dopés rend l'utilisation des grilles métalliques inévitable) et aux définitions des zones actives et des grilles, ces dispositifs restent très attractifs pour les prochaines générations de transistors. La faisabilité technique de ces dispositifs a été démontrée, et leurs performances sont depuis graduellement améliorées. Le compromis courant de sortie/courant de fuite reste à optimiser dans beaucoup de cas. A titre de synthèse, la Fig. 41 (resp. Fig. 42) montre les performances $I_{OFF}(I_{ON})$ pour quelques dispositifs NMOS (resp. PMOS) multigrilles publiés. La normalisation utilisée dans ces deux figures est de type 'physique', c'est à dire faite par la surface totale des canaux, et non par l'aire d'occupation sur la plaque.

La Fig. 43 résume les avantages et inconvénients comparés des structures présentées dans ce chapitre. Le prochain chapitre va s'attacher à présenter plus en détail les transistors verticaux de type Ω FET réalisés dans le cadre de cette thèse.

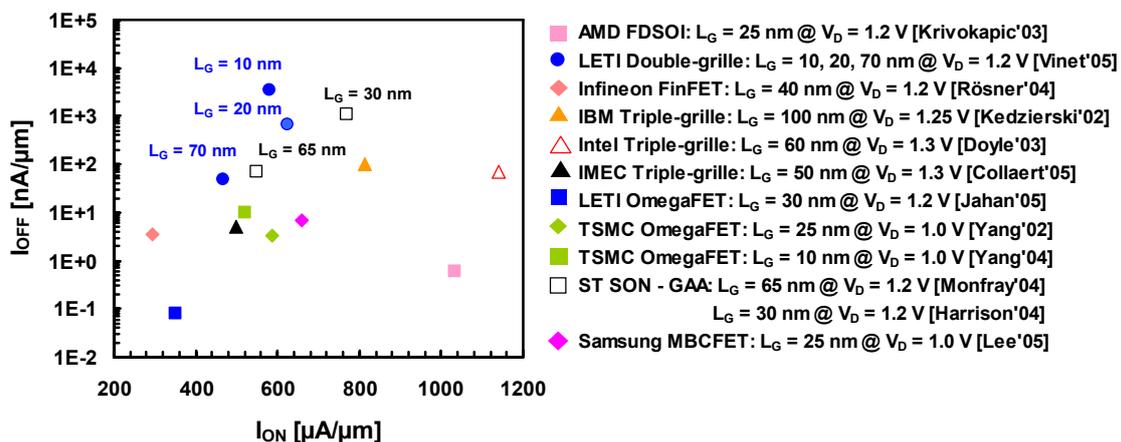


Fig. 41: Performances comparées de quelques composants NMOS multigrille. Les longueurs de grilles et les tensions d'alimentation sont indiquées sur le graphique. La normalisation des courants est faite en considérant la largeur totale des canaux. La référence [Lee'06] pour $L_G = 5$ nm n'est pas montrée en raison de l'absence de performances I_{OFF} publiées.

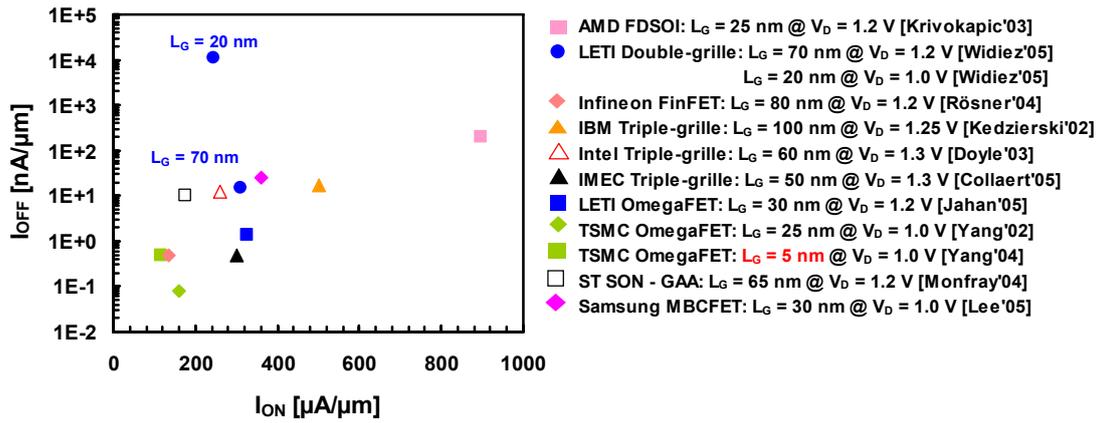


Fig. 42: Performances comparées de quelques composants PMOS multigrille. Les longueurs de grilles et les tensions d'alimentation sont indiquées sur le graphique. La normalisation des courants est faite en considérant la largeur totale des canaux.

		Avancement	Avantages principaux	Inconvénients principaux	Potentiel pour les nœuds $\leq 30 \text{ nm}$
Simple grille bulk		Production	. Procédés bien maîtrisés	. Contrôle des effets de canaux courts	NON
Simple grille contraint		Développement	. Mobilité améliorée	. Relaxation des couches contraintes pour les petites dimensions	OUI
Simple grille PDSOI		Développement / Production	. Technologie pragmatique	. Effets de <i>body</i> flottant	PEUT- ETRE
Simple grille FDSOI		Développement	. Pas d'effet de <i>body</i> flottant	. Faibles épaisseurs (bien contrôlées) de film nécessaires. . Champs dans l'oxyde enterré.	OUI
SON		Recherche	. Technologie innovante . Quasi SOI	. Fortes capacités de recouvrement	OUI
Double Grille planaires		Recherche	. Deux canaux de conduction . Bon contrôle des effets de canaux courts	. Faibles épaisseurs de canal nécessaire . Auto alignement des grilles	OUI
FinFET		Recherche	. Auto-alignement des grilles . Procédé assez proche du silicium massif	. Espacement lithographique . Dopage des extensions . Faibles largeurs d'aileron nécessaires	OUI

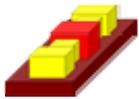
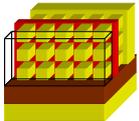
				<ul style="list-style-type: none"> . Les flancs doivent être bien verticaux . Dispersion des résultats . Mobilité sur les flancs 	
Triple-Grille		Recherche	<ul style="list-style-type: none"> . Trois canaux de conduction . Auto-alignement des grilles 	<ul style="list-style-type: none"> . Espacement lithographique . Dopage des extensions . Contrôle des flancs et des coins . Dispersion des résultats . Mobilité sur les flancs 	OUI
Grille enrobée		Recherche	<ul style="list-style-type: none"> . Excellent contrôle électrostatique 	<ul style="list-style-type: none"> . Dopage des extensions . Epaisseurs faibles nécessaires 	OUI
Multicanaux		Recherche	<ul style="list-style-type: none"> . Bonne densité d'intégration . Excellent contrôle électrostatique 	<ul style="list-style-type: none"> . Procédé difficile à maîtriser à l'heure actuelle . Dopage des extensions 	OUI
Nanofils		Recherche	<ul style="list-style-type: none"> . Très forte densité d'intégration . Excellent contrôle électrostatique 	<ul style="list-style-type: none"> . Procédé difficile à maîtriser à l'heure actuelle . Dopage des extensions . Contrôle des petites dimensions 	OUI

Fig. 43: Comparatif des avantages et des inconvénients des dispositifs MOS avancés.

Références du chapitre 2

- [Akarvardar'05] K. Akarvardar, S. Cristoloveanu, B. Dufrene, P. Gentil, R. D. Schrimph, B. J. Blalock, J. A. Chroboczek, M. Mojarradi, "Evidence for Reduction of Noise and Radiation Effects in G4-FET Depletion-All-Around Operation", proceedings of the *35th European Solid-State Device Research conference (ESSDERC'05)*, pp. 89-92, 2005.
- [Assaderaghi'94] F. Assaderaghi, D. Sinitsky, S. Parke, J. Bokor, P. K. Ko, C. Hu, "A Dynamic Threshold Voltage MOSFET (DTMOS) for Ultra-Low Voltage Operation", *IEDM'94 Technical Digest*, pp. 809-812, 1994.
- [Chen'05] S. Chen, J. Vandersand, B. J. Blalock, K. Akarvardar, S. Cristoloveanu, M. Mojarradi, "SOI Four-Gate Transistors (G4-FETs) for High Voltage Analog Applications", proceedings of the *31st European Solid-State Circuits Research conference (ESSCIRC'05)*, pp. 311-314, 2005.
- [Choi'02] Y.-K Choi, T.-J. King, C. Hu, "Nanoscale CMOS spacer FinFET for the terabit era", *IEEE Electron Device Letters*, vol. 23, no. 1, pp. 25-27, 2002.
- [Colinge'88] J.-P. Colinge, "Thin-film, accumulation-mode p-channel SOI MOSFETs", *Electronic letters*, Vol. 24, no. 5, 3 March 1988, pp. 257-258, 1988.
- [Colinge'90] J.-P. Colinge, M.H. Gao, A. Romano-Rodriguez, H. Maes, C. Claeys, "Silicon-on-insulator "Gate-All-Around Device" ", *IEDM'90 Technical Digest*, pp. 595-598, 1990.
- [Colinge'04] J.-P. Colinge, "Multiple-gate SOI MOSFETs", *Solid State Electronics*, vol. 48, no. 6, pp. 897-905, 2004.
- [Collaert'05] N. Collaert, M. Demand, I. Ferain, R. Singanamalla, P. Zimmerman, Y.S. Yim, T. Schram, G. Mannaert, M. Goodwin, J.C. Hooker, F. Neuilly, M.C. Kim, K. De Meyer, S. De Gendt, W. Boullart, M. Jurczak, S. Biesemans, "Tall Triple-Gate Devices with TiN/HfO₂ Gate Stack", *Digest of Technical Papers, 2005 Symposium on VLSI Technology*, pp. 108-109, 2005.
- [Douseki'01] T. Douseki, N. Shibata, J. Yamada, "SIMOX ROM Macro with Low-V_{th} Memory Cells", *IEEE Int. SOI conf. 2001*, pp. 143-144, 2001.

- [Doyle'03] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R. Rios, R. Chau, "Tri-Gate Fully-Depleted CMOS Transistors: Fabrication, Design and Layout", *Digest of Technical Papers, 2003 Symposium on VLSI Technology*, pp. 133-134, 2003.
- [Eminente'05] S. Eminente, K.-I. Na, S. Cristoloveanu, L. Mathew, A. Vandooren, "Lateral and vertical coupling effects in MIGFETs", *IEEE Int. SOI conf.* 2005, pp. 94-95, 2004.
- [Ernst'06] T. Ernst, C. Dupré, C. Isheden, E. Bernard, R. Ritzenthaler, V. Maffini-Alvaro, J. Cluzel, A. Toffoli, C. Vizioz, S. Borel, F. Andrieu, F. De Crecy, V. Delaye, D. Lafond, G. Rabillé, J.-M. Hartmann, M. Rivoire, B. Guillaumot, A. Suhm, P. Rivallin, O. Faynot, G. Ghibaud, S. Deleonibus, "Novel 3D integration process for highly scalable Nano-Beam stacked-channels GAA (NBG) CMOSFETs with HfO₂/TiN gate stack", *IEDM'06 Technical Digest*, 2006. (accepté)
- [Guarini'01] K. W. Guarini, P. M. Solomon, Y. Zhang, K.K. Chan, E. C. Jones, G. M. Cohen, A. Krasnoperova, M. Ronay, O. Dokumaci, J. J. Bucchignano, C. Cabral Jr., C. Lavoie, V. Ku, D. C. Boyd, K. S. Petrarca, I. V. Babich, J. Treichler, P. M. Kozlowski, J. S. Newbury, C. P. D'Emic, R. M. Sicina, H.-S.P. Wong, "Triple-Self-Aligned, Planar Double-Gate MOSFETs: Devices and Circuits", *IEDM'01 Technical Digest*, pp. 425-428, 2001.
- [Harrison'04] S. Harrison, D. Munteanu, J.L. Autran, A. Cros, R. Cerutti, T. Skotnicki, "Electrical characterization and modelling of high-performance SON DG MOSFETs", proceedings of the *34th European Solid-State Device Research conference (ESSDERC'04)*, pp. 373-376, 2004.
- [Harrison04-2] S. Harrison, P. Coronel, A. Cros, R. Cerutti, F. Leverd, A. Beverina, R. Wacquez, J. Bustos, D. Delille, B. Tavel, D. Barge, J. Bienacel, M.-P. Samson, F. Martin, S. Maitrejean, D. Munteanu, J.-L. Autran, T. Skotnicki, "Poly-gate Replacement Through Contact Hole (PRETCH): A new method for High-K/Metal gate and multi-oxide implementation on chip", *IEDM'04 Tech. Dig.*, pp. 291-294, 2004.
- [Hisamoto'89] D. Hisamoto, T. Kaga, Y. Kawamoto, E. Takeda, "A fully depleted lean-channel transistor (DELTA)-a novel vertical ultra thin SOI MOSFET", *IEDM'94 Technical Digest*, pp. 833-836, 1989.
- [Hisamoto'98] D. Hisamoto, W.-C. Lee, J. Kedzierski, E. Anderson, H. Takeuchi, K. Asano, T.-J. King, J. Bokor, C. Hu, "A Folded-channel MOSFET for Deep-sub-tenth Micron Era", *IEDM'98 Technical Digest*, pp. 1032-1034, 1998.

- [Hisamoto'00] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, C. Hu, "FinFET – A Self-Aligned Double-Gate MOSFET Scalable to 20 nm", *Electron Devices, IEEE Transactions on*, vol. 47, no. 12, pp. 2320-2325, 2000.
- [Jacob'05] S. Jacob, B. De Salvo, G. Le Carval, S. Deleonibus, "Modeling of a Double-Gate FinFlash memory", *proceedings of the 1st International Conference on Memory Technology and Design (ICMTD 2005)*, pp. 153-156, 2005.
- [Jahan'05] C. Jahan, O. Faynot, M. Cassé, R. Ritzenthaler, L. Brévard, L. Tosti, X. Garros, C. Vizioz, F. Allain, A.M. Papon, H. Dansas, F. Martin, M. Vinet, B. Guillaumot, A. Toffoli, B. Giffard and S. Deleonibus, "10nm Ω FETs transistors with TiN metal gate and HfO₂", *VLSI Tech. Dig.*, pp. 112-113 ,2005.
- [Jiménez'04] D. Jiménez, B. Iñiguez, J. Suñé, L. F. Marsal, J. Pallarès, J. Roig, D. Flores, "Continuous Analytical I-V Model for Surrounding-Gate MOSFETs", *IEEE Electron Device Letters*, vol. 25, no. 8, pp. 571-573, 2004.
- [Jurczak'00] M. Jurczak, T. Skotnicki, M. Paoli, B. Tormen, J. Martins, J.L. Regolini, D. Dutartre, P. Ribot, D. Lenoble, R. Pantel, S. Monfray, "Silicon-on-nothing (SON) an innovating process for advanced CMOS", *Electron Devices, IEEE Transactions on* , vol. 47, no. 11, pp. 2179-, 2000.
- [Kedzierski'02] J. Kedzierski, E. Nowak, T. Kanarsky, Y. Zhang, D. Boyd, R. Carruthers, C. Cabral, R. Amos, C. Lavoie, R. Roy, J. Newbury, E. Sullivan, J. Benedict, P. Saunders, K. Wong, D. Canaperi, M. Krishnan, K.-L. Lee, B. A. Rainey, D. Fried, P. Cottrell, H.-S. P. Wong, M. Jeong, W. Haensch, "Metal-gate FinFET and fully depleted SOI devices using total gate silicidation", *IEDM'02 Technical Digest*, pp. 247-250, 2003.
- [Kedzierski'04] J. Kedzierski, M. Jeong, T. Kanarsky, Y. Zhang, H-S.P. Wong, "Fabrication of Metal Gates finFETs Through Complete Gate Silicidation With Ni", *Electron Devices, IEEE Transactions on* , vol. 51, no. 12, p. 2115-2120 2004.
- [Kim'04] S.-M. Kim, E.-J. Yoon, H.-J. Jo, M. Li, C.-W. Oh, S.-Y. Lee, K.-H. Yeo, M.-S. Kim, S.-H. Kim, D.-U. Choe, J.-D. Choe, S.-D. Suk, D.-W. Kim, D. Park, K. Kim, B.-I. Ryu, "A Novel Multi-channel Field Effect Transistor (McFET) on Bulk Si for High Performance Sub-80 nm Application", *IEDM'04 Technical Digest*, pp. 639-642, 2004.

- [Kim'05] S.-M. Kim, E.-J. Yoon, M.-S. Kim, C.-W. Oh, S.-D. Suk, M. Li, S.-Y. Lee, K.-H. Yeo, S.-H. Kim, D.-U. Choe, D.-W. Kim, D. Park, K. Kim, B.-I. Ryu, "Fully Working High Performance Multi-channel Field Effect Transistor (McFET) SRAM Cell on Bulk Si substrate Using TiN Single Metal Gate", *Digest of Technical Papers, 2005 Symposium on VLSI Technology*, pp. 196-197, 2005.
- [Kim'06] S.-M. Kim, E.-J. Yoon, M.-S. Kim, M. Li, C.-W. Oh, S.-Y. Lee, K.-H. Yeo, S.-H. Kim, D.-U. Choe, S.-D. Suk, D.-W. Kim, D. Park, "Highly Manufacturable 65 nm McFET (Multi-channel Field Effect Transistor) SRAM Cell with Extremely High Performance", *Institute of Electronics Engineers of Korea's Journal of Semiconductor Technology and Science*, vol. 6, no. 1, March 2006.
- [Kim'06-2] S.-Y. Kim, K.-R. Han, B.-K. Choi, S.-H. Kong, J.-S. Lee, J.-H. Lee, "Negative Bias Temperature Instability of Bulk Fin Field Effect Transistor", *Japanese Journal of Applied Physics*, Vol. 45, no. 3A, pp. 1467-1470, 2006.
- [Krivokapic'03] Z. Krivokapic, V. Moroz, W. Maszara, M.-R. Lin, "Locally Strained Ultra-Thin Channel 25nm Narrow FDSOI Devices with Metal Gate and Mesa Isolation", *IEDM'03 Technical Digest*, pp. 445-448, 2003.
- [Kuriyama'06] A. Kuriyama, O. Faynot, L. Brévard, J. Mitard, V. Vidal, S. Deleonibus, S. Cristoloveanu, H. Iwai, "Work Function Investigation in Advanced Metal Gate-HfO₂-SiO₂ Systems with Bevel Structures", proceedings of the 36th European Solid-State Device Research conference (ESSDERC'06), 2006.
- [Lee'03] S.-Y. Lee, S.-M. Kim, E.-J. Yoon, C.-W. Oh, I. Chung, D. Park, K. Kim, "A Novel Multibrige-Channel MOSFET (MBCFET): Fabrication Technologies and Characteristics", *Nanotechnology, IEEE Transactions on*, vol. 2, no. 04, pp. 253-257, 2003.
- [Lee'04] S.-Y. Lee, E.-J. Yoon, S.-M. Kim, C.-W. Oh, M. Li, J.-D. Choi, K.-H. Yeo, M.-S. Kim, H.-J. Cho, S.-H. Kim, D.-W. Kim, D. Park, K. Kim, "A Novel Sub-50 nm Multi-Bridge-Channel MOSFET (MBCFET) with Extremely High Performance", *Digest of Technical Papers, 2004 Symposium on VLSI Technology*, pp. 200-201, 2004.
- [Lee'04-2] S.-Y. Lee, S.-M. Kim, E.-J. Yoon, C.-W. Oh, I. Chung, D. Park, K. Kim, "Three-Dimensional MBCFET as an Ultimate Transistor", *IEEE Electron Device Letters*, vol. 25, no. 4, pp. 217-219, 2004.

- [Lee'05] S.-Y. Lee, E.-J. Yoon, D.-S. Shin, S.-M. Kim, S.-D. Suk, M.-S. Kim, D.-W. Kim, D. Park, K. Kim, B.-I. Ryu, "Sub-25nm Single-Metal Gate CMOS Multi-Bridge-Channel MOSFET (MBCFET) for High Performance and Low Power Application", *Digest of Technical Papers, 2005 Symposium on VLSI Technology*, pp. 145-155, 2005.
- [Lee'06] H. Lee, L.-E. Yu, S.-W. Ryu, J.-W. Han, K. Jeon, D.-Y. Jang, K.-H. Kim, J. Lee, J.-H. Kim, S. C. Jeon, G. S. Lee, J. S. Oh, Y. C. Park, W. H. Bae, H. M. Lee, J. M. Yang, J. J. Yoo, S. I. Kim, Y.-K. Choi, "Sub-5nm All-Around Gate FinFET for Ultimate Scaling", *VLSI Tech. Dig.*, 2006. (à paraître)
- [Liu'03] Y. X. Liu, M. Masahara, K. Ishii, T. Tsutsumi, T. Sekigawa, H. Takashima, H. Yamauchi, E. Suzuki, "Flexible threshold voltage FinFETs with independent double gates and an ideal rectangular cross-section Si-Fin channel", *IEDM'03 Technical Digest*, pp. 986-988, 2003.
- [Liu'03-2] Y.X. Liu, K. Ishii, T. Tsutsumi, M. Masahara, E. Suzuki, "Ideal Rectangular Cross-Section Si-Fin Channel Double-Gate MOSFETs Fabricated Using Orientation-Dependent Wet Etching", *IEEE Electron Device Letters*, vol. 24, no. 7, pp. 484-486, 2003.
- [Mathew'04] L. Mathew, Y. Du, A. V.-Y. Thean, M. Sadd, A. Vandooren, C. Parker, T. Stephens, R. Mora, R. Rai, M. Zavala, D. Sing, S. Kalpat, J. Hughes, R. Shimer, S. Jallepalli, G. Workman, W. Zhang, J. G. Fossum, B. E. White, B. -Y. Nguyen, J. Mogab, "CMOS Vertical Multiple Independent Gate Field Effect Transistor (MIGFET)", *IEEE Int. SOI conf.* 2004, pp. 187-189, 2004.
- [Mathew'05] L. Mathew, M. Sadd, S. Kalpat, M. Zavala, T. Stephens, R. Mora, S. Bagchi, C. Parker, J. Vasek, D. Sing, R. Shimer, L. Prabhu, G. O. Workman, G. Ablen, Z. Shi, J. Saenz, B. Min, D. Burnett, B.-Y. Nguyen, J. Mogab, M.M. Chowchury, W. Zhang, J. G. Fossum, "Inverted T channel FET (ITFET) – Fabrication and Characteristics of Vertical-Horizontal, Thin Body, Multi-Gate, Multi-Orientation Devices, ITFET SRAM Bit-cell operation. A Novel Technology for 45nm and Beyond CMOS", *IEDM'05 Technical Digest*, pp. 713-717, 2005.
- [Mathew'05-2] L. Mathew, Y. Du, S. Kaipat, M. Sadd, M. Zavala, T. Stephens, R. Mora, R. Rai, S. Becker, C. Parker, D. Sing, R. Shimer, J. Saenz, A.V.-Y. Thean, L. Prabhu, M. Moosa, B.-Y. Nguyen, J. Mogab, G. Workman, A. Vandooren, Z. Shi, M. M. Chowdhury, W. Zhang, J.G. Fossum, " Multiple independent gate field effect transistor (MIGFET) - multi-fin RF mixer architecture, three independent gates (MIGFET-T) operation and temperature characteristics", *Digest of Technical Papers, 2005 Symposium on VLSI Technology*, pp. 200-201, 2005.

- [Miyano'92] S. Miyano, M. Hirose, F. Masuoka, "Numerical analysis of a cylindrical thin-pillar transistor (CYNTHIA)", *Electron Devices, IEEE Transactions on*, Vol. 39, no. 8, pp. 1876-1881, 1992.
- [Monfray'01] S. Monfray, T. Skotnicki, Y. Morand, S. Descombes, M. Paoli, P. Ribot, A. Talbot, D. Dutartre, F. Leverd, Y. Lefric, R. Pantel, M. Haond, D. Renaud, M.-E. Nier, C; Vizios, D. Louis, N. Buffet, "First 80 nm SON (Silicon-On-Nothing) MOSFETs with perfect morphology and high electrical performance", *IEDM'01 Technical Digest*, pp. 645-648, 2001.
- [Monfray'04] S.Monfray, D.Chanemougame, S.Borel, A.Talbot, F. Leverd, N.Planes, D.Delille, D.Dutartre, R. Palla, Y.Morand, S.Descombes, M-P.Samson, N.Vulliet, T.Sparks, A.Vandooren and T. Skotnicki, "SON (Silicon-On-Nothing) technological CMOS Platform: Highly performant devices and SRAM cells", *IEDM'04 Technical Digest*, pp. 635-638, 2004.
- [Monfray'04-2] S. Monfray, T. Skotnicki, C. Fenouillet-Beranger, N. Carriere, D. Chanemougame, Y. Morand, S. Descombes, A. Talbot, D. Dutartre, C. Jenny, P. Mazoyer, R. Palla, F. Leverd, Y. Le Fric, R. Pantel, S. Borel, D. Louis, N. Buffet, "Emerging silicon-on-nothing (SON) devices technology", *Solid State Electronics*, vol. 46, no. 6, pp. 887-895, 2004.
- [Mueller'64] C. W. Mueller, P. H. Robinson, "Grown-film silicon transistors on sapphire", *proceedings of the IEEE*, vol. 52, no. 12, pp. 1497-1490, 1964.
- [Park'01] J.-T. Park, J.-P. Colinge, C.H. Diaz, "Pi-Gate SOI MOSFET", *IEEE Electron Device Letters*, vol. 22, no. 8, pp. 405-406, 2001.
- [Park'01-2] J.-T. Park, C. A. Colinge, J.-P. Colinge, "Comparison of Gate structures for Short-Channel SOI MOSFETs", *IEEE Int. SOI conf.*, pp. 115-116, 2001.
- [Park'04] D. Park, K. Kim, B.-I. Ryu, "3-dimensional nano-CMOS Transistors to Overcome Scaling Limits", *Proceedings of the 2004 7th International Conference on Solid-State and Integrated-Circuit Technology (SSICT'04)*, pp. 35-40, 2004.
- [Park'06] T.-S. Park, H.-J. Cho, J.-D. Choe, S.-Y. Han, D. Park, K. Kim, E. Yoon, J.-H. Lee, "Characteristics of the Full CMOS SRAM Cell Using Body-Tied TG MOSFETs (Bulk FinFETs)", *Electron Devices, IEEE Transactions on*, vol. 53, no. 3, pp. 481-487, 2006.
- [Rösner'04] W. Rösner, E. Landgraf, J. Kretz, L. Dreeskorfeld, H. Schäfer, M. Städele, T. Schulz, F. Hofmann, R. J. Luyken, M. Specht, J. Hartwich, W. Pamler, L. Risch, "Nanoscale FinFETs for low power applications", *Solid-State Electronics*, vol. 48, no. 10-11, pp. 1819-1823, 2004.

- [Vinet'05] M. Vinet, T. Poiroux, J. Widiez, J. Lolivier, B. Previtali, C. Vizioz, B. Guillaumot, Y. Le Tiec, P. Besson, B. Biasse, F. Allain, M. Cassé, D. Lafond, J.-M. Hartmann, Y. Morand, J. Chiaroni, S. Deleonibus, "Bonded Planar Double-Metal-Gate NMOS Transistors Down to 10 nm", *IEEE Electron Device Letters*, vol. 26, no. 5, pp. 317-319, 2005.
- [Wakabayashi'01] H. Wakabayashi, Y. Saito, K. Takeuchi, T. Mogami, T. Kunio, "A Dual-Metal Gate CMOS Technology Using Nitrogen-Concentration-Controlled TiN_x Film", *Electron Devices, IEEE Transactions on*, vol. 48, no. 10, pp. 2363-2369, 2001.
- [Widiez'04] J. Widiez, F. Dauge, M. Vinet, T. Poiroux, B. Previtali, M. Mouis, S. Deleonibus, "Experimental gate misalignment analysis on double gate SOI MOSFETs", *IEEE Int. SOI conf. 2001*, pp. 185-186, 2004.
- [Widiez'05] J. Widiez, "Etude, fabrication et caractérisation de transistors CMOS double grille planaires déca-nanométriques", *thèse de doctorat*, INPG, 2005.
- [Widiez'05-2] J. Widiez, J. Lolivier, M. Vinet, T. Poiroux, B. Previtali, F. Dauge, M. Mouis, S. Deleonibus, "Experimental evaluation of gate architecture influence on DG SOI MOSFETs performance", *Electron Devices, IEEE Transactions on*, vol. 52, no. 8, pp. 1772-1779, 2005.
- [Yang'02] F.-L. Yang, H.-Y. Chen, F.C. Chen, C.-C. Huang, C.-Y. Chang, H.-K. Chiu, C.-C. Lee, C.-C. Chen, H.-T. Huang, C.-J. Chen, H.-J. Tao, Y.-C. Yeo, M.-S. Liang, C. Hu, "25 nm CMOS Omega FETs", *IEDM'02 Technical Digest*, pp. 255-258, 2002.
- [Yang'04] F.-L. Yang, D.-H. Lee, H.-Y. Chen, C.-Y. Chang, S.-D. Liu, C.-C. Huang, T.-X. Chung, H.-W. Chen, C.-C. Huang, Y.-H. Liu, C.-C. Wu, C.-C. Chen, S.-C. Chen, Y.-T. Chen, Y.-H. Chen, C.-J. Chen, B.-W. Chan, P.-F. Hsu, J.-H. Shieh, H.-J. Tao, Y.-C. Yeo, Y. Li, J.-W. Lee, P. Chen, M.-S. Liang, C. Hu, "5nm-Gate Nanowire FinFET", *Digest of Technical Papers, 2004 Symposium on VLSI Technology*, pp. 196-197, 2004.
- [Yoon'04] E.-J. Yoon, S.-Y. Lee, S.-M. Kim, M.-S. Kim, S.-H. Kim, M. Li, S. Suk, K. Yeo, C.-W. Oh, J.-D. Choe, D. Choi, D.-W. Kim, D. Park, K. Kim, B.-I. Ryu, "Sub-30 nm Multi-Bridge-Channel MOSFET (MBCFET) with Metal Gate Electrode for Ultra High Performance Application", *IEDM'04 Technical Digest*, pp. 627-630, 2004.
- [Yu'02] B. Yu, L. Chang, S. Ahmed, H. Wang, S. Bell, C.-Y. Yang, C. Tabery, C. Ho, Qi. Xiang, T.-J. King, J. Bokor, C. Hu, M.-R. Lin, D. Kyser, "FinFET Scaling to 10nm Gate Length", *IEDM'02 Technical Digest*, pp. 251-254, 2002.

- [Zhang'03] S. Zhang, X. Lin, R. Huang, R. Han, M. Chan, "A Self-Aligned, Electrically Separable Double-Gate MOS Transistor Technology for Dynamic Threshold Voltage Application", *Electron Devices, IEEE Transactions on* , vol. 50, no. 11, pp. 2297-2299, 2003.
- [Zhang'05] W. Zhang, J. G. Fossum, L. Mathew, Y. Du, "Physical Insights Regarding Design and Performance of Independent-Gate FinFETs", *Electron Devices, IEEE Transactions on*, vol. 52, no. 10, pp. 2198-2206, 2006.

Chapitre 3

Fabrication et performances électriques des transistors réalisés

L'objectif de ce chapitre est de présenter la fabrication et les performances des transistors Ω FETs réalisés dans le cadre de cette thèse. La première partie du chapitre s'attachera aux différentes étapes technologiques nécessaires à la fabrication des transistors. Ensuite, les performances électriques relatives aux performances de base des dispositifs réalisés seront présentées. En particulier, le courant de sortie, le courant de fuite, le contrôle de la tension de seuil et de la pente sous le seuil pour des dispositifs ultra-courts seront analysés.

Sommaire du chapitre 3

1	Fabrication d'un lot de transistors Ω FETs.....	75
1.1	Description schématique des étapes de fabrication	75
1.2	Description des étapes critiques.....	76
1.2.1	Gravure de l'aileron.....	76
1.2.2	Gravure de la grille.....	77
1.3	Morphologie des transistors	78
2	Performances électriques du lot réalisé.....	79
2.1	Courant de drain en fonction de la tension de drain $I_D(V_D)$	79
2.2	Courant de sortie (I_{ON}) et courant à l'état bloqué (I_{OFF})	81
2.3	Tensions de seuil.....	83
2.3.1	Méthodes d'extraction de la tension de seuil.....	83
2.3.2	Contrôle de la tension de seuil lors de la réduction d'échelle.....	84
2.4	Pente sous le seuil et <i>DIBL</i>	85
2.4.1	Cas d'un transistor large de type 'complètement déserté'.....	85
2.4.2	Cas des transistors Ω FETs	86
3	Conclusions	88

1 Fabrication d'un lot de transistors Ω FETs

1.1 Description schématique des étapes de fabrication

Le substrat de base est une plaque SOI dont l'épaisseur t_{Si} a été réduite à environ 30 nm par oxydation puis désoxydation contrôlée (Fig. 44, étapes 1-2). En utilisant une lithographie 'deep-UV' ou 'e-beam' pour les dimensions critiques, la zone active et l'aileron de silicium sont gravés (3).

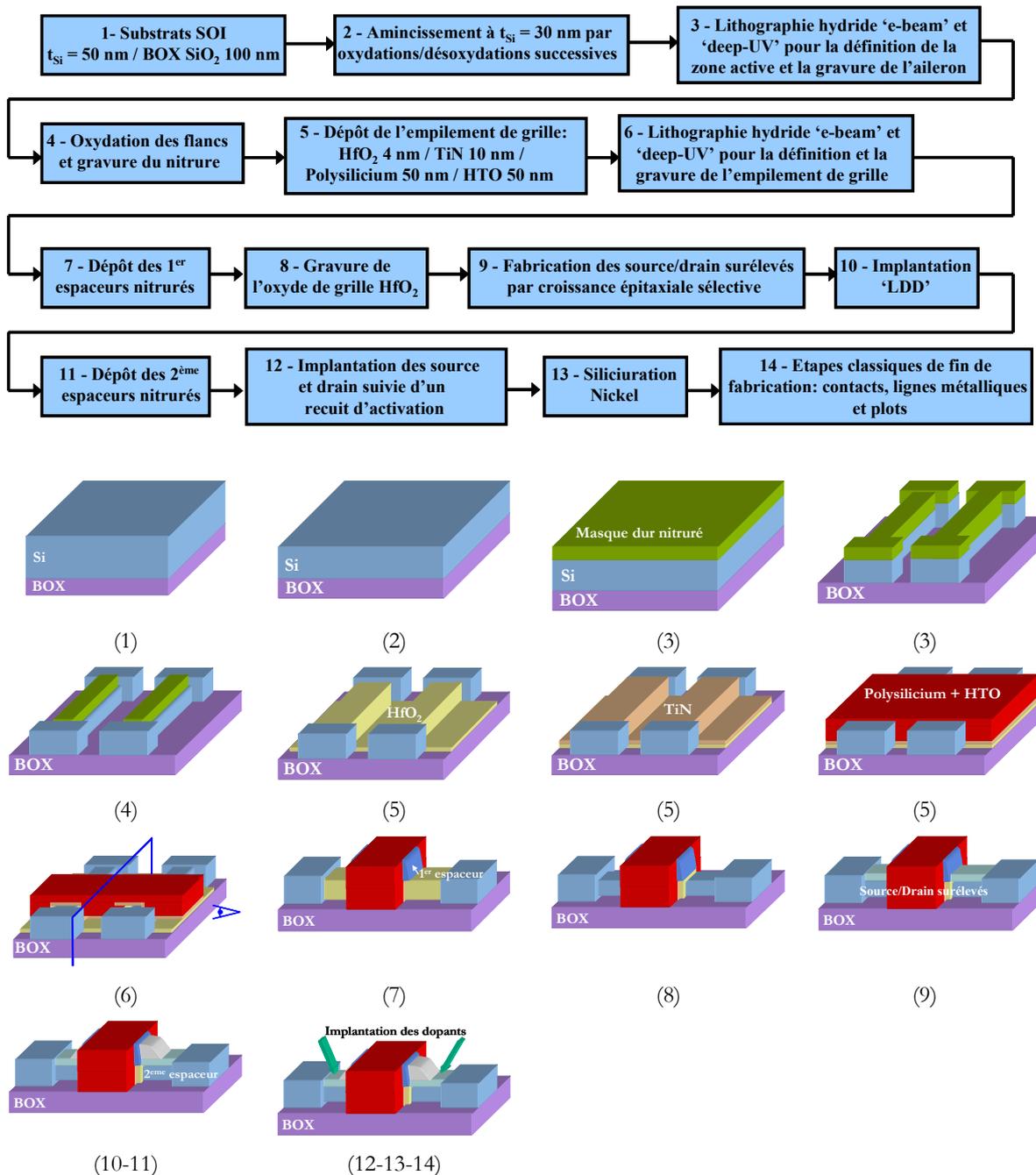


Fig. 44: Description schématique des principales étapes effectuées pour la fabrication du lot Ω FET au LETI.

Une oxydation sacrificielle (c'est-à-dire une oxydation suivie d'une désoxydation) est ensuite réalisée pour régénérer les flancs verticaux et éliminer au maximum les défauts induits par la gravure (4). Le masque dur nitruré est ensuite enlevé (configuration Triple-grille, 4). L'oxyde de grille est réalisé avec une couche d'oxyde d'hafnium HfO_2 de 4 nm déposée avec la technique *ALD* (*Atomic Layer Deposition*, i.e. dépôt par couches atomiques, 5). La grille métallique en nitrure de titane TiN est ensuite déposée par *CVD* (*Chemical Vapor Deposition*, i.e. dépôt en phase vapeur, 5). Son épaisseur est d'environ 10 nm. Une couche de polysilicium de 50 nm et un oxyde déposé à haute température (*HTO*, *High Temperature Oxide*) de 50 nm complètent l'empilement de grille (5). L'empilement de la grille est ensuite gravé (6). Des espaceurs nitrurés de 15 nm sont déposés avant la gravure de l'oxyde de grille (7-8).

Les zones de sources et drain surélevées sont fabriquées par croissance épitaxiale sélective (*SEG*, *Selective Epitaxial Growth*, 9). La structure *LDD* (*Lightly Doped Drain*, ou 'Drain Faiblement Dopé') a pour objectif de réduire le champ électrique induit par les jonctions source/substrat et drain/substrat. Elle est constituée d'une région faiblement dopée et étroite afin de limiter les effets de canaux courts (10). Un deuxième dépôt d'espaceurs nitrurés (11) est ensuite réalisé avant l'implantation des zones de source et drain (12). Le rôle de ces *offset spacer* est de repousser les zones de source et drain afin de contrôler les courants de fuite sans trop dégrader le courant de saturation. La capacité de recouvrement entre la grille et les extensions est diminuée. La longueur effective du canal est néanmoins augmentée. Un recuit rapide (de type '*spike*') est effectué, suivi d'une siliciuration au nickel (13). La fin du procédé est standard et consiste à fabriquer les contacts, les lignes de métal (1 niveau) et les plots (14).

1.2 Description des étapes critiques

1.2.1 Gravure de l'aileron

Traditionnellement, la longueur de grille est la plus petite dimension lithographique à prendre en compte dans la fabrication d'un transistor. Dans le cas d'un transistor FinFET, la largeur de l'aileron est du même ordre de grandeur ou même inférieure à la longueur de grille; sa largeur peut donc également être problématique lors de la fabrication du transistor. Pour les dimensions peu agressives, une lithographie optique '*deep UV*' a été utilisée avec une résine à amplification chimique. Les dimensions critiques du transistor, c'est-à-dire principalement les courtes longueurs et largeurs de grille, ont été réalisées avec une lithographie utilisant un faisceau électronique (technologie '*e-beam*') et une résine chimique (limitée cependant à une résolution de 40 nm). L'inconvénient de cette méthode est la très longue durée de cette étape de lithographie.

Afin d'assurer des performances électriques optimales (en particulier la pente sous le seuil, donc les fuites à l'état bloqué), il faut s'assurer d'avoir des flancs les plus verticaux possibles (voir chapitre 4, section 0). Les coupes morphologiques réalisées montrent une bonne verticalité des flancs (Fig. 45). De fortes variations de la largeur de l'aileron sont également à éviter absolument, puisqu'elles entraînent une indétermination dans les paramètres électriques du transistor. Le comportement électrostatique de la région des coins

est également une question en débat (voir chapitre 4, section 4); pour ce lot, l'oxydation sacrificielle des flancs verticaux assure un léger arrondi des coins (Fig. 45). Cette oxydation sacrificielle a également pour but de réparer les défauts générés par la gravure de l'aileron.

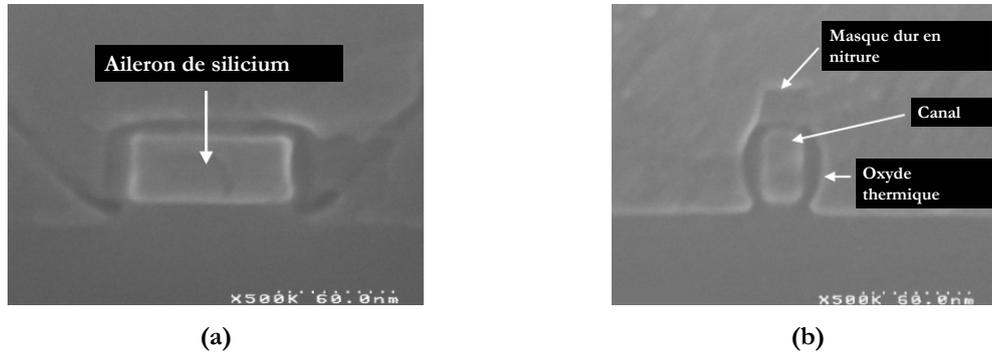


Fig. 45: Photos MEB de transistors FinFETs sans (a) et avec (b) masque dur en nitrure, montrant l'arrondi des coins généré lors de la fabrication.

La gravure de l'aileron est réalisée en deux étapes: une gravure du nitrure et de l'oxyde 'piédestal', avec ensuite la gravure du silicium avec arrêt sur l'oxyde enterré. Le masque dur est ensuite enlevé à la fin des étapes de gravure des tranchées. A cause de la consommation de l'oxyde enterré lors de la gravure des zones actives et des nettoyages effectués avant le dépôt de la grille, le canal se présente sous la forme d'un aileron de silicium posé sur un piédestal de silice (Fig. 46.a). Cette configuration porte le nom de 'ΩFET' à cause de la ressemblance avec la lettre grecque (voir chapitre 2, section 3.3).

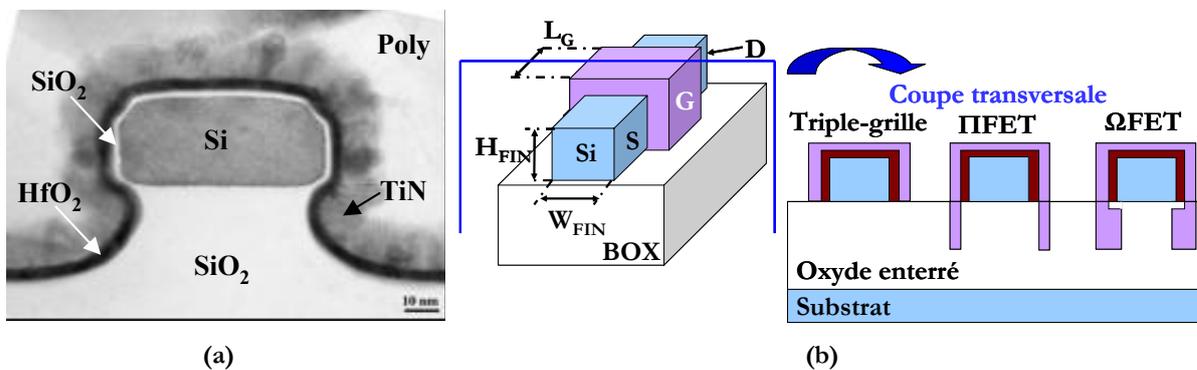


Fig. 46: Photo TEM (a) et illustration schématique (b) des types de morphologies selon le procédé de gravure utilisé.

1.2.2 Gravure de la grille

Un des problèmes majeurs pouvant se poser dans l'étape de gravure de la grille est la présence de résidus après gravure (Fig. 47). Ces résidus s'avèrent très handicapants pour les étapes ultérieures. En utilisant une combinaison de gravure sèche et de gravure humide, l'empilement de grille a été gravé sans résidus.

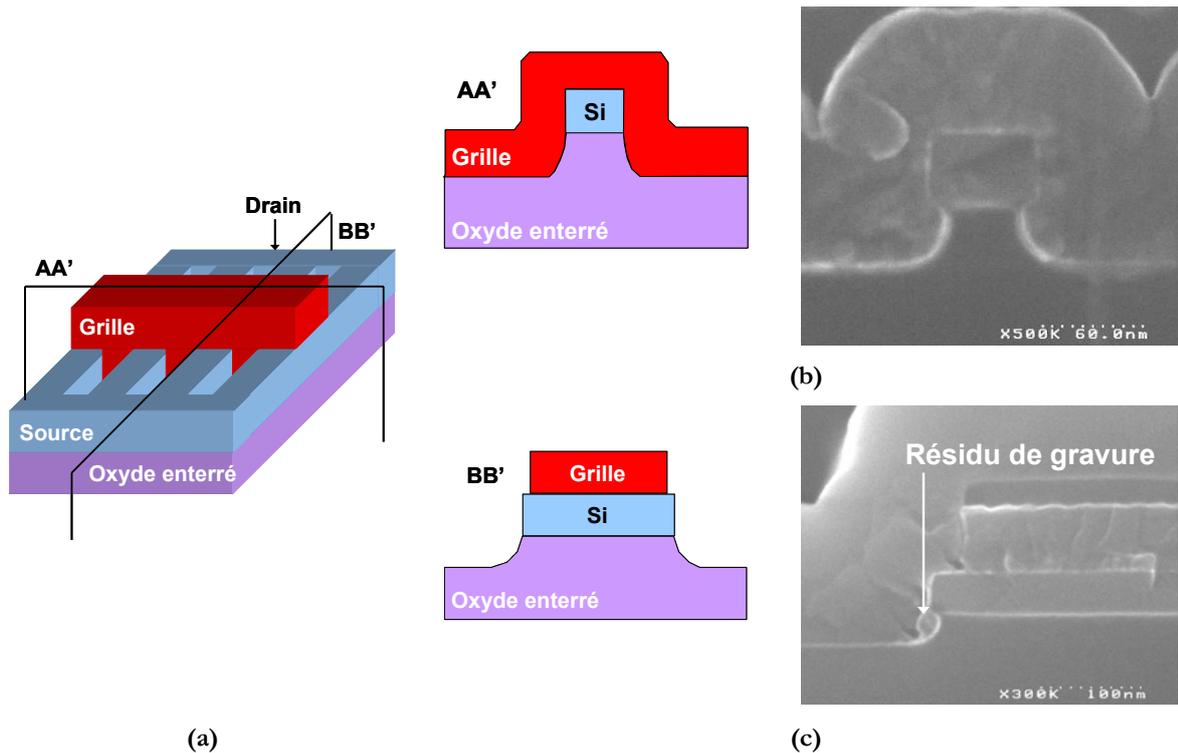


Fig. 47: Schéma d'un réseau d'ailerons après gravure de la grille (a). Coupes transversale (axe de la grille, b) et longitudinale (axe source/drain, c) associées.

1.3 Morphologie des transistors

Les clichés réalisés au microscope à balayage électronique (Fig. 48) montrent la morphologie des transistors après l'étape de gravure de la grille. On reconnaît les zones de source, drain et les deux contacts de grille (Fig. 48.a). Les réseaux d'une cinquantaine d'ailerons ont également été gravés avec succès (Fig. 48.b).

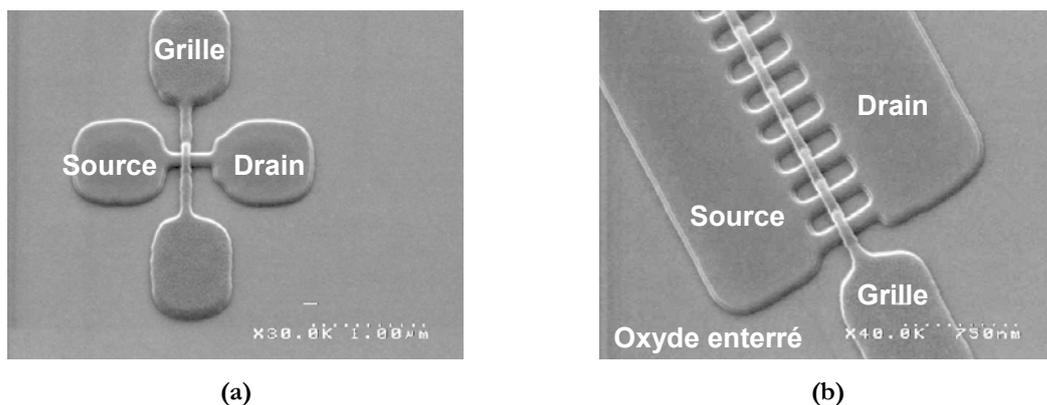


Fig. 48: Photo MEB d'un transistor isolé d'une longueur de grille de 20 nm (a) et d'un réseau d'ailerons de silicium ($L_G = 250$ nm, $W_{FIN} = 50$ nm, espace entre chaque aileron = 250 nm) après la gravure de la grille.

Les clichés obtenus au microscope électronique en transmission (Fig. 49) permettent d'explorer plus en détail la topologie des transistors. L'empilement de grille présente une bonne conformité (coupe transversale, Fig. 49.a). Il n'y a pas de gros résidus visibles entre

chaque aileron. La topologie du canal en Ω est également évidente. Les flancs sont bien verticaux, avec néanmoins la présence d'un chanfrein au sommet des canaux (insert de la Fig. 49.a). Sur la coupe longitudinale (selon l'axe source/drain, Fig. 49.b), on reconnaît l'oxyde enterré, l'aileron de silicium, les deux espaceurs nitrurés, les zones de source et drain siliciurées (NiSi) ainsi que l'empilement de grille.

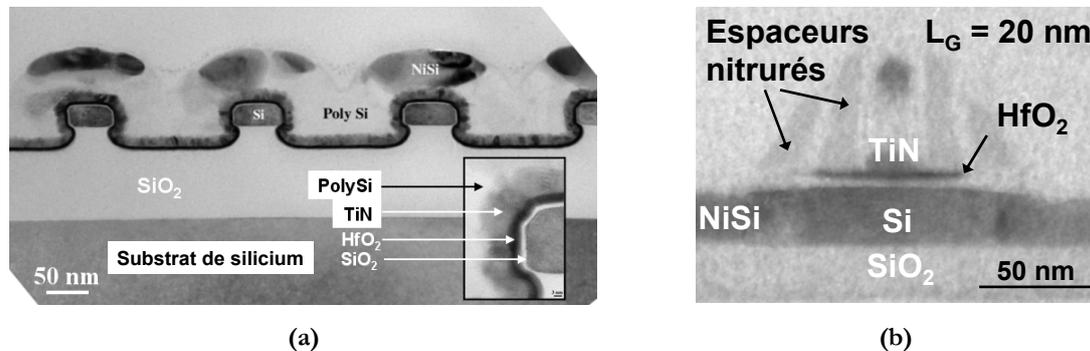


Fig. 49: Photos TEM des transistors Ω FETs. (a): Coupe transversale d'un réseau de transistors ($W_{FIN} = 60$ nm). L'insert souligne l'empilement de grille SiO_2 interfacial/ HfO_2 / TiN /Polysilicium. (b): Coupe longitudinale (axe source/drain) d'un transistor d'une longueur de grille de 20 nm. L'empilement de grille, les deux espaceurs et la siliciuration sont visibles.

Sur la coupe TEM agrandie de la Fig. 46.a, on constate qu'il n'y a pas de résidus visibles au pied des ailerons. L'oxyde interfacial (c'est à dire la couche de silice se formant entre l'oxyde et le canal lors du dépôt de l'oxyde *high-k*) est clairement visible. Son épaisseur a été mesurée à 1.3 nm. L'EOT (*Equivalent Oxide Thickness*, ou épaisseur équivalente d'oxyde) a été également mesuré électriquement à 1.95 nm. Au final, on dispose de transistors dont la longueur de grille varie de 10 μm à 10 nm, et dont la largeur peut varier de 25 μm à 40 nm. Les plus petites tailles correspondent au meilleur niveau de l'état de l'art international. Des dispositifs en réseaux sont également disponibles, ainsi que des transistors orientés à 45° par rapport à l'encoche (le *notch*). Nous allons maintenant examiner les caractéristiques de base observées pour ces transistors.

2 Performances électriques du lot réalisé

2.1 Courant de drain en fonction de la tension de drain $I_D(V_D)$

Classiquement, les caractéristiques $I_D(V_D)$ d'un transistor montrent un courant nul en dessous du seuil, qui croit ensuite régulièrement quand la tension de grille augmente. En augmentant V_D , le régime de saturation s'établit lorsque la tension de drain est supérieure à la tension de pincement. Pour de fortes polarisations de drain, la longueur effective du canal diminue et en conséquence le courant de drain augmente graduellement dans les canaux courts (effet de modulation de longueur de canal). De plus, le courant de drain augmente considérablement si on est en condition de perçage puisque des porteurs majoritaires issus de la source sont injectés dans la zone de désertion puis collectés dans le drain. Si la grille perd le

contrôle du canal, on ne peut plus couper le transistor et le courant de drain n'est pas nul pour une tension de grille nulle. Aux faibles longueurs de grille, l'examen de la courbe $I_D(V_D)$ permet donc de caractériser le contrôle électrostatique de la grille sur le canal.

La Fig. 50 montre les performances de sortie des transistors en fonction du courant de drain pour des transistors NMOS à faible longueur de grille ($L_G = 30$ et 10 nm). La normalisation est faite par la largeur effective du canal, i.e. $W_{FIN} + 2H_{FIN}$. Pour un transistor isolé (Fig. 50.a) d'une longueur de grille de 30 nm, le contrôle du perçage est correct puisque le courant de drain à V_G nul est très faible et que la conductance $\frac{\partial I_D}{\partial V_D}$ reste raisonnable à fort V_D . Les conclusions sont les mêmes pour un transistor en réseau (Fig. 50.b), même avec une longueur de grille de 10 nm!

Les mesures effectuées sur un transistor PMOS court ($L_G = 50$ nm) avec la même largeur d'aileron ($W_{FIN} = 40$ nm) sont montrées sur la Fig. 51. De même que dans le cas des NMOS, le contrôle du perçage est très correct.

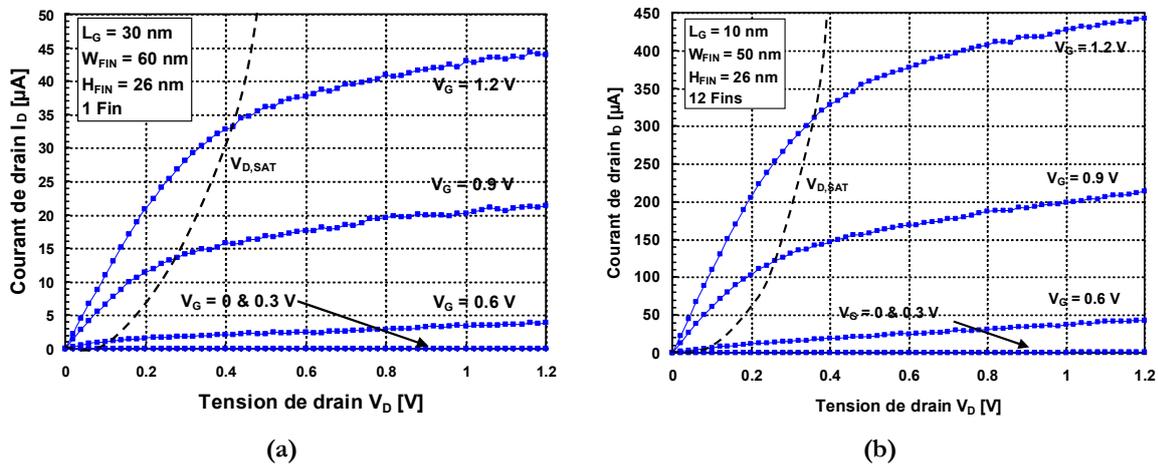


Fig. 50: (a): Courant de drain en fonction de la tension de drain pour un transistor NMOS court isolé ($L_G = 30$ nm). (b): Courant de drain en fonction de la tension de drain pour un transistor NMOS court en réseau ($L_G = 10$ nm).

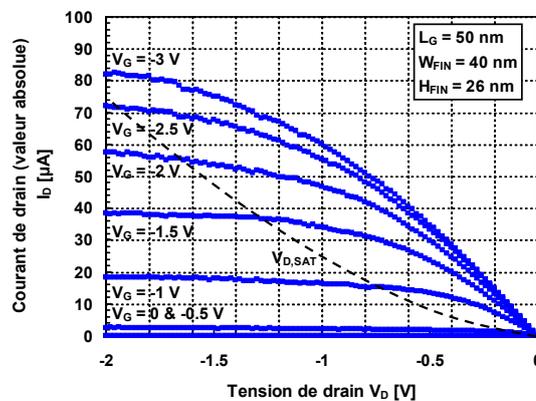


Fig. 51: Courant de drain en fonction de la tension de drain pour un transistor PMOS isolé court ($L_G = 50$ nm).

2.2 Courant de sortie (I_{ON}) et courant à l'état bloqué (I_{OFF})

Un autre aspect du contrôle des effets de canaux courts réside dans la mesure du courant de drain en fonction de la tension de grille. Cette mesure permet d'une part d'examiner la dégradation de la pente sous le seuil et le décalage de la tension de seuil à fort V_D (DIBL) quand la longueur de grille est réduite.

La Fig. 52.a présente les résultats obtenus pour des transistors NMOS et PMOS ayant respectivement une longueur de grille de 30 et 50 nm. Au vu de la longueur de grille, les courants obtenus et le contrôle des effets de canal court sont bons. La tension de seuil vaut environ 0.5 V pour le transistor NMOS et -0.2 V pour le PMOS. La pente sous le seuil et le DIBL mesurés sont respectivement de l'ordre de 80 mV/décade et de 100 mV/V. Pour un transistor NMOS ultra-court ($L_G = 10$ nm, Fig. 52.b) ayant une largeur d'aileron W_{FIN} de 50 nm, la pente sous le seuil et le DIBL sont dégradés mais le transistor reste tout à fait opérationnel.

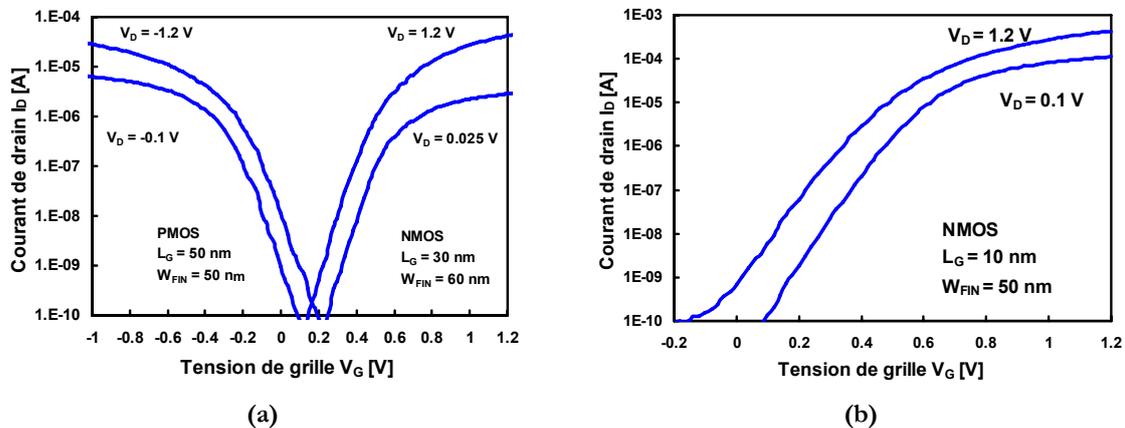


Fig. 52: Courant de drain en fonction de la tension de grille pour des transistors NMOS ($L_G = 30$ nm (a) et 10 nm (b)) et PMOS ($L_G = 50$ nm).

On constate également sur la Fig. 52 que la géométrie en Ω du dispositif permet de garantir un excellent contrôle du courant à l'état bloqué. D'autre part, l'oxyde de grille utilisé (oxyde *high-k* HfO_2) est connu pour présenter de faibles courants de fuite mais également une mobilité considérablement dégradée par rapport à l'oxyde de silicium [Chang'04]. Par rapport à la bibliographie, on attend donc un très bon courant à l'état bloqué avec des courants de sortie un peu en retrait. Le tableau de la Fig. 53 est une synthèse bibliographique des performances obtenues pour des composants avancés (silicium massif, FinFET, Triple-grille, Ω FET). Dans notre cas, le courant I_{ON} est effectivement légèrement plus faible que ceux déjà publiés mais parallèlement le courant I_{OFF} est bien meilleur.

Ref	ITRS'05	[Yu'01] <i>bulk</i>	[Xiang'01] <i>bulk</i>	[Doyle'03] Triple- grille	[Yang'02] Ω FET	[Yu'02] FinFET	[Kedzierski'02] FinFET	[Collaert'05] Ω FET	[Ce travail] Ω FET
L_G [nm]	45	35	40	60	25	10	100	40	30 (N) 50 (P)
W [nm]								10	60
H [nm]								60	25
Type de grille		PolySi	PolySi	PolySi	PolySi	PolySi	NiSi	TiN (MOCVD)	TiN
Type d'oxyde		Nitruure/ oxynitruure	Nitruure/ oxynitruure	SiO ₂	SiO ₂	Nitruure	SiO ₂	HfO ₂	HfO ₂
EOT [nm]	1.9	0.7/0.12	0.9	1.3	1.8	1.7	1.6	1.9	1.9
V _{DD} [V]	1.2	0.85	0.9	1.3	1.0	1.2	1.25	1.3	1.2
NMOS									
N I _{ON} [μ A/ μ m]	519	580	657	1140	587	446 ^a	814	~350	370
N I _{OFF} [nA/ μ m]	0.01	50	100	69	3.3	n.c ^b	100	~1	0.05
NMOS I _{ON} /I _{OFF}	5.19 10 ⁷	1.16 10 ⁴	6.57 10 ³	1.6 10 ⁴	1.8 10 ⁵		8 10 ³	~3.5 10 ⁵	7.4 10 ⁶
PMOS									
P I _{ON} [μ A/ μ m]	260	250	290	520	320	356 ^a	500	~350	325
P I _{OFF} [nA/ μ m]	0.005	90	100	24	0.16	n.c ^b	17	~0.1	1.4
PMOS I _{ON} /I _{OFF}	5.19 10 ⁷	2.7 10 ³	2.9 10 ³	2.2 10 ⁴	2 10 ⁶		3 10 ⁴	3.5 10 ⁶	2.3 10 ⁵

^a: à $V_G - V_T = 1$ V

^b: dispositif normalement à l'état ouvert

Fig. 53: Synthèse bibliographique des structures et performances obtenues pour des composants avancés. Les spécifications requises par l'ITRS pour les applications basse consommation du nœud 65 nm sont également montrées. La normalisation est faite avec la largeur effective de l'aileron, c'est-à-dire $2H_{FIN}$ pour des FinFETs et $W_{FIN} + 2H_{FIN}$ pour des dispositifs Triple-grille et Ω FET.

Le rapport I_{ON}/I_{OFF} en fonction de la longueur de grille L_G (Fig. 54) montre que le compromis "état passant – état bloqué" obtenu avec une architecture de type Ω FET [Jahan'05] est excellent (Fig. 53), et révèle une très bonne tenue des dispositifs quand la longueur de grille est réduite. Selon ce critère, nos résultats sont même généralement légèrement meilleurs que ceux déjà publiés dans la bibliographie.

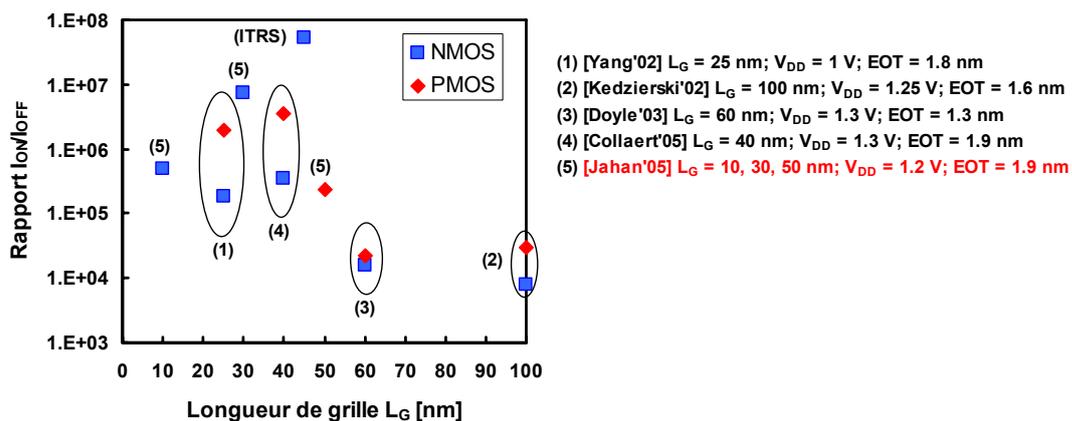


Fig. 54: Rapport I_{ON}/I_{OFF} en fonction de la longueur de grille pour les dispositifs Ω FET réalisés et pour quelques dispositifs Triple-grille et Ω FET publiés dans la bibliographie.

2.3 Tensions de seuil

2.3.1 Méthodes d'extraction de la tension de seuil

Il existe beaucoup de méthodes d'extraction de la tension de seuil (une description exhaustive peut être trouvée dans [Rochette'04]). Ces méthodes dépendent de la définition choisie, et sont donc susceptibles de donner des résultats différents. La plus utilisée car la plus simple consiste à mesurer la tension de grille pour laquelle le courant de drain est égal à $0.1 \mu\text{A}$ (à normaliser par les dimensions du dispositif, i.e. $(W_{\text{FIN}} + 2H_{\text{FIN}}) / L_G$). C'est la méthode dite 'à courant constant' (Fig. 55.a).

La méthode du ' V_T extrapolé' consiste à approximer le fonctionnement d'un transistor en régime linéaire au dessus du seuil à une droite. Ainsi, en traçant la tangente à la courbe $I_D(V_G)$ au maximum de transconductance (i.e. la dérivée première), l'ordonnée à l'origine de la tangente permet de mesurer une tension de seuil (Fig. 55.b).

La méthode de 'la dérivée seconde' consiste à prendre comme tension de seuil le maximum de la dérivée seconde (Fig. 55.b). Cette méthode permet de s'affranchir de l'effet des résistances séries. Enfin, certaines fonctions définies pour l'étude des caractéristiques électriques permettent également de mesurer une tension de seuil. Les fonctions à partir des dérivées première et seconde de l'inverse du courant de drain (voir chap. 5 section 3 et Fig. 55.c) sont les plus couramment utilisées. Ces courbes sont linéaires en V_G au dessus du seuil et leur extrapolation à V_G nul fournit une tension de seuil.

Les méthodes utilisées dans ce chapitre sont les méthode du ' V_T extrapolé' et la méthode 'à courant constant'.

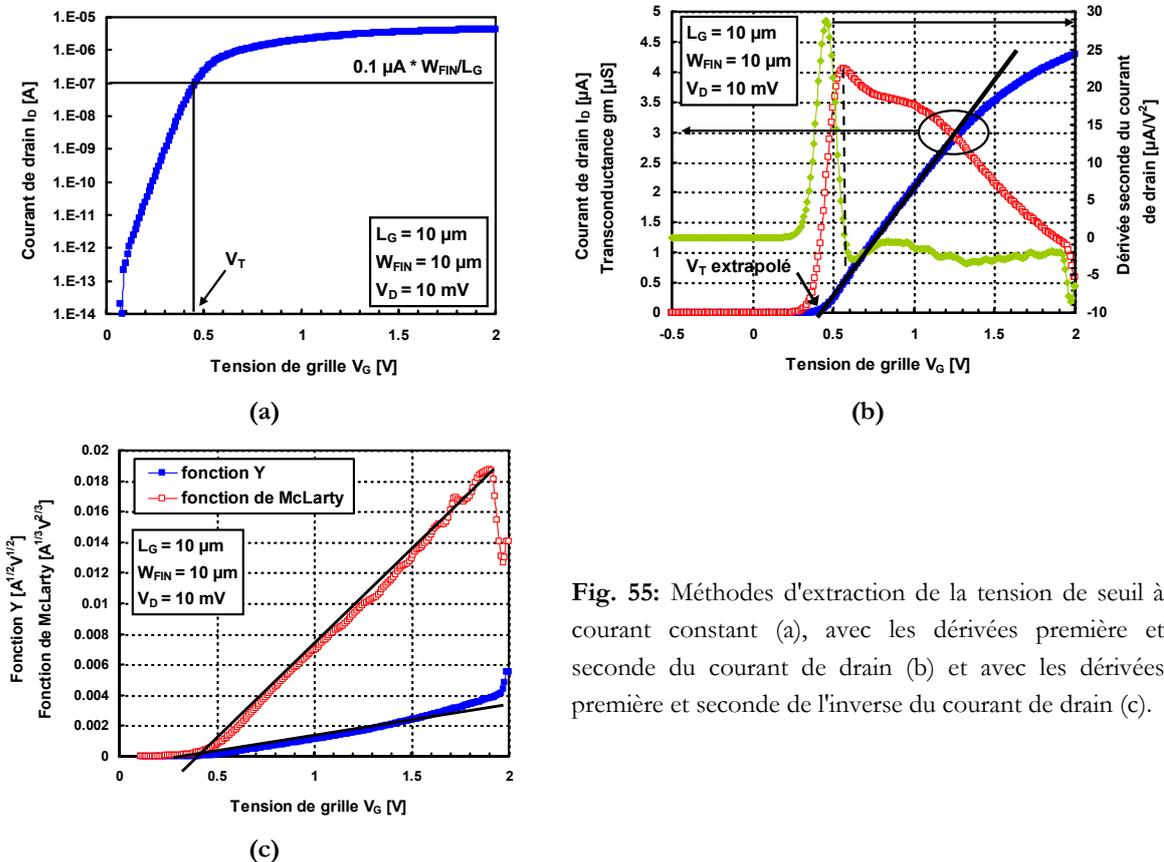


Fig. 55: Méthodes d'extraction de la tension de seuil à courant constant (a), avec les dérivées première et seconde du courant de drain (b) et avec les dérivées première et seconde de l'inverse du courant de drain (c).

2.3.2 Contrôle de la tension de seuil lors de la réduction d'échelle

En améliorant le contrôle du potentiel dans le canal, une architecture multigrille va également faire évoluer la tension de seuil. La tension de seuil peut être définie en utilisant le critère capacitif qui définit le passage entre inversion faible et inversion forte [Tsividis'99]:

$$C_{INV} = C_{OX} + C_{DEP} \quad \text{Eq. (13)}$$

où C_{INV} est la capacité associée à la charge d'inversion, C_{DEP} celle associée à la charge de désertion et C_{OX} la capacité d'oxyde.

Cette approche est valable pour tous types de transistors. En partant de cette relation, [Poiroux'05] a montré que la tension de seuil d'un transistor 'complètement déserté' faiblement dopé pouvait se mettre sous la forme:

$$V_{th} = \Delta\Phi_M + \frac{kT}{q} \ln\left(\frac{kT}{q} \frac{\eta C_{OX}}{q n_i t_{Si}}\right) + \frac{\hbar^2 \pi^2}{2.m^* t_{Si}^2} \quad \text{Eq. (14)}$$

où $\Delta\Phi_M$ est la différence entre le travail de sortie du matériau de grille et du silicium intrinsèque, n_i la concentration de porteurs dans le silicium intrinsèque, t_{Si} l'épaisseur du film de silicium entre les grilles et m^* la masse de confinement des porteurs dans la direction transverse. η est un paramètre capacitif valant 1 pour un transistor simple grille et 2 pour les transistors double grille. Le troisième terme de l'équation traduit le confinement induit par le puits de potentiel formé par la couche du canal située entre les deux oxydes. Ce terme ne devient important qu'en dessous de $t_{Si} = 5$ nm.

En réduisant la largeur d'aileron d'un transistor Ω FET, on passe ainsi d'une configuration 'simple-grille horizontale' à une configuration 'double-grille verticale'. Ce passage conduit donc d'une part à avoir le paramètre η tendant vers 2 et d'autre part à réduire le paramètre ' t_{Si} ' (épaisseur du film ou largeur d'aileron suivant la configuration). Par conséquent, la réduction de la largeur de l'aileron s'accompagne d'une remontée de la tension de seuil (Fig. 56.a).

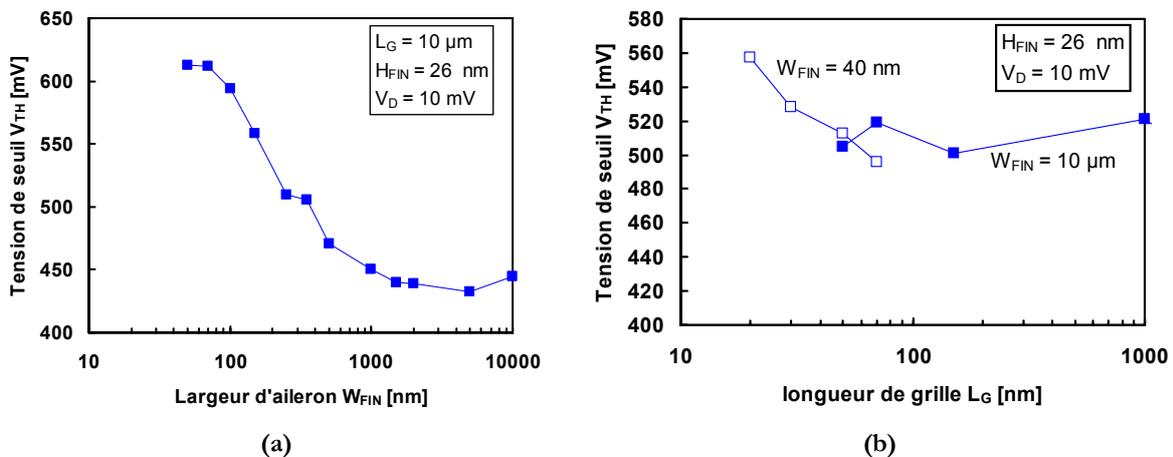


Fig. 56: Effet de la réduction de la largeur d'aileron pour un transistor long (a) et de la réduction de la longueur de grille (b). Les tensions de seuil ont été extraites avec la méthode du maximum de transconductance.

De même, quand la longueur de grille diminue, on observe un très bon contrôle des effets de canaux courts pour un NMOS large jusqu'à une longueur de grille de 40 nm (carrés pleins, Fig. 56.b). En utilisant une largeur d'aileron de 40 nm (carrés ouverts, Fig. 56.b), le contrôle de la capacité d'oxyde reste important; on observe une légère remontée de la tension de seuil en régime linéaire (environ 60 mV) en diminuant la longueur de grille de 70 à 20 nm. On note la même tendance pour des transistors dont la largeur de grille est égale à 50 et 80 nm ($V_D = 20$ mV, Fig. 57.a).

La tension de seuil en régime de saturation est plus sensible à la réduction de la longueur de grille à cause de l'abaissement de la barrière d'injection par le potentiel du drain. La tension de seuil en saturation (Fig. 57.a) présente une particularité puisque son évolution en fonction de la longueur de grille semble dépendante de la largeur de l'aileron. Une mesure pour différentes largeurs d'ailetons a été effectuée ($V_D = 1.2$ V, Fig. 57.b); malgré l'imprécision de ce type d'extraction, des tendances se dégagent. On observe ainsi une forte réduction de la tension de seuil avec la longueur de grille pour les transistors les plus larges ($W_{FIN} = 80$ nm). Cependant, la réduction de la largeur de l'aileron revient à écranter l'influence du drain par le couplage électrostatique latéral. Ainsi, des transistors dont la largeur d'aileron vaut 40 ou 50 nm vont être assez peu sensibles à la réduction des dimensions et présentent un comportement similaire celui de la tension de seuil en régime linéaire. C'est un bon argument pour la viabilité des transistors Ω FET.

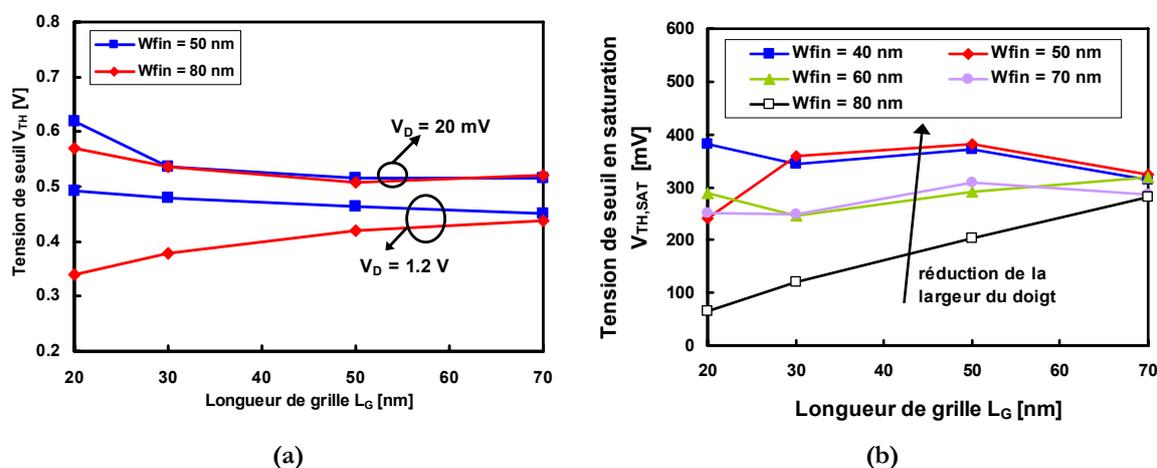


Fig. 57: Tension de seuil extraite en régime linéaire et en régime saturé en fonction de la longueur de grille (la largeur d'aileron W_{FIN} vaut 50 et 80 nm) (a). Tension de seuil extraite en régime saturé ($V_D = 1.2$ V) en fonction de la longueur de grille pour différentes largeurs d'aileton (b). Les tensions de seuil ont été extraites avec la méthode à courant constant.

2.4 Pente sous le seuil et *DIBL*

2.4.1 Cas d'un transistor large de type 'complètement déserté'

Pour un transistor long la pente mesurée à 300 K est d'environ 65 mV/décade, proche de la valeur idéale (Fig. 58). Ceci illustre la qualité des matériaux SOI et de la technologie utilisée. En réduisant la longueur de grille, l'influence accrue des régions de désertion aux

jonctions source et drain conduit à une augmentation de la pente sous le seuil. On retrouve ici classiquement l'effet de canal court, dû au fait qu'aux faibles longueurs de grilles la proximité des zones de source et drain rend le contrôle du canal par la grille de plus en plus ardu. On observe ainsi une pente sous le seuil pouvant atteindre 150 mV/décade pour un transistor de 50 nm de longueur de grille. Dans ce cas, l'épaisseur de film ($H_{\text{FIN}} = 26$ nm) n'est pas suffisante pour avoir de bonnes performances.

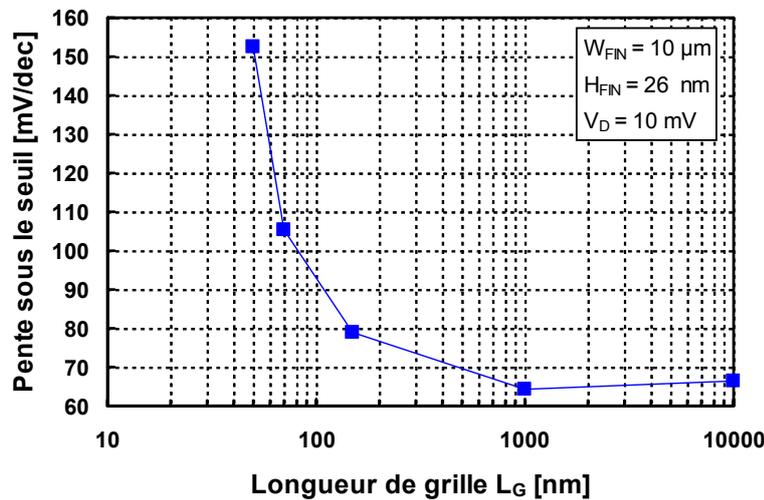


Fig. 58: Pente sous le seuil en fonction de la longueur de grille pour un transistor de type 'complètement déserté'. L'épaisseur H_{FIN} et la largeur du transistor W_{FIN} sont respectivement de 26 nm et 10 μm .

2.4.2 Cas des transistors ΩFET s

Dans le cas d'un transistor ΩFET , une faible largeur d'aileron induit un fort couplage entre les deux faces latérales de la grille. Par conséquent, la capacité de grille augmente et permet de contrebalancer plus efficacement la capacité de désertion des jonctions (partage de charge). Pour une longueur de grille donnée la pente sous le seuil diminue clairement avec la largeur de l'aileron de silicium. On observe ce phénomène pour toutes les longueurs de grille étudiées, avec un effet plus marqué pour les petites longueurs de grille. Ainsi, entre la pente obtenue pour un aileron de 40 nm et celle obtenue pour un aileron de 80 nm, la dégradation de la pente sous le seuil varie seulement de 10 mV/décade avec les transistors de longueur de grille 70 nm et de 20 mV/décade avec les transistors de longueur de grille 20 nm (Fig. 59.a).

Plusieurs constatations peuvent être faites à partir de la courbe Fig. 59.a. En particulier, la relation d'échelle pour l'optimisation des transistors de type 'Triple-grille' ($L_G \sim W_{\text{FIN}} \sim H_{\text{FIN}}$, voir chap. 4) est ici bien respectée. La hauteur de l'aileron H_{FIN} est ici toujours égale à 26 nm. Des caractéristiques presque idéales (pente proche de 60 mV/décade) sont obtenues pour une longueur de grille de 70 nm. Pour une longueur de grille de 50 nm, tant que $W_{\text{FIN}} < L_G$, les caractéristiques sont très bonnes. Pour les très courtes longueurs de grilles ($L_G = 20$ et 30 nm), il n'est pas possible de tester la validité de relations d'échelle sur les dispositifs. Cependant, il est intéressant de noter que bien que dégradées les performances de pente sous le seuil n'en

restent pas moins convenables (pour $L_G = 20$ nm, $SS = 75$ mV/décade avec $W_{FIN} = 40$ nm et $SS = 100$ mV/décade avec $W_{FIN} = 80$ nm).

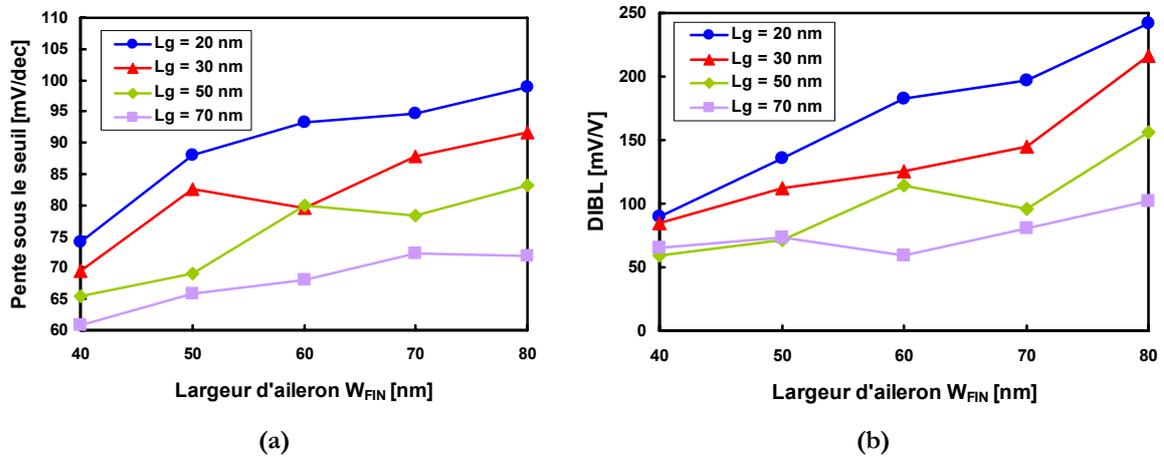


Fig. 59: Pente sous le seuil à $V_D = 50$ mV (a) et DIBL $\Delta V_T / \Delta V_D$ (b) en fonction de la largeur d'aileron W_{FIN} obtenus pour des dispositifs avec des longueurs de grille L_G de 70 nm (carrés), 50 nm (losanges), 30 nm (triangles) et 20 nm (cercles).

L'effet DIBL, dont le mécanisme a été exposé au chapitre 1, permet quant à lui d'étudier la dégradation des caractéristiques sous le seuil quand la tension de drain V_D augmente. En utilisant les mêmes mesures que précédemment (Fig. 59.b), on constate que le DIBL augmente pour une longueur de grille donnée avec la largeur de l'aileron. Pour une largeur d'aileron donnée, il augmente quand la longueur de grille diminue. Ce sont les mêmes tendances que pour les pentes sous le seuil (Fig. 59.a). Même à fort V_D , on continue donc de garder un bon contrôle électrostatique du canal par la grille en utilisant des canaux étroits.

3 Conclusions

Ce chapitre a présenté les étapes de fabrication d'un lot Ω FET. Les caractéristiques du procédé ont été détaillées et les étapes critiques mises en lumière.

Les transistors sont parfaitement fonctionnels jusqu'à une longueur de grille de 10 nm. Les transistors Ω FET montrent des performances de courant de sortie très honorables compte tenu de l'empilement '*high-k* – grille métallique' utilisé. De plus, les courants de fuite à l'état bloqué sont très faibles. Ceci est dû à la structure multigrille Ω FET qui assure un excellent contrôle du canal par les grilles latérales. On observe ainsi un excellent rapport courant de sortie/courant de fuite à l'état bloqué par rapport aux autres types de structures prospectives reportées dans la bibliographie.

De même, la tension de seuil est bien contrôlée par les grilles latérales. Ainsi, la variation de tension de seuil en réduisant la longueur de grille, qui est traditionnellement un problème important, reste assez faible dans le cas des transistors Ω FET. L'effet bénéfique du canal étroit a été montré, ainsi que le contrôle des effets de canaux courts.

La pente sous le seuil et le *DIBL* sont très bien contrôlés jusqu'à de faibles longueurs de grille à condition de respecter les règles d'échelle. Dans le cas présent, il s'agit surtout de pouvoir réduire la largeur des ailerons conjointement avec la longueur de grille.

Le prochain chapitre s'attachera à dégager plus spécifiquement les intérêts et les mécanismes physiques des architectures multigrilles, entrevus dans ce chapitre.

Références du chapitre 3

- [Chang'04] L. Chang, M. Jeong, M. Yang, "CMOS Circuit Performance Enhancement by Surface Orientation Optimization", *Electron Devices, IEEE Transactions on*, vol. 51 no. 10, pp. 1621-1627, 2004.
- [Collaert'05] N. Collaert, M. Demand, I. Ferain, J. Lisoni, R. Singanamalla, P. Zimmerman, Y.S. Yim, T. Schram, G. Mannaert, M. Goodwin, J.C. Hooker, F. Neuilly, M.C. Kim, K. De Meyer, S. De Gendt, W. Boullart, M. Jurczak, S. Biesemans, " Tall triple-gate devices with TiN/HfO₂/ gate stack", *Digest of Technical Papers, 2005 Symposium on VLSI Technology*, pp. 108-109, 2005.
- [Doyle'03] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Harelend, B. Jin, J. Kavalieros, T. Linton, R. Rios, R. Chau, "Tri-Gate Fully-Depleted CMOS Transistors: Fabrication, Design and Layout", *Digest of Technical Papers, 2003 Symposium on VLSI Technology*, pp. 133-134, 2003.
- [ITRS'05] International Technology Roadmap for Semiconductors, Process Integration Devices, and Structures, Low Standby Power Requirements, p. 17, 2005.
- [Jahan'05] C. Jahan, O. Faynot, M. Cassé, R. Ritzenthaler, L. Brévard, L. Tosti, X. Garros, C. Vizioz, F. Allain, A.M. Papon, H. Dansas, F. Martin, M. Vinet, B. Guillaumot, A. Toffoli, B. Giffard and S. Deleonibus, "10nm Ω FETs transistors with TiN metal gate and HfO₂", *Digest of Technical Papers, 2005 Symposium on VLSI Technology*, pp. 112-113, 2005.
- [Kedzierski'02] J. Kedzierski, E. Nowak, T. Kanarsky, Y. Zhang, D. Boyd, R. Carruthers, C. Cabral, R. Amos, C. Lavoie, R. Roy, J. Newbury, E. Sullivan, J. Benedict, P. Saunders, K. Wong, D. Canaperi, M. Krishnan, K.-L. Lee, B. A. Rainey, D. Fried, P. Cottrell, H.-S. P. Wong, M. Jeong, W. Haensch, "Metal-gate FinFET and fully depleted SOI devices using total gate silicidation", *IEDM'02 Technical Digest*, pp. 247-250, 2002.
- [Poiroux'05] T. Poiroux, M. Vinet, O. Faynot, J. Widiez, J. Lolivier, T. Ernst, B. Previtali, S. Deleonibus, "Multiple gate devices: advantages and challenges", *Microelectronic Engineering*, vol. 80, pp. 378-385, 2005.
- [Rochette'04] F. Rochette, "Caractérisation et simulation électriques de transistors MOS SOI complètement désertés", rapport de stage DEA, CEA-LETI/INSA Lyon, 2004.

- [Sze'81] S. M. Sze, "Physics of Semiconductor Devices (2nd ed.)", *John Wiley and Sons (WIE)*. ISBN 0-47-105661-8, p. 477, 1981.
- [Tsvividis'99] Y.P. Tsvividis, "Operation and modelling of the MOS transistor", ed. McGraw-Hill, 1999.
- [Xiang'01] Q. Xiang, B. Yu, H. Wang, M.-R. Lin, "High Performance Sub-50nm CMOS with Advanced Gate Stack", *Digest of Technical Papers, 2001 Symposium on VLSI Technology*, pp. 23-24, 2001.
- [Yang'02] F.-L. Yang, H.-Y. Chen, F.C. Chen, C.-C. Huang, C.-Y. Chang, H.-K. Chiu, C.-C. Lee, C.-C. Chen, H.-T. Huang, C.-J. Chen, H.-J. Tao, Y.-C. Yeo, M.-S. Liang, C. Hu, "25 nm CMOS Omega FETs", *IEDM'02 Technical Digest*, pp. 255-258, 2002.
- [Yu'01] B. Yu, H. Wang, Q. Xiang, J.X. An, J. Jeon, M.-R. Lin, "Scaling Towards 35nm Gate Length CMOS", *Digest of Technical Papers, 2001 Symposium on VLSI Technology*, pp. 9-10, 2001.
- [Yu'02] B. Yu, L. Chang, S. Ahmed, H. Wang, S. Bell, C.-Y. Yang, C. Tabery, C. Ho, Qi. Xiang, T.-J. King, J. Bokor, C. Hu, M.-R. Lin, D. Kyser, "FinFET Scaling to 10nm Gate Length", *IEDM'02 Technical Digest*, pp. 251-254, 2002.

Chapitre 4 :

Analyse des mécanismes physiques spécifiques aux transistors multigrilles: aspects dimensionnels et capacitifs, modélisation des effets de couplage

Dans ce chapitre, les aspects propres aux structures Triple-grille sont étudiés. Le contrôle des effets de canaux courts par les grilles latérales est analysé, ainsi que les dimensions à respecter (règles d'échelle). Le couplage géométrique dans les zones de coin est également mis en lumière. Il est montré que son influence sur le décalage de la tension de seuil est faible, à moins d'avoir de très forts dopages. Nous démontrons que le couplage latéral permet de supprimer certains effets observés dans les structures planaires 'complètement désertées'. L'affinement de l'aile de silicium permet d'annuler l'influence de la face arrière, conduisant à des dispositifs naturellement très résistants aux radiations. Nous proposons également un modèle original pour l'effet de couplage du drain à travers l'oxyde enterré. Cet effet est très bien contrôlé pour les structures Triple-grille.

Sommaire du chapitre 4

1	Introduction	95
2	Modélisation du fonctionnement d'un transistor	95
	2.1 Equation classiques	95
	2.2 Effets quantiques	96
	2.2.1 Confinement quantique dans un semiconducteur.....	96
	2.2.2 Approximation du puit triangulaire	97
	2.2.3 Simulation des effets quantiques.....	98
3	Influence de la structure des transistors sur leurs performances électriques	99
	3.1 Règles d'échelle	99
	3.2 Impact du nombre de grilles.....	102
	3.3 Impact de la non-verticalité des flancs.....	104
	3.4 Effets des résistance d'accès et des contacts.....	105
	3.5 Conclusions sur le contrôle des effets de canaux courts.....	106
4	Effets tridimensionnels dans les structures Triple-grille: les effets de coins.....	107
	4.1 Aspects classiques.....	107
	4.1.1 Impact sur la tension de seuil	107
	4.1.2 Effet de l'arrondi de coin	110
	4.2 Impact de la prise en compte des effets quantiques	111
5	Couplages d'interfaces dans les Ω FETs.....	113
	5.1 Cas des transistors larges	113
	5.1.1 Modèle unidimensionnel de Lim et Fossum.....	113
	5.1.2 Mesures expérimentales sur des transistors larges	114
	5.2 Cas des transistors étroits.....	119
	5.2.1 Couplage bidimensionnel	119
	5.2.2 Mesures expérimentales sur des transistors Ω FET étroits	122
	5.2.3 Application: la résistance aux radiations	126
6	DIVSB (<i>Drain Induced Virtual Substrate Biasing</i>)	128
	6.1 Effet DIVSB dans les structures Triple-grille.....	128
	6.1.1 Définition.....	128
	6.1.2 Mesure expérimentale	128
	6.1.3 Simulations numériques: effet de la largeur de la grille	129
	6.2 Modélisation analytique du DIVSB.....	133
	6.2.1 Cas bidimensionnel: transistor 'complètement déserté'.....	133
	6.2.2 Cas tridimensionnel: transistors Triple-grilles.....	135
	6.3 Discussion.....	139
7	Conclusions	141
	Références du chapitre 4	142

1 Introduction

Le caractère novateur des structures multigrilles soulève un certain nombre de questions. Dans les transistors Triple-grille, le comportement électrostatique des zones de coin est ainsi un effet à étudier. Les effets induits par le fort couplage latéral sont également à même de causer de nouveaux phénomènes, ou de désactiver certains phénomènes observés dans les transistors simple-grille. Ce chapitre, s'appuyant largement sur des mesures électriques et des simulations numériques, se propose d'étudier ces phénomènes.

Après être revenu sur les lois d'échelle, les effets de coins sont analysés. Les effets de couplage latéraux, et notamment leur influence sur le couplage vertical vu par la grille avant, seront présentés. Le DIVSB est un couplage drain/canal propre aux structures sur SOI. Après avoir mis en lumière les différents effets inhérents aux structures Triple-grille, une modélisation analytique de ce phénomène est conduite.

2 Modélisation du fonctionnement d'un transistor

2.1 Equations classiques

La physique des semi-conducteurs se formalise en termes d'équations aux dérivées partielles, d'ordres et de complexités différentes. Les trois équations de base sont l'équation de Poisson, l'équation de transport et l'équation de continuité. Les coefficients et les conditions aux limites de ces équations (tels que la mobilité, le taux de génération-recombinaison, les paramètres du matériau, les conditions limites sur les contacts et aux interfaces) dépendent de la microphysique, de la structure du dispositif et de la tension appliquée.

L'équation de Poisson est fondamentale, puisque c'est la méthode habituelle pour trouver la distribution de potentiel électrique à partir de la distribution des charges:

$$\Delta\varphi = \frac{\rho}{\varepsilon} \quad \text{Eq. (15)}$$

avec φ le potentiel électrostatique, ρ la densité de charges et ε la permittivité diélectrique du matériau

Par définition, la mobilité $\mu_{e,h}$ relie la vitesse d'un porteur (électron ou trou) avec le champ électrique qui lui est appliqué (voir chapitre 5 pour une définition plus complète):

$$\begin{aligned} \vec{v}_e &= -\mu_e \vec{E} \\ \vec{v}_h &= +\mu_h \vec{E} \end{aligned} \quad \text{Eq. (16)}$$

Le courant total de conduction s'écrit donc:

$$\vec{j}_C = \sum_i qn_i \vec{v} = e(n_e \mu_e + n_h \mu_h) \vec{E} \quad \text{Eq. (17)}$$

où e est la charge de l'électron, n_e et n_h les densités d'électrons et de trous.

Dans un semiconducteur, la non-uniformité du dopage entraîne un courant de diffusion. En exprimant la loi de Fick, il vient:

$$\vec{j}_D = e(D_e \vec{\nabla} n_e - D_h \vec{\nabla} n_h) \quad \text{Eq. (18)}$$

où D_e et D_h sont les coefficients de diffusion dans le semiconducteur pour les électrons et les trous.

Les courants de conduction et de diffusion existant simultanément, la densité de courant totale s'écrit:

$$\vec{j}_{TOT} = \vec{j}_C + \vec{j}_D = e(n_e \mu_e + n_h \mu_h) \vec{E} + e(D_e \vec{\nabla} n_e - D_h \vec{\nabla} n_h) \quad \text{Eq. (19)}$$

Dans un semi-conducteur les phénomènes de conduction du courant par diffusion et par entraînement cohabitent avec le phénomène de génération-recombinaison des paires électrons-trous. L'équation de conservation de la matière, tenant compte de tous ces phénomènes, régissant l'évolution de la charge au cours du temps, est appelée équation de continuité:

$$\begin{aligned} \frac{\partial \bar{n}}{\partial t} &= \frac{1}{q} \vec{\nabla} j_e + (G_e - R_e) \\ \frac{\partial \bar{p}}{\partial t} &= -\frac{1}{q} \vec{\nabla} j_h + (G_h - R_h) \end{aligned} \quad \text{Eq. (20)}$$

où G_e (respectivement G_h) et R_e (respectivement R_h) sont les taux de génération et de recombinaison pour les électrons (respectivement les trous).

Les équations précédentes constituent des équations fondamentales de fonctionnement des dispositifs à semi-conducteurs. Elles sont associées dans les simulateurs numériques pour déterminer les courants et charges dans les dispositifs.

2.2 Effets quantiques

2.2.1 Confinement quantique dans un semiconducteur

Dans les transistors MOS modernes, certaines dimensions (épaisseur de l'oxyde, longueur, largeur et épaisseur de canal) ont atteint une taille proche des longueurs d'onde des électrons. Les effets de confinement sont donc à prendre en compte. Les porteurs de charge dans la couche d'inversion (sous l'influence d'un fort champ électrique) sont libres de se mouvoir parallèlement à la surface du canal mais leur mouvement perpendiculaire est confiné dans un puits de potentiel étroit, ou puits de confinement.

Sous l'hypothèse de la masse effective (chapitre 5), l'équation de Schrödinger est de la forme:

$$\left[-\frac{\partial}{\partial z} \frac{\hbar^2}{2m_z^*} \frac{\partial}{\partial z} + E_C(z) \right] \psi_{v,i}(z) = E_{v,i}(z) \psi_{v,i}(z) \quad \text{Eq. (21)}$$

Dans cette expression, E_C est le champ électrique de confinement, $\psi_{v,i}(z)$ (fonction propre) représente la fonction enveloppe des électrons selon la direction 'z' et $E_{v,i}(z)$ (valeur propre) le niveau d'énergie associé à l'état (v,i) du puits de potentiel.

L'idée consiste à traiter les porteurs du canal comme un gaz d'électron 2D. De même qu'en régime classique, la densité de porteur s'écrit:

$$n_{v,i}(E) = \int_{E_{v,i}}^{+\infty} N_C^{2D} f(E) dE \quad \text{Eq. (22)}$$

avec $f(E)$ la fonction de distribution de Fermi-Dirac et N_C^{2D} la densité d'état d'une sous bande.

Finalement, la densité de charge des différents niveaux peut s'écrire selon l'expression suivante:

$$n_{v,i} = \frac{g_v m_{//}^{*v} k_B T}{\pi \hbar^2} \ln \left(1 + \exp \left(\frac{E_F^{SC} - E_{v,i}}{k_B T} \right) \right) \quad \text{Eq. (23)}$$

avec g_v le degré de dégénérescence du type de vallée v [Mathieu'96]

Notons bien que la masse à prendre en compte dans l'évaluation des niveaux discrets (équation de Schrödinger 1D) est la masse m_{\perp}^* des électrons dans la direction perpendiculaire aux couches, tandis que la masse à considérer pour l'évaluation des densités de charge des niveaux discrets est la masse $m_{//}^*$ dans le plan des couches. Ainsi, dans le cas du silicium, le nombre de types de vallées équivalentes à prendre en compte ainsi que les valeurs de ces masses ($m_{//}^*$ et m_{\perp}^*) varient selon l'orientation cristallographique du silicium par rapport à l'interface (Si/SiO₂). Dans le cas de l'orientation $\langle 100 \rangle$, les six vallées Δ ellipsoïdales du bas de la bande de conduction du silicium sont à considérer en deux groupes conduisant à deux séries de niveaux :

- ✓ Les deux vallées Δ pour lesquelles la masse longitudinale m_{\perp}^* des électrons correspond à la direction perpendiculaire aux couches sont caractérisées par $m_{\perp}^* = m_{\perp}^*$ et $m_{//}^* = m_{\perp}^*$ (vallées normales).
- ✓ Les quatre vallées Δ pour lesquelles l'axe longitudinal de la surface ellipsoïdale iso-énergie est parallèle au plan des couches sont caractérisées par $m_{\perp}^* = m_{\perp}^*$ et $m_{//}^* = \sqrt{m_{\perp}^* m_{\perp}^*}$ (vallées parallèles).

La concentration électronique des états 2D au point z de l'espace est finalement évaluée en sommant les densités de charge des différents niveaux pondérées par le carré du module des fonctions enveloppes associées :

$$n_{2D}(z) = \sum_{v,i} n_{v,i} |\psi_{v,i}(z)|^2 \quad \text{Eq. (24)}$$

2.2.2 Approximation du puits triangulaire

Dans l'approximation du puits de potentiel triangulaire, la charge d'inversion est supposée faible devant la charge de désertion [Mathieu'96] donc on considère $V(z) = V_{dep}(z)$. Pour déterminer son expression, nous devons résoudre :

$$\frac{d^2 \phi_{dep}(z)}{dz^2} = -\frac{qN_A}{\epsilon_0 \epsilon_{SC}} \quad \text{Eq. (25)}$$

$$\phi_{\text{dep}}(z) = -\frac{qN_A}{\epsilon_0 \epsilon_{\text{sc}}} z_{\text{dep}} z \left(1 - \frac{z}{2z_{\text{dep}}} \right) \quad \text{Eq. (26)}$$

En utilisant l'Eq. (25), l'énergie potentielle $V(z)$ est égale (quand $z \rightarrow 0$) à :

$$V(z) = \frac{qN_A z_{\text{dep}}}{\epsilon_0 \epsilon_{\text{sc}}} z = qE_s z \quad \text{Eq. (27)}$$

L'énergie potentielle des porteurs à l'interface forme ainsi un puits de potentiel triangulaire dissymétrique; on parle alors d'*approximation du puits triangulaire*. Maintenant que l'énergie potentielle $V(z)$ est définie, l'équation de Schrödinger peut être résolue afin de déterminer la fonction d'onde $\psi_i(z)$ et le niveau d'énergie E_i associé à l'état 'i' du puits de potentiel. La probabilité de présence d'une particule est donnée par $|\varphi(y)|^2$. Celle-ci n'est pas maximale à l'interface oxyde-semiconducteur comme dans un cas classique, mais à une distance (appelée *dark space*) typiquement de l'ordre de quelques Å.

En utilisant une approche simplifiée comme la méthode variationnelle [Greiner'94], on peut par exemple calculer l'épaisseur du *dark space* pour le premier niveau d'énergie :

$$y_{\text{DARK SPACE}} = 2.3 \sqrt{\frac{\hbar^2}{12m^* qE}} \quad \text{Eq. (28)}$$

avec E le champ électrique transverse appliqué.

De manière générale, dans le cas d'une barrière de potentiel infinie (ce qui revient à éliminer toute possibilité de transmission à travers la barrière - effet tunnel) une particule a une probabilité nulle de se trouver au bord du puits de potentiel (i.e. à l'interface oxyde/semiconducteur). D'où ce phénomène de répulsion quantique et la création d'un *dark space*. Si la barrière n'est pas infinie, on autorise une légère pénétration des fonctions d'onde dans l'oxyde; la fonction d'onde de la particule ne s'annule pas contre la barrière, mais le confinement quantique existe toujours.

2.2.3 Simulation des effets quantiques

Le développement de modèles de transport quantique rigoureux est difficile à traiter numériquement. Le couplage Poisson/Schrödinger est utilisé pour prendre en compte les effets de confinement quantique dans la couche d'inversion. On calcule les niveaux d'énergie des différentes sous bandes mais aussi, à l'aide des fonctions d'onde, la densité de charges des différents niveaux. Cette approche très efficace en régime statique ne permet cependant pas d'intégrer les phénomènes de transport pourtant indispensables pour étudier le comportement électrique d'un dispositif. On utilise donc d'autres modèles, et la résolution directe Poisson/Schrödinger ne sert qu'à la calibration du modèle.

L'approche *Density-Gradient* est basée sur un modèle de dérive-diffusion quantique [Ancona'87] [Ancona'89] [ISE]. C'est une approximation du couplage des effets de confinement quantiques avec les équations de transport. Pour inclure les effets quantiques dans une simulation, une approche usuelle consiste à inclure un potentiel additionnel Λ dans l'expression classique de la densité:

$$n = N_C \exp\left(\frac{E_{FN} - E_C - \Lambda}{kT}\right) \quad \text{Eq. (29)}$$

où n est la densité électronique, k la constante de Boltzmann, T la température, N_C la densité d'état dans la bande de conduction, E_C l'énergie de la bande de conduction et E_{FN} le niveau de Fermi des électrons.

Dans le modèle *Density-Gradient*, Λ est donné par la relation:

$$\Lambda = -\frac{\hbar^2}{12m} \left(\nabla^2 \log(n) + \frac{1}{2} (\nabla \log n)^2 \right) = -\frac{\gamma \hbar^2}{6m} \frac{\nabla^2 \sqrt{n}}{\sqrt{n}} \quad \text{Eq. (30)}$$

où γ est un paramètre d'ajustement.

Comparé aux autres modèles de quantification, le modèle *Density-Gradient* peut également décrire les effets de quantification en deux ou trois dimensions [Toyabe'04].

Dans ce chapitre, nous avons utilisé le modèle de résolution '*Dérive-Diffusion*' qui résout les équations classiques (Eq. (15), Eq. (19) et Eq. (20)) et le modèle '*Density-Gradient*' (Eq. (29) et Eq. (30)) pour la section sur les effets quantiques.

3 Influence de la structure des transistors sur leurs performances électriques

3.1 Règles d'échelle

Nous rappelons tout d'abord les notations utilisées dans ce manuscrit (Fig. 60.a). Pour un transistor Triple-grille, nous parlerons dans ce chapitre de 'configuration FinFET' si la largeur de l'aileton est faible devant sa hauteur, de 'configuration Triple-grille' si largeur et hauteur de l'aileton sont du même ordre de grandeur et de 'configuration planaire' si la largeur de l'aileton est grande devant sa hauteur (Fig. 60.b). Ce sont des notations communément utilisées. Par rapport aux procédés exposés dans le chapitre 2, ces 'configurations' ne dépendent pas de la présence ou non d'un masque dur.

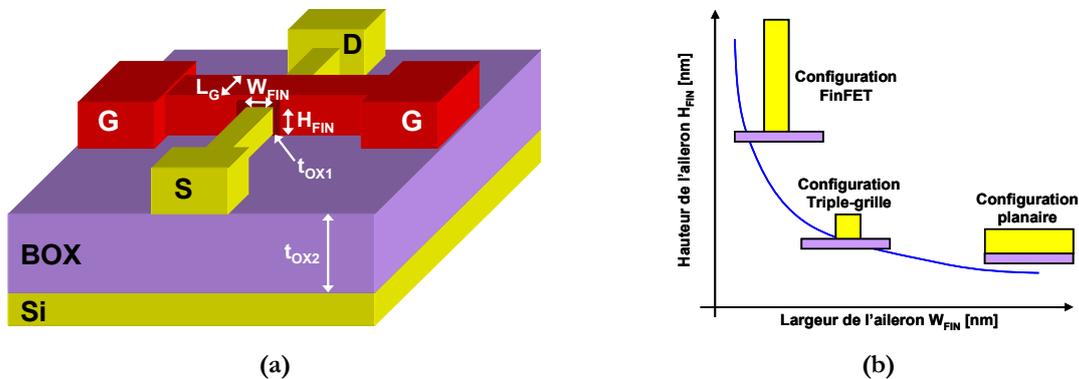


Fig. 60: (a): Schéma d'un transistor Triple-grille. (b): Les différentes 'configurations' d'un transistor Triple-grille.

Les plans d'expériences sont très utiles dans le cas d'une étude systématique d'une réponse dépendant de beaucoup de paramètres. Ici, la réponse est un paramètre électrique du

transistor (pente sous le seuil, DIBL, I_{OFF} , etc.) et les paramètres sont les dimensions du transistor. Pour des structures tridimensionnelles comme les Triple-grilles, la longueur de grille bien sûr mais également la largeur et la hauteur de l'aileron sont des paramètres de la réponse. En menant toute une série de simulations en fonction des paramètres (ici la largeur et la hauteur de l'aileron), on obtient ainsi une cartographie du comportement du transistor. Un 'modèle de réponse des surfaces', c'est-à-dire une interpolation sur les points simulés, permet ensuite d'obtenir une formule caractérisant la réponse en fonction des paramètres. Bien sûr, ce type de formule peut être qualifiée de 'semi-empirique' et n'a pas la rigueur d'une modélisation analytique compacte.

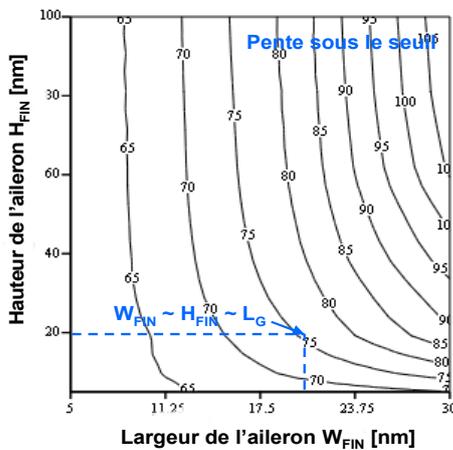
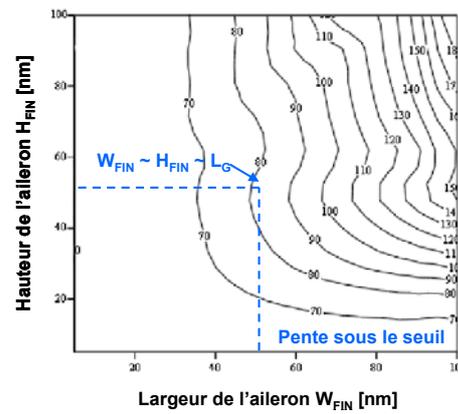
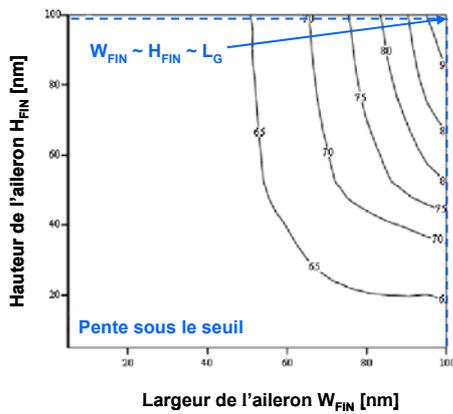


Fig. 61: Tracés de contours de la pente sous le seuil en fonction de la largeur et de la hauteur de l'aileron. (a) donne les résultats obtenus pour $L_G = 100$ nm, (b) pour $L_G = 50$ nm et (c) pour $L_G = 20$ nm. $t_{OX1} = 2$ nm, $t_{OX2} = 100$ nm, $N_A = 10^{15}$ cm⁻³, $\Phi_M = 4.5$ eV, dopage source/drain = $2 \cdot 10^{20}$ cm⁻³.

Sur la Fig. 61, le tracé de contours de la pente sous le seuil en fonction de la largeur et de la hauteur de l'aileron pour un transistor Triple-grille est représenté. Différentes longueurs de grille ont été étudiées ($L_G = 100$ nm sur la Fig. 61.a, 50 nm sur la Fig. 61.b et 20 nm sur la Fig. 61.c). Pour une longueur de grille donnée, les courbes 'iso-pente sous le seuil' présentent une allure similaire. Ainsi, la pente sous le seuil est bien contrôlée pour un transistor étroit (configuration 'FinFET') ou large et peu haut (configuration 'planaire complètement déserté'). Dans un cas ou dans l'autre, les couplages latéraux ou verticaux permettent de maintenir la pente sous le seuil à des valeurs acceptables. Ces deux configurations ne sont pas équivalentes:

la solution consistant à avoir un dispositif étroit est plus souple en terme d'épaisseur du film. C'est un résultat logique dans la mesure où le couplage latéral s'exerce avec deux grilles. Dans une configuration 'complètement désertée', $H_{FIN} = \frac{1}{4} L_G$ environ (voir Fig. 63.b); la solution de type Triple-grille, où $W_{FIN} \sim H_{FIN} \sim L_G$ s'avère ainsi être la moins contraignante au niveau des dimensions.

On obtient un autre résultat intéressant en examinant les courbes pour différentes longueurs de grille (de la Fig. 61.a à la Fig. 61.c). En diminuant la longueur de grille, la pente sous le seuil se dégrade fortement (effet de canal court). Le moyen de garder les performances sous contrôle consiste à diminuer conjointement largeur et hauteur de l'aileron. Le facteur à respecter entre les différentes grandeurs électriques pour les structures Triple-grilles est environ $L_G \sim W_{FIN} \sim H_{FIN}$. Ceci constitue ce qu'on appelle les 'lois d'échelle' [Colinge'04] [Poiroux'05].

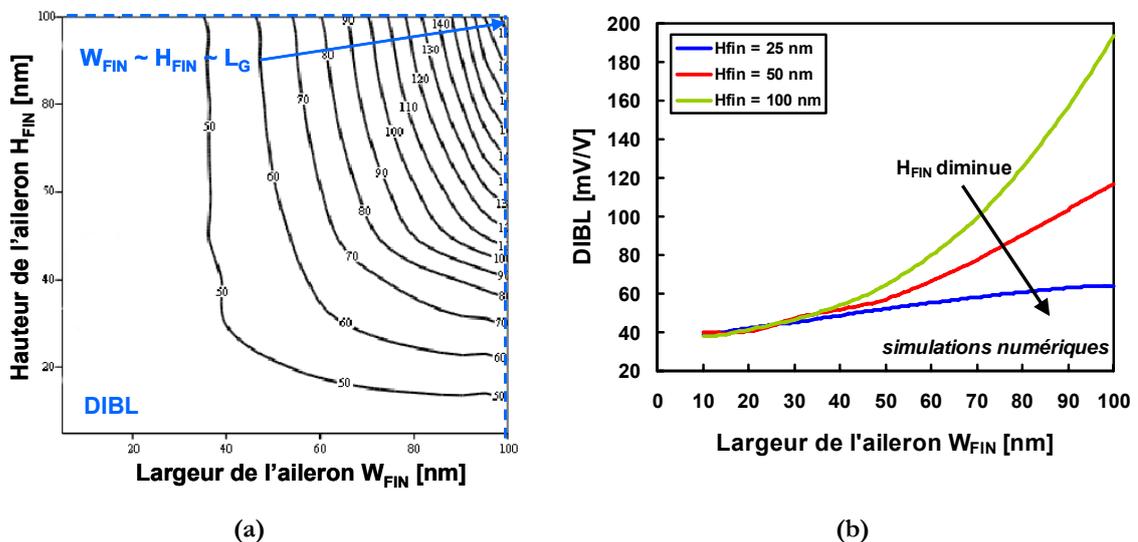


Fig. 62: Tracés de contours du DIBL en fonction de la largeur et de la hauteur de l'aileron pour $L_G = 100$ nm. (b) montre la prédiction obtenue pour une longueur de grille de 100 nm. $t_{OX1} = 2$ nm, $t_{OX2} = 100$ nm, $N_A = 10^{15}$ cm⁻³, $\Phi_M = 4.5$ eV, dopage source/drain = $2 \cdot 10^{20}$ cm⁻³.

On obtient des résultats très similaires concernant le DIBL (Fig. 62.a). La Fig. 62.b montre l'interpolation obtenue pour le DIBL en fonction de la largeur de l'aileron pour différentes hauteurs d'aileron. Suivant la largeur et la hauteur de l'aileron, on est soit dans une configuration 'planaire', 'Triple-grille' ou 'FinFET'. En dessous d'une certaine largeur le couplage latéral devient tel que les réponses électriques du dispositif ne dépendent plus de la hauteur de l'aileron. Au dessus de cette largeur d'aileron, le contrôle du DIBL nécessite d'apporter un meilleur couplage vertical avec un film plus mince.

Au final, il est intéressant de représenter les dimensions géométriques à respecter pour différents types de configuration en fonction de la longueur de grille (Fig. 63.a et Fig. 63.b; ces figures sont réalisées avec un canal fortement dopé afin de rendre les effets recherchés plus visibles). Sur la Fig. 63.a, la largeur d'aileron W_{FIN} nécessaire pour assurer une pente sous le seuil de 75 mV/décade est tracée en fonction de la longueur de grille pour les architectures

Triple-grille ($L_G \sim W_{FIN} \sim H_{FIN}$) et FinFET ($H_{FIN}/L_G \sim 3$; on considère qu'avec de tels rapports d'aspect le comportement électrostatique ne dépend plus de la hauteur de l'aileron). De manière attendue, la largeur d'aileron W_{FIN} à respecter dans le cas d'un FinFET est plus petite que pour un transistor Triple-grille (relaxation des contraintes d'échelle). Par contre, la réduction des longueurs de grille s'accompagne également d'un 'durcissement' des dimensions à respecter. Ainsi, dans le cas d'un Triple-grille, le rapport W_{FIN}/L_G à respecter à $L_G = 20$ nm est d'environ 0.7 contre 1 à $L_G = 30$ nm.

Sur la Fig. 63.b, la hauteur d'aileron H_{FIN} nécessaire pour assurer une pente sous le seuil de 75 mV/décade est tracée en fonction de la longueur de grille pour les architectures Triple-grille ($L_G \sim W_{FIN} \sim H_{FIN}$) et 'FDSOI planaire' (ici, $W_{FIN}/L_G = 3$). On constate qu'un transistor planaire est contraignant au niveau de l'épaisseur de film H_{FIN} requise, puisque le rapport H_{FIN}/L_G à respecter s'avère être de l'ordre de 0.25. Ceci signifie que pour une longueur de grille de 20 nm, on doit pouvoir contrôler avec une bonne précision un film de 5 nm d'épaisseur! Le couplage latéral du Triple-grille permet une bien plus grande souplesse, puisque le rapport H_{FIN}/L_G est compris entre 0.7 et 1 pour une longueur de grille comprise entre 20 et 30 nm. La partie hachurée sur la figure montre quant à elle la région où H_{FIN} est supérieur au cas d'un transistor Triple-grille, c'est-à-dire une configuration de type 'FinFET'.

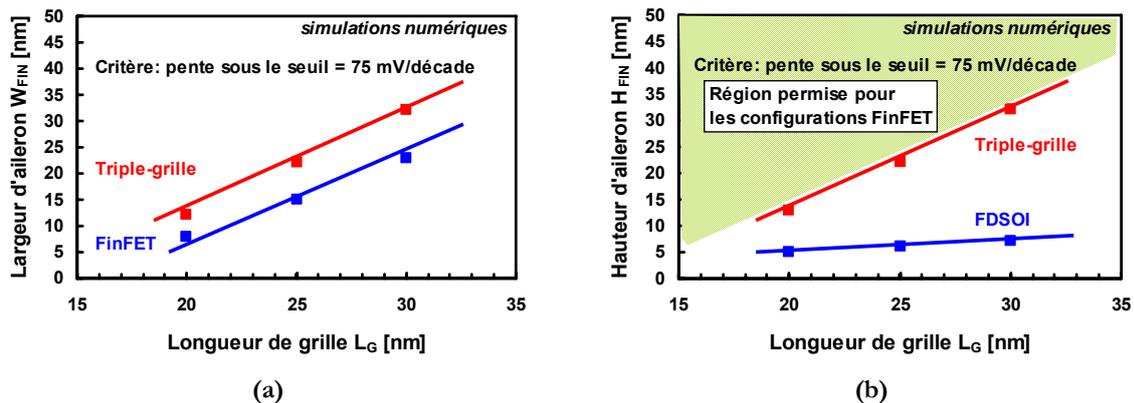


Fig. 63: (a): Largeurs d'aileron nécessaires pour conserver une pente sous le seuil de 75 mV/décade en fonction de la longueur de grille pour des transistors FinFET et Triple-grille. (b): Hauteurs d'aileron permises pour conserver une pente sous le seuil de 75 mV/décade en fonction de la longueur de grille pour des transistors complètement désertés, FinFET et Triple-grille. Pour ces simulations, l'épaisseur du BOX t_{OX2} est de 100 nm, le dopage du canal de l'ordre de $6 \cdot 10^{18} \text{ cm}^{-3}$ et le dopage des zones source/drain de $2 \cdot 10^{20} \text{ cm}^{-3}$.

3.2 Impact du nombre de grilles

Le couplage électrostatique des grilles du canal conditionne le contrôle des effets de canaux courts. Il existe toute une variété d'architectures à grilles multiples (voir chapitre 2). La Fig. 64 présente les résultats obtenus en simulation pour différents types d'architectures: Simple-grille (configuration 'complètement déserté'), Triple-grille, Quadruple-grille (dispositif 'à grille enrobante'), multicanaux sans oxyde enterré (quatre Quadruple-grille) et multicanaux avec un oxyde enterré (trois Quadruple-grille et un Triple-grille). La pente sous le seuil a été

extraite en fonction de la longueur de grille. La configuration 'complètement déserté' est la plus défavorable et la dégradation de la pente sous le seuil quand la longueur de grille est réduite est très nette (rappelons que $H_{FIN} = 25$ nm, et que la loi d'échelle $H_{FIN} \sim \frac{1}{4} L_G$ n'est pas respectée pour les petites longueurs de grille). La pente sous le seuil obtenue pour un transistor Triple-grille est bien meilleure grâce au couplage électrostatique latéral. La solution optimale consiste à utiliser un transistor Quadruple-grille: on maximise ainsi le couplage de la grille. Les résultats obtenus avec les multicanaux avec/sans oxyde enterré sont du même ordre que ceux obtenus avec un dispositif Quadruple-grille. La gravure de l'oxyde enterré et le dépôt d'une grille sous le canal (étape de procédé complexe) n'améliorent que très légèrement la pente sous le seuil. Un dispositif à multicanaux avec BOX a une meilleure pente sous le seuil qu'un transistor Triple-grille: les effets de canaux courts ne sont donc pas directement liés au 'pire contrôle électrostatique'. C'est un résultat que nous allons retrouver dans le paragraphe suivant.

De manière synthétique, on peut conclure qu'augmenter le nombre de grilles améliore la pente sous le seuil en optimisant le contrôle électrostatique du canal.

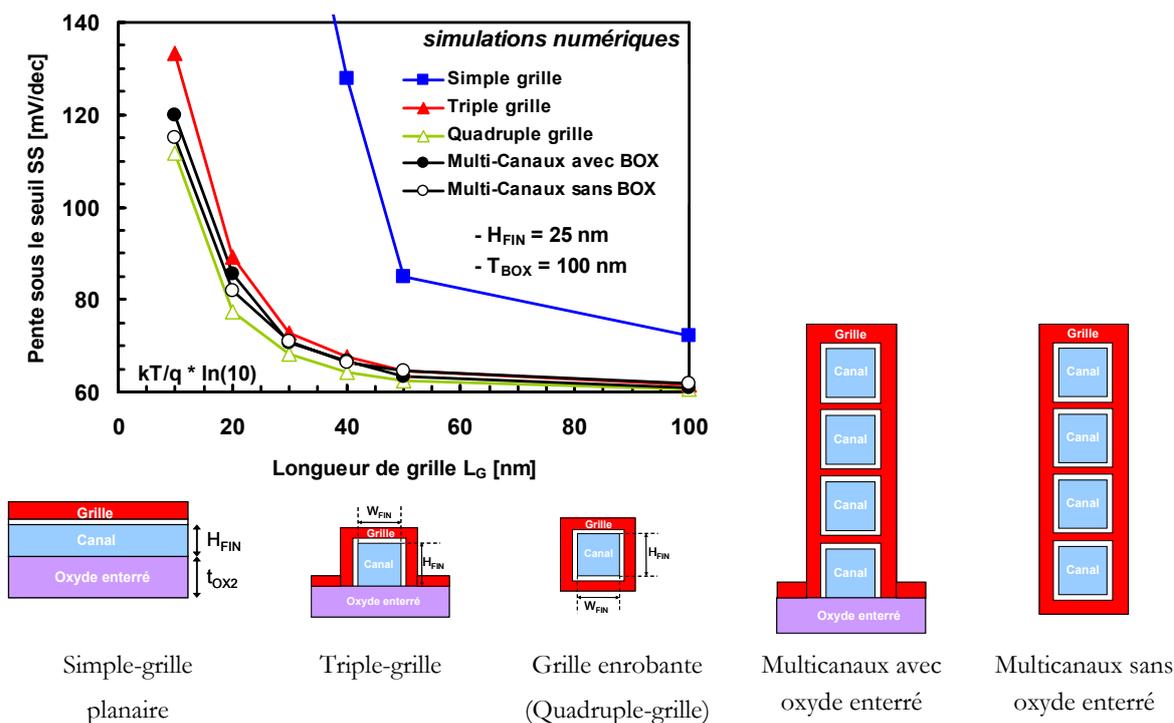


Fig. 64: Simulations numériques montrant l'évolution de la pente sous le seuil en fonction de la longueur de grille pour différents types d'architecture (FDSOI, Triple-grille, Quadruple-grille, multicanaux avec et sans BOX). La largeur de l'aileron W_{FIN} est de 500 nm pour les structures planaires, et de 25 nm pour les autres structures. L'épaisseur de l'oxyde enterré t_{OX2} est de 100 nm, le dopage du canal est de 10^{15} cm⁻³ et le dopage des zones de source et drain est de $2 \cdot 10^{20}$ cm⁻³.

3.3 Impact de la non-verticalité des flancs

Lors de l'étape de gravure de l'aileron, la côte lithographique peut ne pas être respectée et conduire à des flancs non verticaux. La base du transistor sera alors plus large que le sommet du transistor, et les performances de pente sous le seuil se trouveront alors dégradées (Fig. 65.a). Plus le rapport de forme du transistor (rapport H_{FIN}/W_{FIN}) est grand, plus la base du transistor sera large et plus la pente sous le seuil est dégradée (Fig. 65.b). En conséquence, augmenter le rapport H_{FIN}/W_{FIN} du transistor pour améliorer densité d'intégration et couplage électrostatique diminue la tolérance sur l'angle θ ; en d'autres termes plus H_{FIN}/W_{FIN} est grand et plus l'angle θ doit être proche de 90° , ce qui est très délicat du point de vue de la gravure de l'aileron.

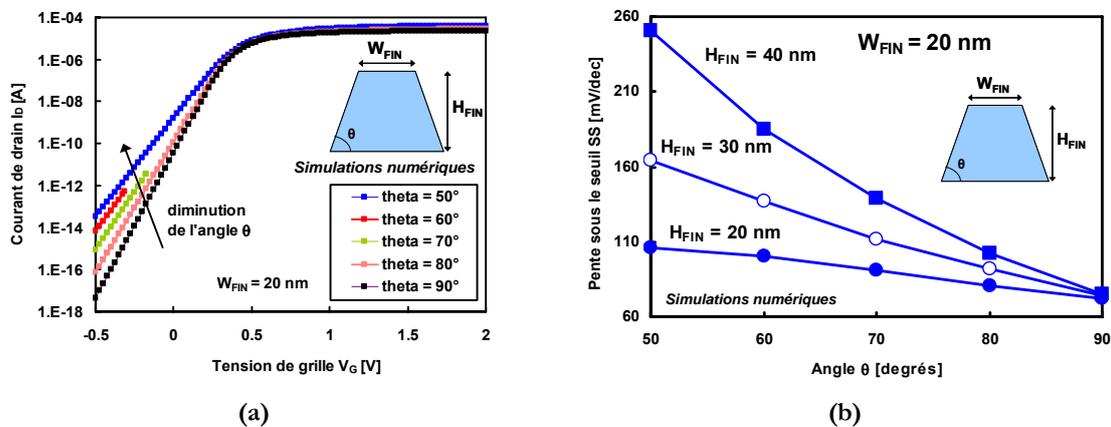


Fig. 65: Caractéristiques $I_D(V_G)$ simulées pour différents angles à la base de l'aileron (a) et pente sous le seuil correspondante (b). La longueur de grille L_G est de 40 nm. L'épaisseur de l'oxyde enterré t_{OX2} est de 100 nm, le dopage du canal de 10^{15} cm^{-3} et le dopage des zones source/drain de $2 \cdot 10^{20} \text{ cm}^{-3}$.

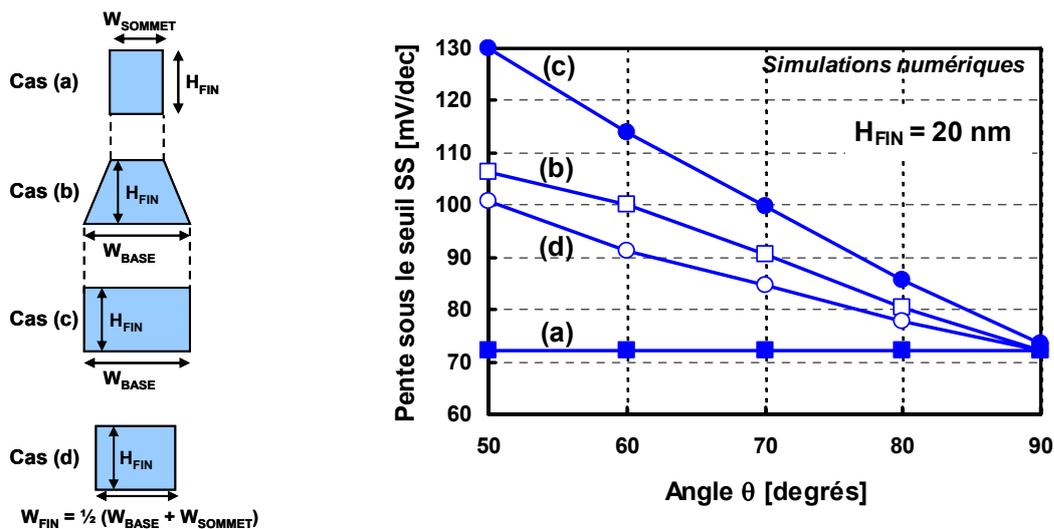


Fig. 66: Pente sous le seuil pour les différentes configurations étudiées: largeur constante égale à la largeur du sommet du trapèze (a), configuration trapézoïdale (b), largeur constante égale à la largeur de la base (c), largeur constante égale à la moyenne de la largeur de la base avec celle du sommet (d). $L_G = 40$ nm, $t_{OX2} = 100$ nm, $N_A = 10^{15} \text{ cm}^{-3}$ et le dopage des zones source/drain est de $2 \cdot 10^{20} \text{ cm}^{-3}$.

Sur la Fig. 66, la pente sous le seuil pour différentes configurations est tracée afin d'examiner la grandeur caractéristique pilotant sa dégradation. Le cas (a) est la référence et correspond à un transistor rectangulaire de largeur W_{SOMMET} . La pente sous le seuil d'un transistor rectangulaire (cas (c)) dont la largeur est la même que la base d'un transistor trapézoïdal est considérablement plus grande que la pente sous le seuil du trapèze correspondant (cas (b)). De même, en considérant une largeur égale à la moyenne algébrique du sommet et de la base (cas (d)) la variation de la pente sous le seuil obtenue est légèrement plus faible que celle du trapèze. Le comportement sous le seuil d'un trapèze est donc plus complexe que la simple dépendance dans la largeur de la base du transistor. La stratégie consistant à dire que la configuration la plus défavorable (ici, la largeur maximale W_{BASE}) va gouverner le comportement de la pente sous le seuil n'est donc pas la bonne.

3.4 Effets des résistance d'accès et des contacts

De par leur structure tridimensionnelle, on peut également se poser la question du comportement des régions de source et drain. La position des contacts métalliques peut se faire de plusieurs manières dans un transistor Triple-grille (Fig. 67). En modifiant la résistance source/drain du transistor, on modifie les performances électriques des dispositifs.

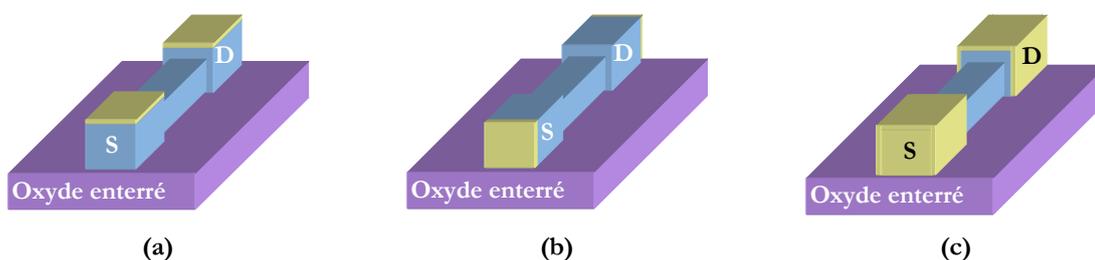


Fig. 67: Schéma des différents contacts étudiés. Le contact est au dessus des zones sources/drain (a), à l'arrière (b) ou enrobant (c).

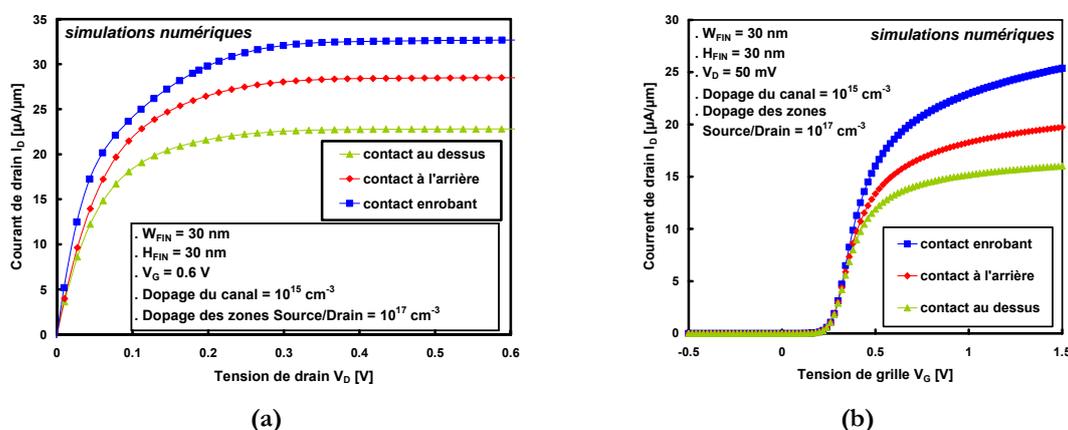


Fig. 68: Courant de drain simulé en fonction de la tension de drain (a) ou de la tension de grille (b) pour différents types de contacts. Dans les deux cas, le canal est non dopé ($N_A = 10^{15} \text{ cm}^{-3}$) et les zones de source et drain faiblement dopées ($N_D = 10^{17} \text{ cm}^{-3}$).

La Fig. 68 présente les résultats obtenus pour un canal intrinsèque ($N_A = 10^{15} \text{ cm}^{-3}$) et des zones de source et drain faiblement dopées ($N_D = 10^{17} \text{ cm}^{-3}$). Au dessus du seuil et en saturation, la perte de courant par rapport à un contact enrobant est d'environ 15 % pour un contact à l'arrière du piédestal source/drain et de 30 % pour un contact situé au dessus (Fig. 68.a). Le dépôt d'un contact sur le dessus d'un transistor est donc la situation la plus défavorable, suivie du contact à l'arrière du piédestal; la solution optimale est le contact enrobant. Cette variation des résistances source/drain suivant la topologie des contacts est également très visible sur la Fig. 68.b.

Cependant, ces résultats spectaculaires sont spéculatifs car ils sont obtenus pour des dopages des zones source/drain faibles. En augmentant le dopage des zones de source et drain (de $N_D = 1 \cdot 10^{17} \text{ cm}^{-3}$ à $2 \cdot 10^{20} \text{ cm}^{-3}$), le courant de sortie obtenu pour un transistor augmente fortement puisque l'on diminue les résistances source/drain (Fig. 69). De même, l'influence de la zone de dépôt du contact diminue quand le dopage des zones source/drain augmente. En arrivant dans la plage des dopages utilisés dans nos transistors (dose de $5 \cdot 10^{12}$ à $3 \cdot 10^{15} \text{ cm}^{-2}$ pour les NMOS, de $1 \cdot 10^{13}$ à $3 \cdot 10^{15} \text{ cm}^{-2}$ pour les PMOS), elle n'a quasiment plus aucune importance.

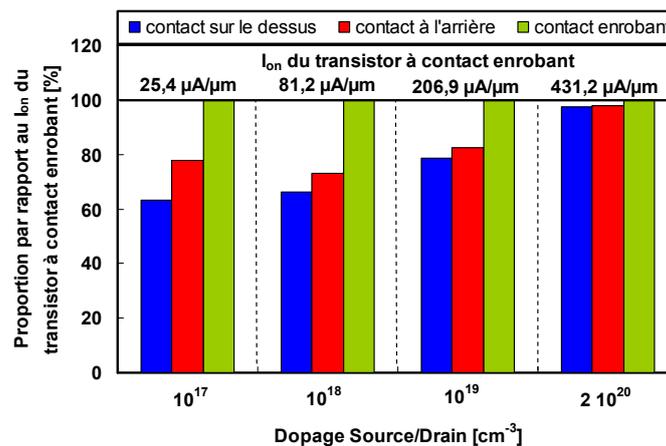


Fig. 69: Rapport entre les courants obtenus selon les différents types de contact en fonction du dopage dans les zones de source et drain.

3.5 Conclusions sur le contrôle des effets de canaux courts

Ce paragraphe a mis en lumière certains aspects de la réduction des dimensions dans les composants multigrille:

- ✓ En augmentant le nombre de grille, on améliore fortement le contrôle électrostatique. Par conséquent, les règles d'échelle (c'est-à-dire le rapport entre les différentes dimensions d'un transistor pour conserver de bonnes performances) sont considérablement assouplies. A titre d'exemple, un transistor Triple-grille d'une longueur de grille de 30 nm doit avoir une largeur et une hauteur de 30 nm

au plus, alors que pour un transistor 'complètement déserté' il faut une épaisseur de film d'environ 8 nm!

- ✓ Pour une structure de type Triple-grille, le cas optimal en simulation correspond à des flancs verticaux. Tout écart s'accompagne d'une dégradation des performances.
- ✓ La position des contacts métalliques sur le piédestal source/drain n'a pas d'influence à condition d'avoir des dopages des zones sources/drain importants.

4 Effets tridimensionnels dans les structures Triple-grille: les effets de coins

4.1 Aspects classiques

4.1.1 Impact sur la tension de seuil

Malgré leur excellent contrôle électrostatique, les architectures Triple-grille ont été pressenties comme problématiques à cause de la région des coins (Fig. 70.a). En effet, cette zone est soumise à l'influence de deux grilles adjacentes et s'avère électrostatiquement favorable au passage des porteurs. En d'autres mots, il peut se créer un canal de coins parasites s'activant avec le seuil principal du transistor. Ce canal parasite pourrait se révéler très gênant pour le contrôle du courant à l'état bloqué. Cet effet porte le nom 'd'effet de coin'. Cette zone favorable au passage des porteurs peut être modélisée en résolvant l'équation de Poisson dans une structure Triple-grille:

$$\begin{cases} \nabla^2\varphi = -\frac{q}{\epsilon_{Si}}(n + p + N_D - N_A) \text{ dans le canal} \\ \nabla^2\varphi = 0 \text{ dans l'oxyde enterré et l'oxyde de grille} \end{cases} \quad \text{Eq. (31)}$$

La solution obtenue par simulation numérique (Fig. 70.b) montre que le potentiel électrostatique est plus élevé dans les régions de coins.

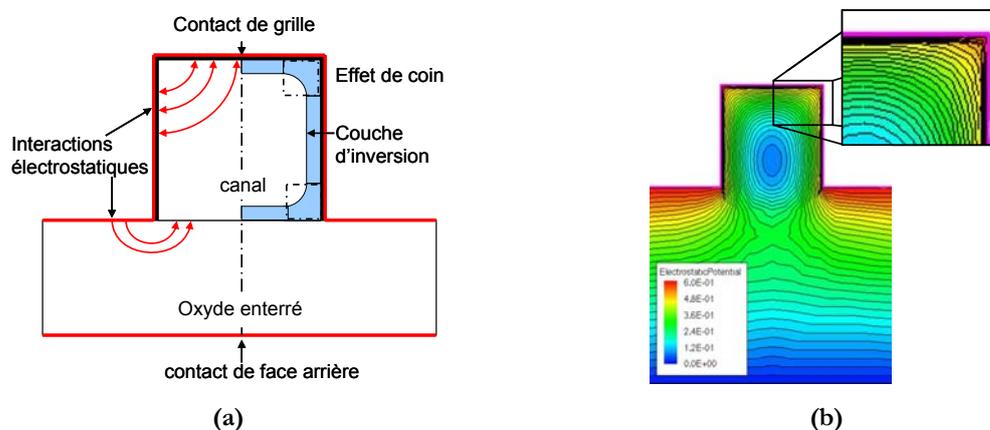


Fig. 70: Coupe d'un transistor Triple-grille montrant la configuration conduisant aux effets de coins (a). Simulation numérique DESSIS illustrant ce phénomène (b). $N_A = 10^{15} \text{ cm}^{-3}$, $L_G = 500 \text{ nm}$, $W_{FIN} = 25 \text{ nm}$, $H_{FIN} = 30 \text{ nm}$, $t_{OX2} = 100 \text{ nm}$.

Plus généralement, on peut également considérer qu'à cause de l'influence de la grille supérieure le potentiel de surface et la tension de seuil vont varier le long d'une grille latérale. Ainsi, la tension de seuil est bidimensionnelle et dépend de l'endroit du canal considéré.

Cependant, la caractéristique électrique d'un transistor ne permet pas de remonter directement à ces tensions de seuil 'locales'. Nous avons envisagé une autre méthode. Pour chaque dispositif simulé, la tension de seuil 'globale' est calculée par une méthode classique comme la deuxième dérivée du courant de drain (Fig. 71). Ensuite, la densité électronique dans le canal est extraite pour cette tension de grille (Fig. 72.a) puis intégrée sur de petites zones (Fig. 72.b) afin de calculer les charges d'inversion locales.

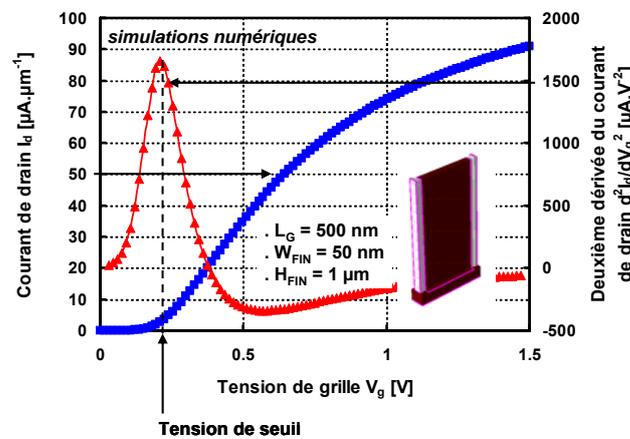


Fig. 71: Courbe $I_D(V_G)$ simulée d'un transistor Triple-grille ayant un très fort rapport de forme ($W_{FIN} = 50$ nm, $L_G = 500$ nm, $H_{FIN} = 1$ µm, $N_A = 10^{15}$ cm⁻³). Le pic de dérivée seconde (triangles) indique la tension de seuil.

En utilisant la relation suivante, on peut ainsi calculer les tensions de seuil 'locales':

$$V_{TH}^{LOCAL} = V_G - \frac{Q_{INV}^{LOCAL}}{C_{OX}} \quad \text{Eq. (32)}$$

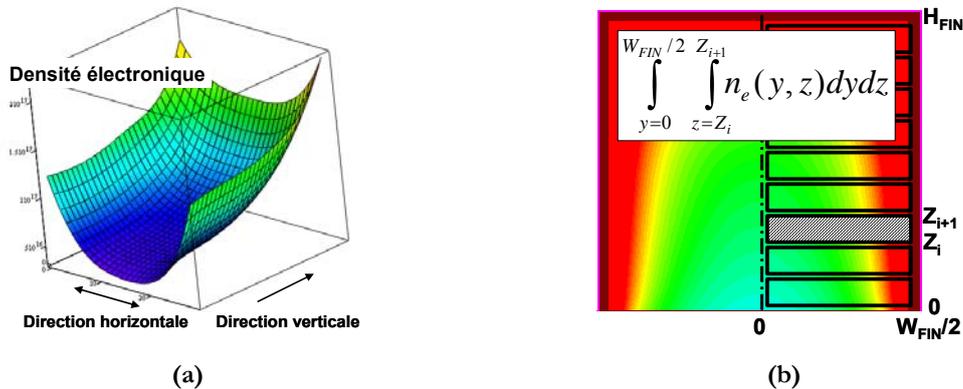


Fig. 72: Interpolation bidimensionnelle de la densité électronique extraite (a), et schéma de l'intégration réalisée pour obtenir les tensions de seuil 'locales' (b).

Par rapport à une méthode comme la dérivée seconde du courant de drain, cette approche présente l'inconvénient de se situer au dessus du seuil et donc d'aplanir les variations

observées. Par contre, elle permet d'obtenir la tension de seuil partout dans le canal. Les simulations réalisées portent sur des transistors de type FinFET (avec masque dur) et Triple-grille pour différents dopages et différentes dimensions. Pour éviter les effets de canaux courts, la longueur de la grille est de $0.5 \mu\text{m}$. La largeur de l'aileron est fixe et égale à 50 nm , et la hauteur du canal est variable de 25 à 200 nm (on change donc le rapport de forme). Le canal est "intrinsèque" ($N_A = 10^{15} \text{ cm}^{-3}$) ou très fortement dopé ($N_A = 10^{18} \text{ cm}^{-3}$). On peut noter que ce dopage est uniforme, ce qui ne serait pas le cas dans la réalité pour les forts dopages à cause du procédé d'implantation des dopants (en particulier, on néglige ici tout effet de désertion des dopants dans les coins avec le dopage). L'oxyde de grille et l'oxyde enterré font respectivement 1.9 et 100 nm .

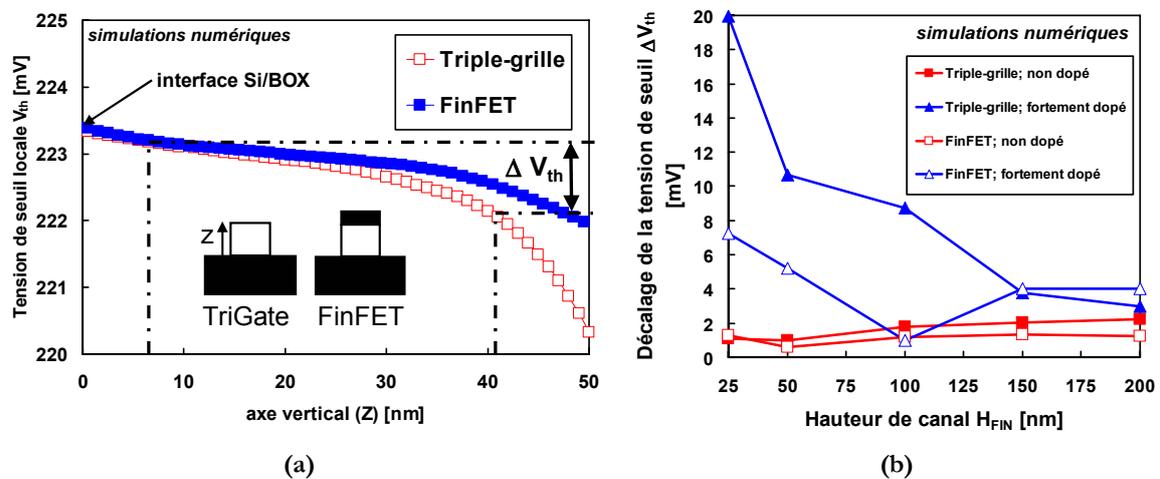


Fig. 73: Tension de seuil locale le long de l'axe vertical pour un transistor FinFET et un transistor Triple-grille (a). Le canal n'est pas dopé ($N_A = 10^{15} \text{ cm}^{-3}$), $W_{FIN} = 50 \text{ nm}$, $H_{FIN} = 50 \text{ nm}$ et $L_G = 0.5 \mu\text{m}$. courbe $\Delta(V_{TH})$ vs. H_{FIN} pour des transistors FinFETs et des transistors Triple-grille (b) avec un canal non dopé ($N_A = 10^{15} \text{ cm}^{-3}$) ou fortement dopé ($N_A = 10^{18} \text{ cm}^{-3}$). $W_{FIN} = 50 \text{ nm}$ et $L_G = 0.5 \mu\text{m}$.

La tension de seuil locale le long de la grille verticale pour un FinFET et un Triple-grille non dopés est représentée sur la Fig. 73.a. La tension de seuil du transistor Triple-grille décroît à proximité de la grille supérieure à cause de l'influence de celle-ci. Pour le transistor FinFET, la présence du masque dur réduit considérablement cette chute. On constate d'emblée que la chute de tension de seuil n'est que de l'ordre du mV.

ΔV_{th} est défini (Fig. 73.a) comme la chute de tension maximale le long de l'axe vertical. Sur la Fig. 73.b, ΔV_{th} est tracé en fonction de la hauteur du canal pour des transistors FinFETs et Triple-grilles dopés et non dopés. Pour un fort rapport de forme (i.e. une grande hauteur de canal H_{FIN}), ΔV_{th} sera logiquement plus faible puisque l'influence de la grille supérieure s'en trouve d'autant atténuée. En comparant FinFET et Triple-grille pour un même dopage, on remarque que ΔV_{th} est plus faible pour les FinFETs à cause de la désactivation de la grille supérieure. Un faible dopage réduit également le décalage de la tension de seuil pour les transistors FinFETs (Fig. 73.b, carrés) et Triple-grille (Fig. 73.b, triangles). Ceci est dû à la

réduction du terme source dans l'équation de Poisson (Eq. (31)) conduisant à une plus faible influence de la grille supérieure.

Cependant, le décalage de la tension de seuil maximal est seulement de 20 mV environ pour la pire configuration, un transistor Triple-grille dopé ayant un rapport de forme de 0.5 ($H_{FIN} = 25$ nm et $W_{FIN} = 50$ nm). Le décalage de la tension de seuil observé est donc faible quelque soit la configuration utilisée. On peut en conclure que l'impact des effets de coins sur le décalage de la tension de seuil sera faible pour les dopages utilisés dans nos transistors expérimentaux.

4.1.2 Effet de l'arrondi de coin

La configuration des coins dans un transistor Triple-grille conduit à l'activation préférentielle de cette région. Cependant, cet effet électrostatique peut également être supprimé en changeant la forme du canal. Si le coin est remplacé par un arrondi, la couche d'inversion est uniforme tout le long du canal (Fig. 74.a). Il en résulte que pour un transistor très fortement dopé, le canal de coin visible pour un canal carré est supprimé pour un canal arrondi. Cet effet est visible sur la Fig. 74.b, où l'activation des coins se manifeste pas un pic de dérivée seconde à une plus faible tension de seuil que le pic principal. Pour un canal intrinsèque, l'effet de coin est quasi-inexistant. L'impact de l'application d'un arrondi de coin n'est alors pas visible.

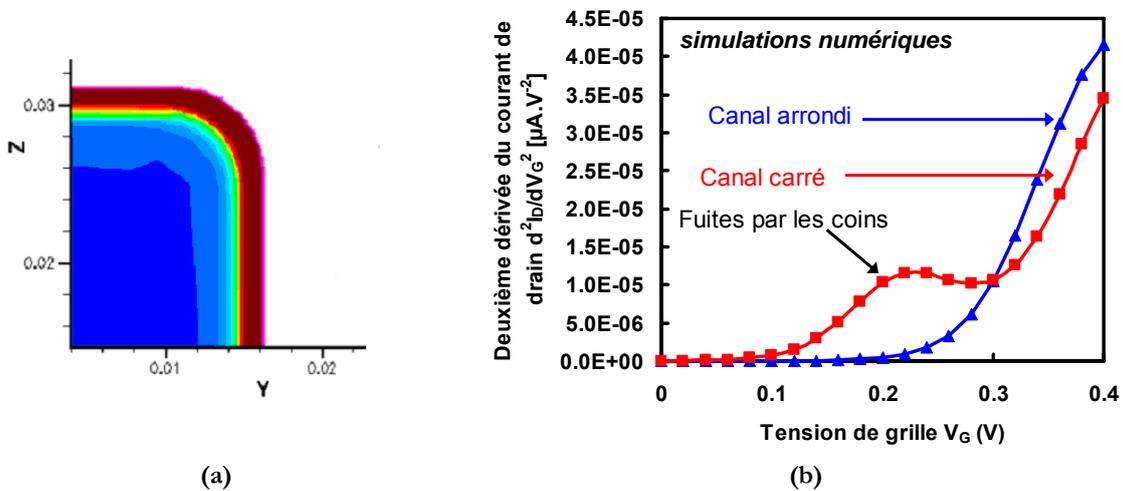


Fig. 74: Coupe transversale de la densité électronique pour un canal arrondi (a) et deuxième dérivée du courant de drain pour des canaux carrés et arrondis et pour un dopage très fort ($N_A = 10^{19} \text{ cm}^{-3}$) (b)

On peut se poser la question de la variation de la tension de seuil du transistor avec l'arrondi de coin. Avec un canal fortement dopé, l'effet de coin n'est pas négligeable et conduit à des variations importantes de tension de seuil. La Fig. 75 montre l'évolution de la tension de seuil en saturation avec le rayon de courbure de l'arrondi appliqué sur une structure Triple-grille fortement dopée. Le rayon de courbure zero correspond à un transistor sans arrondi: l'effet de coin se traduit par une activation du transistor par les coins et donc par une faible tension de seuil mesurée. En augmentant le rayon de courbure du coin, l'activation des coins

(autrement formulé, on peut également parler de fuites par les coins) est graduellement supprimée et la tension de seuil mesurée augmente. Avec de telles structures, le rayon de courbure à appliquer est néanmoins à optimiser car le gain en I_{OFF} obtenu avec de forts rayons de courbure est à tempérer par la perte de I_{ON} consécutive au plus faibles surfaces de canaux de conduction.

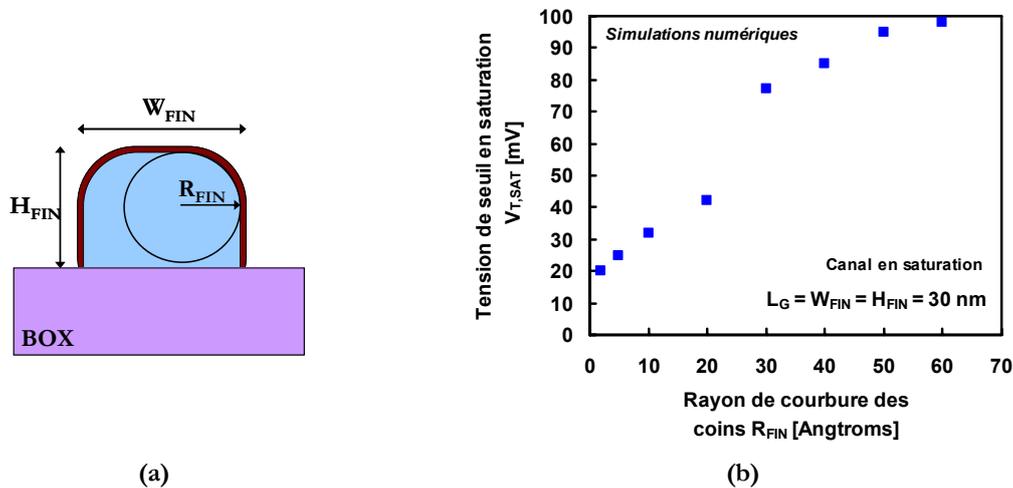


Fig. 75: (a): Illustration schématique du transistor Triple-grille arrondi considéré. (b): Evolution de la tension de seuil en saturation en fonction du rayon de courbure appliqué sur les coins d'un transistor Triple-grille ($L_G = W_{FIN} = H_{FIN} = 30$ nm). $N_A = 10^{19}$ cm⁻³.

4.2 Impact de la prise en compte des effets quantiques

La densité électronique avec le modèle *Density-Gradient* sur un transistor 'à grille enrobante' de faible section (10 nm * 10 nm) est montrée sur la Fig. 76. Basiquement, on se retrouve dans une situation de concurrence entre l'interaction électrostatique Coulombienne et la répulsion quantique. Pour de faibles tensions de grilles (Fig. 76.a), la répulsion quantique renforce donc le phénomène d'inversion volumique [Balestra'87]. Pour des tensions de grilles plus élevées, la conjonction de la conduction préférentielle dans les coins et de la répulsion quantique conduit à la création de quatre zones à plus fortes densités de porteurs dans les coins (Fig. 76.b et Fig. 76.c).

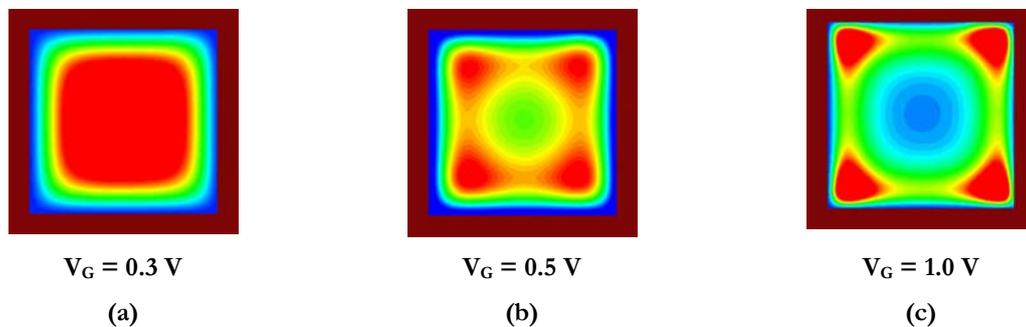


Fig. 76: Coupe transversale de la densité électronique dans le canal d'un transistor 'à grille enrobante' ($W_{FIN} = H_{FIN} = 10$ nm) pour une tension de grille V_G de 0.3 V (a), 0.5 V (b) et 1.0 V (c). $N_A = 10^{15}$ cm⁻³; $L_G = 500$ nm. Les échelles ne sont pas les mêmes pour les différentes figures ; la couleur rouge traduit une plus forte densité électronique que la couleur bleue.

On retrouve ce phénomène en extrayant la densité électronique le long d'une section diagonale du transistor (Fig. 77.a). L'inversion en volume est marquée pour un canal intrinsèque (Fig. 77.b) mais peu importante pour un canal fortement dopé (Fig. 77.c). Dans les régions de coins, on retrouve le fait que la différence 'centre du canal/coins du canal' diminue très fortement avec le dopage (absence d'effet de coins pour des canaux non dopés).

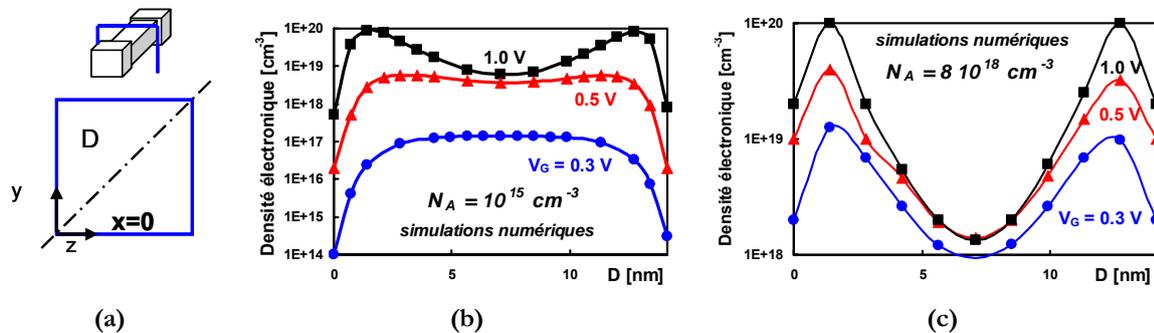


Fig. 77: Densité électronique en fonction de la distance le long d'une diagonale (a) d'un transistor 'à grille enrobante' ($W_{FIN} = H_{FIN} = 10$ nm) pour un dopage intrinsèque ($N_A = 10^{15} \text{ cm}^{-3}$, (b)) ou fort ($N_A = 8.10^{18} \text{ cm}^{-3}$, (c)).

Au niveau des caractéristiques électriques (Fig. 78.a), la prise en compte des effets quantiques conduit à d'importantes variations pour les petites dimensions. La tension de seuil augmente quand la section du dispositif diminue à cause de la plus importante séparation des sous bandes, et du couplage des grilles. La prise en compte des effets quantiques revient à considérer le remplissage des sous bandes quantiques. L'énergie de la première sous bande augmente avec la concentration électronique [Colinge'06]; cela crée plus de répulsion électrostatique dans la sous bande et plus d'énergie est nécessaire pour augmenter la densité électronique dans la sous bande. Par conséquent, la différence des tensions de seuil d'un dispositif simulé avec/sans effet quantique augmente pour les faibles sections (Fig. 78.b). On observe également à un léger abaissement du courant de sortie avec la prise en compte des effets quantiques. Cependant, les dimensions à atteindre pour que ces phénomènes prennent de l'ampleur sont très petites (< 5 nm de section).

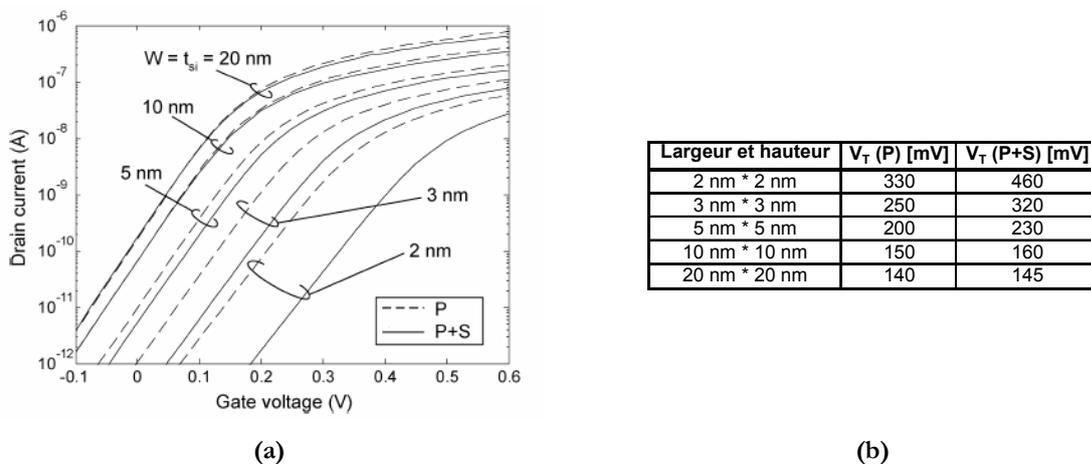


Fig. 78: Evolution des caractéristiques électriques d'un transistor Triple-grille en fonction de sa section et du modèle de résolution utilisé (P: Poisson; P+S: Poisson + Schrödinger) (a); valeurs des tensions de seuil en régime linéaire avec ou sans effets quantiques (b). $N_A = 5 \cdot 10^{17} \text{ cm}^{-3}$; $L_G = 1 \mu\text{m}$. (d'après [Colinge'06])

5 Couplages d'interfaces dans les Ω FETs

5.1 Cas des transistors larges

5.1.1 Modèle unidimensionnel de Lim et Fossum

Cette partie reprend les principaux résultats du modèle de couplage d'interfaces de Lim et Fossum [Lim'83] [Cristoloveanu'95]. Le modèle de Lim et Fossum s'applique dans le cas des MOSFETs complètement désertés. Les hypothèses considérées sont un canal long, une faible tension de drain, un dopage constant, une approximation de charges en surface (c'est-à-dire qu'on néglige l'épaisseur de la couche d'accumulation et d'inversion), et aucune influence de la part de l'interface située sous l'oxyde enterré.

L'intégration de l'équation de Poisson dans le film déserté et l'application du théorème de Gauss aux interfaces avant et arrière conduisent à la relation:

$$V_{G_{1,2}} = \varphi_{fb_{1,2}} + \psi_{s_{1,2}} \left(1 + \frac{C_{Si} + C_{it_{1,2}}}{C_{ox_{1,2}}}\right) - \psi_{s_{2,1}} \frac{C_{Si}}{C_{ox_{1,2}}} - \frac{2Q_{c_{1,2}} + Q_{Si}}{2C_{ox_{1,2}}} \quad \text{Eq. (33)}$$

où $\psi_{s_{1,2}}$ est le potentiel (défini en supposant un point de référence neutre dans le film) aux

interfaces avant et arrière, $C_{Si} = \frac{\epsilon_{Si}}{t_{Si}}$ la capacité du film déserté, $C_{it_{1,2}} = qD_{it_{1,2}}$ la capacité

associée aux états d'interface, $Q_{c_{1,2}}$ les charges de surface dans le canal et $Q_{Si} = -qN_A t_{Si}$ la charge de désertion.

Le potentiel de bandes plates dans le silicium massif s'exprime sous la forme:

$$\varphi_{fb_{1,2}} = \phi_{ms_{1,2}} - \frac{Q_{f_{1,2}}}{C_{ox_{1,2}}} \quad \text{Eq. (34)}$$

La définition standard de la tension de seuil V_T du canal avant implique que $\psi_{s_1} = 2\varphi_F$ et $Q_{inv_1} = 0$ dans l'Eq. (33). On peut alors distinguer trois cas, suivant l'état de la face arrière :

Cas 1 : l'interface arrière est accumulée ($\psi_{s_2} = 0$)

$$V_{T_1}^{acc} = \phi_{fb_1} + \left(1 + \frac{C_{Si} + C_{it_1}}{C_{ox_1}}\right) 2\varphi_F - \frac{Q_{Si}}{2C_{ox_1}} \quad \text{Eq. (35)}$$

Cas 2 : l'interface arrière est inversée ($\psi_{s_2} = 2\varphi_F$)

$$V_{T_1}^{inv} = \varphi_{fb_1} + \left(1 + \frac{C_{it_1}}{C_{ox_1}}\right) 2\varphi_F - \frac{Q_{Si}}{2C_{ox_1}} \quad \text{Eq. (36)}$$

Cas 2 : l'interface arrière est désertée ($0 < \psi_{s_2} < 2\varphi_F$)

En éliminant ψ_{s_2} dans l'Eq. (33), il vient :

$$V_{T_1}^{dep} = V_{T_1}^{acc} - \frac{C_{Si} C_{ox_2}}{C_{ox_1} (C_{ox_2} + C_{Si} + C_{it_2})} (V_{G_2} - V_{G_2}^{acc}) \quad \text{Eq. (37)}$$

$V_{G_2}^{acc}$ représente la tension en dessous de laquelle la face arrière est toujours accumulée, quelque soit la valeur de V_{G_1} . On obtient cette valeur dans l'Eq. (33) en prenant $\varphi_{S_2} = 0$ et $\varphi_{S_1} = 2\phi_F$:

$$V_{G_2}^{acc} = \phi_{fb_2} - \frac{C_{Si}}{C_{ox_2}} 2\phi_F - \frac{Q_{Si}}{2C_{ox_2}} \quad \text{Eq. (38)}$$

De manière similaire, l'interface arrière est définitivement inversée pour $V_{G_2} > V_{G_2}^{inv}$, où $V_{G_2}^{inv}$ n'est rien d'autre que la tension de seuil du canal arrière $V_{T_2}^{inv}$:

$$V_{G_2}^{inv} = V_{T_2}^{inv} = \phi_{fb_2} + \left(1 + \frac{C_{it_2}}{C_{ox_2}}\right) 2\phi_F - \frac{Q_{Si}}{2C_{ox_2}} \quad \text{Eq. (39)}$$

Le comportement des expressions calculées ci-dessus est montré sur la Fig. 79. La tension de seuil du canal avant décroît linéairement entre deux plateaux correspondant à l'accumulation et à l'inversion de la face arrière.

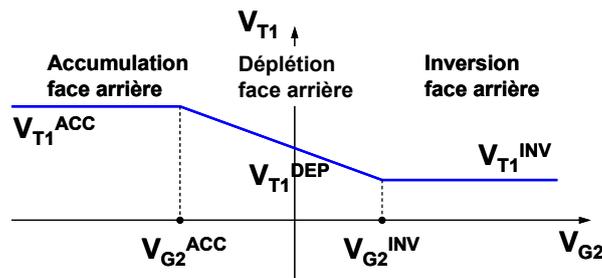


Fig. 79: Influence de la tension de face arrière sur la tension de seuil de la face avant dans le cas d'un transistor NMOS 'complètement déserté'.

La différence entre les plateaux $\Delta V_{T_1} = \frac{C_{Si}}{C_{ox_1}} 2\phi_F$ est légèrement dépendante du dopage, alors que la pente de la courbe $V_{T_1}(V_{G_2})$ ne l'est pas. Aucun de ces paramètres ne dépend de la qualité de l'interface oxyde-semiconducteur avant, mais ils peuvent être contrôlés en ajustant l'épaisseur de silicium.

Dans les films très fins, le coefficient de couplage est ramené au rapport entre l'oxyde de grille et l'oxyde enterré $\frac{t_{ox_1}}{t_{ox_2}}$.

5.1.2 Mesures expérimentales sur des transistors larges

Un transistor Ω FET large se comporte comme un transistor de type totalement déserté. En polarisant la face arrière, le fonctionnement du transistor s'en trouvera donc modifié selon le modèle de Lim et Fossum. La tension de seuil diminue quand la polarisation de la face arrière va de l'accumulation vers l'inversion face arrière (Fig. 80).

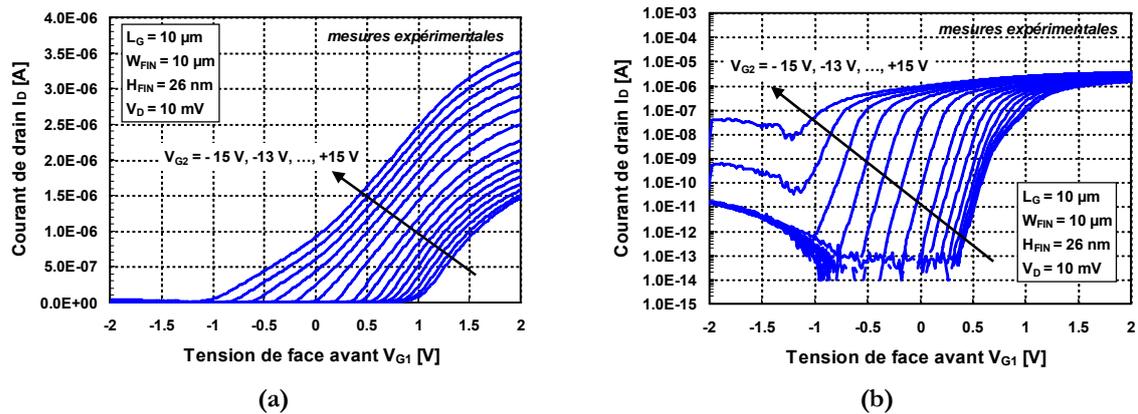


Fig. 80: Courant de drain en fonction de la tension de grille face avant pour différentes polarisations de face arrière en coordonnées linéaire (a) et logarithmique (b).

Quand la tension de face arrière est suffisante pour passer en régime d'inversion face arrière, la caractéristique présente deux tensions de seuil correspondant à l'inversion du canal arrière puis à l'inversion de tous les canaux de conduction. Ceci est très visible quand on représente la transconductance en fonction de la tension de grille (Fig. 81.a); le palier de transconductance (visible dans la plage $-1 \text{ V} < V_{G1} < 0 \text{ V}$) correspond à la conduction en face arrière (Fig. 81.b). Le pic observé correspond à l'activation de tous les canaux (avant et latéraux). Pour la zone d'inversion face arrière, ce pic diminue par rapport au pic de transconductance observé pour la désertion face arrière. Cette diminution peut s'expliquer par l'effet des résistances séries du canal arrière ou encore par les porteurs minoritaires dans le volume du film, dont la présence peut modifier la capacité de grille apparente [Cristoloveanu'95].

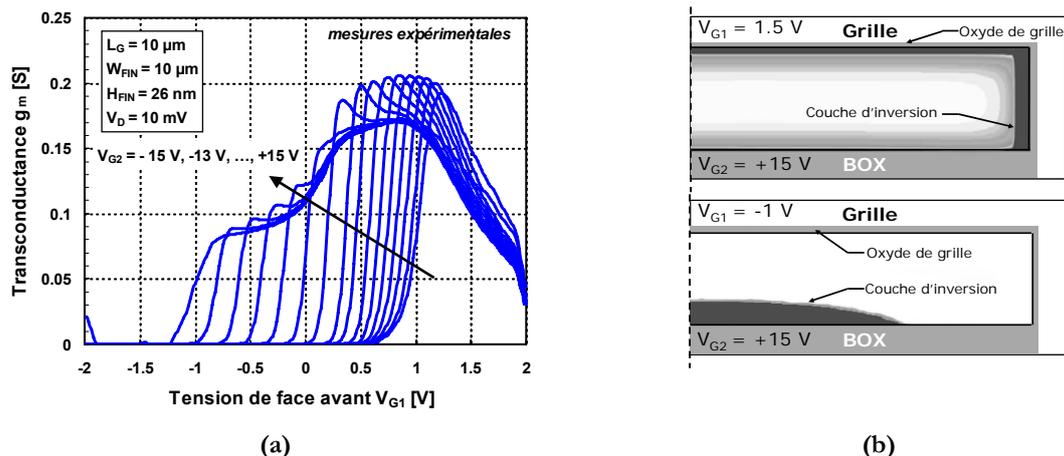


Fig. 81: Dérivée première (transconductance) du courant de drain en fonction de la tension de grille avant pour différentes polarisations de face arrière (a). Coupes transversales représentant la densité électronique dans le canal pour $V_{G1} = -1 \text{ V}$ et $V_{G1} = 1.5 \text{ V}$ (b).

En exploitant la dérivée seconde du courant de drain (Fig. 82.a), les pics correspondant à l'activation des différents canaux (Fig. 82.b) sont directement visibles. On observe le décalage de la tension de seuil correspondant au modèle de Lim et Fossum quand la face arrière est accumulée (zone 1) ou désertée (zone 2). Cependant, en inversant la face arrière deux pics apparaissent; le premier pic correspond à l'activation du canal arrière (zone 4) et le deuxième à l'activation du canal avant (zone 3).

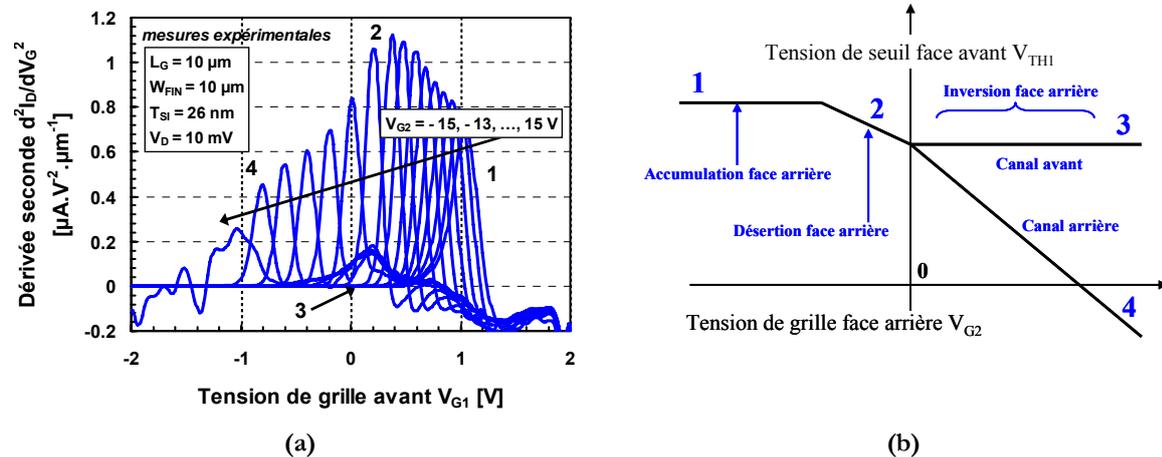


Fig. 82: Dérivée seconde du courant de drain en fonction de la tension de grille face avant pour différentes polarisations de face arrière (a). Représentation schématique des différentes zones de conduction selon la tension de grille face arrière (b).

On retrouve ce comportement en extrayant expérimentalement la tension de seuil en fonction de la polarisation de la face arrière (Fig. 83). Les méthodes d'extraction à courant constant et avec la tension de seuil extrapolée ne permettent d'extraire que le pic face arrière en régime d'inversion face arrière. La méthode de la dérivée seconde permet d'observer les pics correspondant aux activations des canaux avant et arrière. En inversion face arrière, la méthode du ' V_T extrapolé' donne une valeur de la tension de seuil surestimée car la position du maximum de la transconductance (Fig. 81.a) conduit à ignorer en partie l'activation du canal arrière.

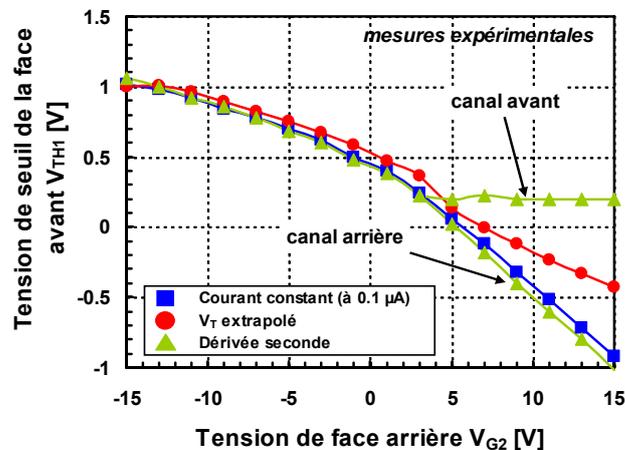


Fig. 83: Tension de seuil face avant en fonction de la tension de grille face arrière avec la méthode d'extraction à courant constant, la méthode extrapolée et la méthode de la deuxième dérivée.

En observant la Fig. 81.a, on peut se demander comment se comporte le courant en régime d'inversion face arrière (là où se trouve le plateau d'inversion). Dans cette zone, on peut exprimer le courant sous la forme [Cristoloveanu'95]:

$$I_{D2} = \frac{\mu_{0,2}}{1 + \theta_{1,2}(V_{G2} - V_{T2}(V_{G1}))} \frac{C_{OX2} W V_D}{L} (V_{G2} - V_{T2}(V_{G1})) \quad \text{Eq. (40)}$$

avec $\theta_{1,2}$ les coefficients d'atténuation de la mobilité.

En utilisant maintenant la relation réciproque de l'Eq. (37), c'est-à-dire qu'on se place dans la zone de désertion face arrière, il vient:

$$V_{T2}^{DEP} = V_{T2}^{ACC} - \gamma_2 (V_{G1} - V_{T2}^{DEP}) \quad \text{Eq. (41)}$$

où γ_2 est le terme représentant le couplage entre face avant et face arrière.:

$$\gamma_2 = \frac{C_{Si} C_{OX1}}{C_{OX2} (C_{OX1} + C_{Si})}; C_{Si} = \frac{\epsilon_{Si}}{T_{Si}}; C_{OX1,2} = \frac{\epsilon_{OX1,2}}{t_{OX1,2}} \quad \text{Eq. (42)}$$

On peut ainsi exprimer le courant de drain en fonction de la tension de grille appliquée sur la face avant:

$$I_{D2} = \frac{\mu_{0,2}}{1 + \theta_{1,2}(V_{G1} - V_{T1}'(V_{G2}))} \frac{\gamma_2 C_{OX2} W V_D}{L} (V_{G1} - V_{T1}'(V_{G2})) \quad \text{Eq. (43)}$$

avec:

$$V_{T1}'(V_{G2}) = V_{G2}^{ACC} + \frac{1}{\gamma_2} (V_{T2}^{ACC} - V_{G2}) \quad \text{Eq. (44)}$$

où $V_{T1}'(V_{G2})$ représente une 'pseudo-tension de seuil' de la face avant.

A la limite entre la désertion et l'inversion de la face arrière, on peut donc définir la valeur du plateau de transconductance:

$$g_{m,PLATEAU} = \frac{\partial I_{D2}}{\partial V_{G1}} = \frac{\gamma_2 \mu_2 C_{OX2} W V_D}{L} \quad \text{Eq. (45)}$$

En mesurant le plateau sur la Fig. 81.b, on calcule la mobilité face arrière:

$$\mu_2 = 301 \pm 60 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$$

On peut comparer cette valeur avec une extraction directe de la mobilité de la face arrière, qui consiste à mesurer le courant du transistor en fonction de la tension de face arrière pour différentes tensions de grille avant (Fig. 84.a). En polarisant la face arrière, l'oxyde enterré joue le rôle d'oxyde de grille et la grille avant va moduler l'état du canal avec un couplage très important. En conséquence, par rapport à la Fig. 80.a, en inversion face avant le transistor conduit toujours, ce qui mène aux importants niveaux de fuites que l'on observe. La transconductance est représentée sur la Fig. 84.b.

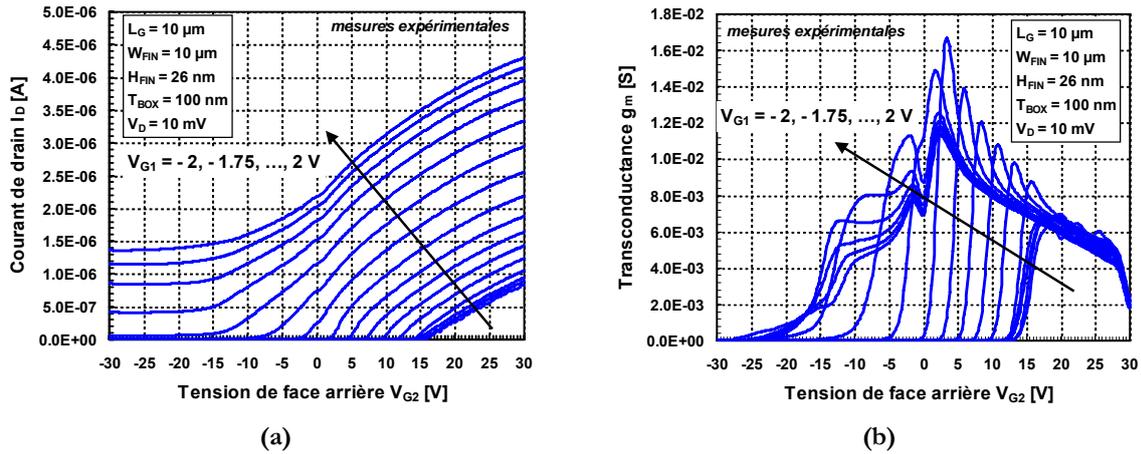


Fig. 84: Courant de drain (a) et transconductance (b) en fonction de la tension de grille face arrière pour différentes polarisations de face avant.

Avec le pic de transconductance de la Fig. 84.b, on calcule directement la mobilité de la face arrière :

$$\mu_{BG,DIRECT} = 395 \pm 50 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$$

Le calcul de la mobilité avec les méthodes directe (μ_2) et indirecte ($\mu_{BG,DIRECT}$) donne des résultats du même ordre de grandeur. La différence obtenue vient de l'imprécision du calcul.

Il existe une différence entre les courbes de transconductance $g_m(V_{G1})$ et $g_m(V_{G2})$. A proximité de $V_{G2} = 0$, on observe systématiquement une chute de la transconductance (Fig. 84.b). Cet effet peut être attribué à la formation d'une zone de désertion dans le substrat à proximité de la tension de bandes plates. En effet, la couche de désertion va former une capacité en série avec la capacité de BOX (Fig. 85.a). La capacité résultante C' sera donc plus faible que C_{OX2} . Les porteurs dans le canal arrière voient ainsi une capacité d'oxyde résultante plus faible, et le courant (donc la transconductance) s'en trouve d'autant diminué.

On peut évaluer la profondeur de la zone de désertion du substrat à partir de la chute de transconductance observée. En considérant l'établissement de deux capacités disposées en série (Fig. 85.a), il vient:

$$W_{DEP,SUB} = \frac{\epsilon_{Si}}{\epsilon_{BOX}} \left(\frac{g_m}{g'_m} - 1 \right) t_{BOX} \quad \text{Eq. (46)}$$

où g_m et g'_m sont respectivement la transconductance sans et avec la désertion du substrat. En mesurant cette chute de transconductance (Fig. 85.b), le calcul conduit à:

$$W_{DEP,SUB} = 141 \pm 30 \text{ nm}$$

qui correspond à un dopage effectif du substrat d'environ $5 \cdot 10^{16} \text{ cm}^{-3}$ (voir chapitre 1, Eq. (11)).

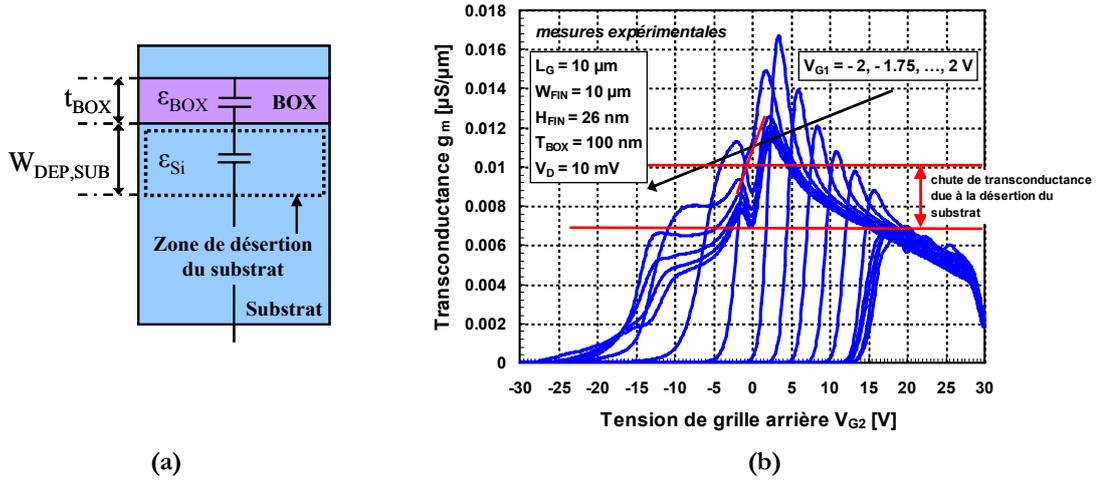


Fig. 85: Schéma de la mise en série de deux capacités en régime de désertion du substrat (a) et chute de transconductance associée (b).

5.2 Cas des transistors étroits

5.2.1 Couplage bidimensionnel

Dans le cas d'un transistor multigrille, le couplage intervenant va dépendre bien sûr du couplage vertical entre la face arrière et la face avant, mais également du couplage entre les grilles latérales. On attend ainsi une variation du coefficient de couplage suivant la largeur du dispositif testé (exemple sur les simulations de la Fig. 86.a).

Récemment, un modèle bidimensionnel pour un transistor long et faiblement dopé a été proposé [Akarvardar'04]. Ce modèle s'appuie sur la résolution de l'équation de Poisson avec les conditions aux limites propres au transistor Triple-grille (Fig. 86.b). Il permet de calculer la variation de la tension de seuil de la grille avant en fonction de la polarisation de la face arrière. Le couplage bidimensionnel avec prise en compte du couplage latéral s'écrit ainsi sous la forme:

$$\alpha(W_{FIN}, H_{FIN}) = \frac{\partial V_{T1}}{\partial V_{G2}} = \left[\frac{2\sqrt{2}}{\sinh(2\sqrt{2} \frac{H_{FIN}}{W_{FIN}})} \right] \frac{\frac{C_W(W_{FIN})}{C_{OX1}}}{1 + \frac{2\sqrt{2}}{\tanh(2\sqrt{2} \frac{H_{FIN}}{W_{FIN}})} \frac{C_W(W_{FIN})}{C_{OX2}}} \quad \text{Eq. (47)}$$

où W_{FIN} est la largeur du dispositif, H_{FIN} sa hauteur, $C_W = \epsilon_{Si}/W_{FIN}$ la capacité latérale du film de silicium, et C_{OX1} et C_{OX2} les différents couplages capacitifs que l'on trouve dans le modèle unidimensionnel.

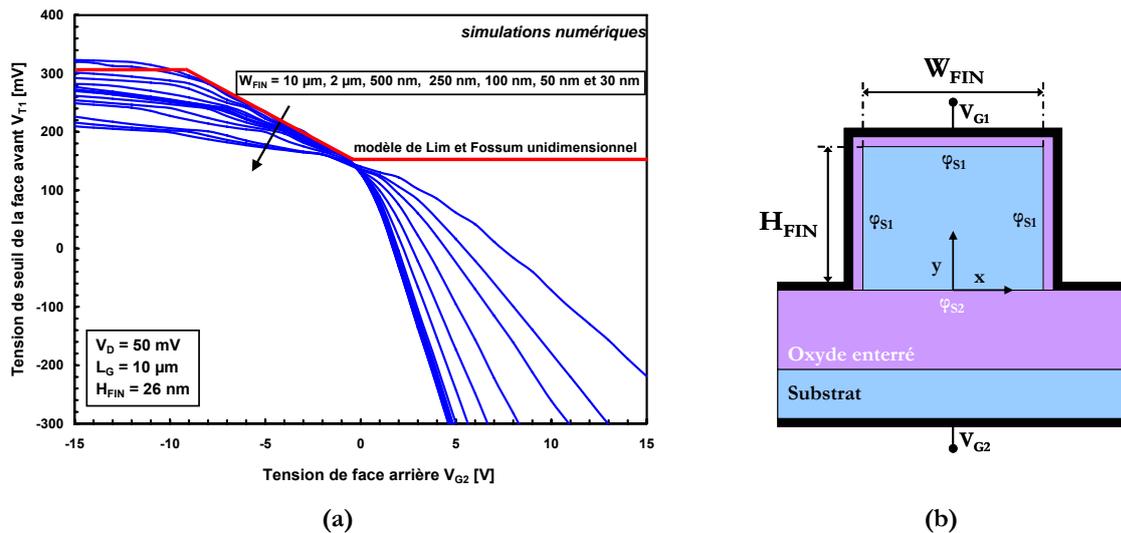


Fig. 86: (a): Simulation numérique de l'évolution de la tension de seuil en fonction de la polarisation de la face arrière pour différentes largeurs de l'aileron. La tension de seuil a été extraite avec la méthode à courant constant. (b): Représentation schématique d'un transistor Triple-grille.

Quand $W_{FIN} \ll H_{FIN}$ (configuration FinFET), le couplage latéral est très fort et le couplage vertical de la face arrière est complètement écranté. Le coefficient de couplage tend donc alors vers 0. Quand $W_{FIN} \gg H_{FIN}$, le coefficient de couplage tend vers le rapport t_{OX1}/t_{OX2} . On est alors dans une configuration planaire 'complètement désertée' et on retrouve le modèle unidimensionnel de Lim et Fossum (Fig. 87). Pour une largeur d'aileron donnée, le coefficient de couplage diminue quand la hauteur de l'aileron augmente (Fig. 88). On peut interpréter cela par le fait que le couplage n'est pas uniforme le long d'un canal vertical et qu'il est maximal à proximité de l'interface silicium/BOX. Par conséquent, en augmentant la hauteur du canal la zone soumise à l'influence du BOX devient de plus en plus faible par rapport au reste du transistor et le coefficient de couplage vertical diminue.

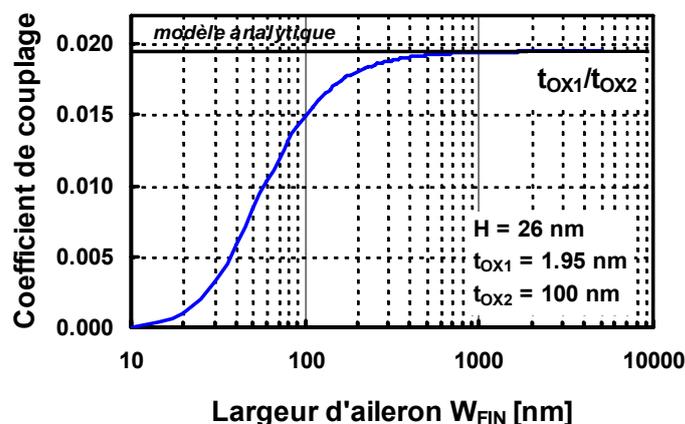


Fig. 87: Coefficient de couplage en fonction de la largeur de l'aileron pour $H_{FIN} = 26 \text{ nm}$.

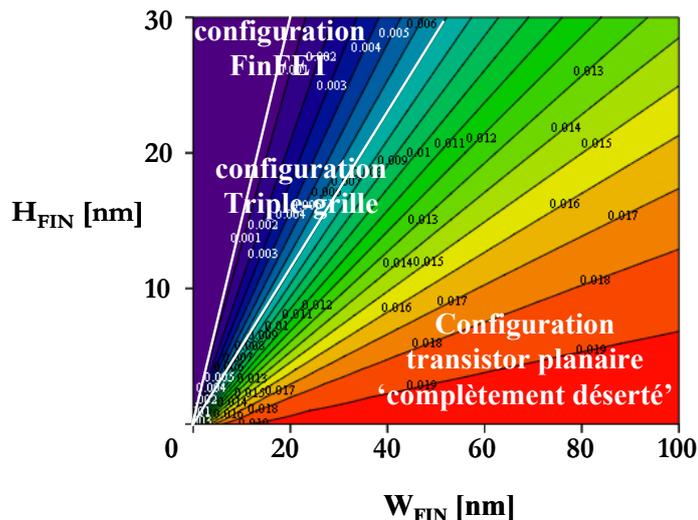


Fig. 88: Tracé de contour du coefficient de couplage bidimensionnel en fonction de la largeur et de la hauteur de l'aileron.

En utilisant une approximation très grossière, on peut également modéliser le couplage d'un transistor PFET en considérant que le couplage vertical réside dans la mise en série de la capacité du silicium et d'une capacité d'oxyde enterré (Fig. 89.a):

$$C_{\Pi} = \frac{C_{Si} C_{OV}}{C_{Si} + C_{OV}} \quad \text{Eq. (48)}$$

avec $C_{Si} = \epsilon_{Si} / H_{FIN}$ et $C_{OV} = \epsilon_{BOX} / t_{OV}$.

En remplaçant la capacité du silicium par cette nouvelle capacité dans l'Eq. (47), on peut calculer le coefficient de couplage d'une structure PFET. Les coefficients calculés analytiquement et extraits par simulation numérique pour les structures Triple-grille et PFET sont comparés dans la Fig. 89.b. On observe les tendances prédites. L'accord est bon pour les faibles largeurs d'aileron, et ce dans le cas des Triple-grilles et des PFET's.

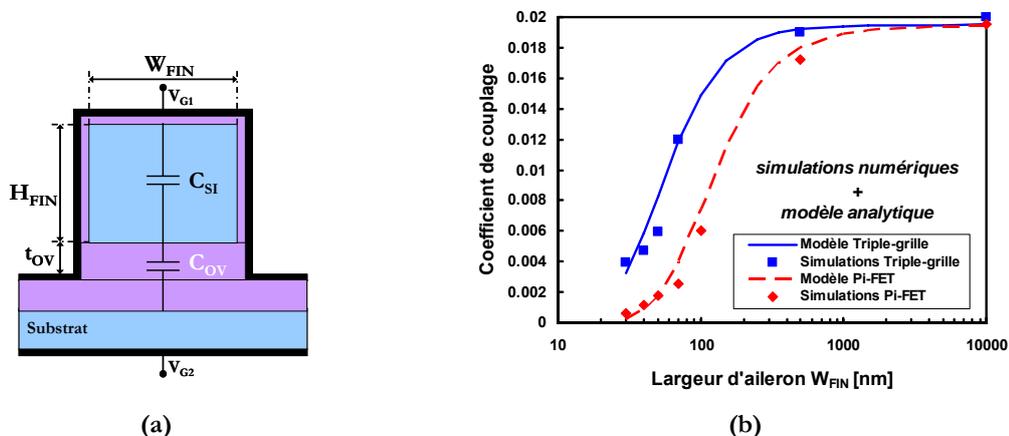


Fig. 89: Représentation schématique d'une structure PFET (a). Comparaison entre le modèle analytique 2D du couplage (Eq. (47)) et le couplage extrait après simulation numérique pour des structures Triple-grille et PFET (b).

5.2.2 Mesures expérimentales sur des transistors Ω FET étroits

5.2.2.1 Ecrantage de l'influence de la face arrière pour des transistors étroits

Dans le cas où la largeur des ailerons W_{FIN} de silicium est réduite, on passe d'un régime de couplage vertical (configuration 'complètement déserté') à un régime de couplage latéral (configuration 'FinFET'). Cette transition graduelle vers une insensibilité à la tension de face arrière est visible sur les caractéristiques électriques (Fig. 90) [Daugé'04] [Préret'03]. En réduisant la largeur de l'aileron, le pic (ou plutôt le plateau ici) de transconductance correspondant à l'activation du canal arrière est de plus en plus faible; le décalage de la tension de grille avant en polarisant différemment la face arrière est également très visible. D'un régime de fort décalage de la tension de seuil ($W_{FIN} = 10 \mu\text{m}$, Fig. 80.a), on passe à un régime intermédiaire ($W_{FIN} = 100 \text{ nm}$, Fig. 90.a) puis à une complète insensibilité à la tension de face arrière ($W_{FIN} = 40 \text{ nm}$, Fig. 90.b).

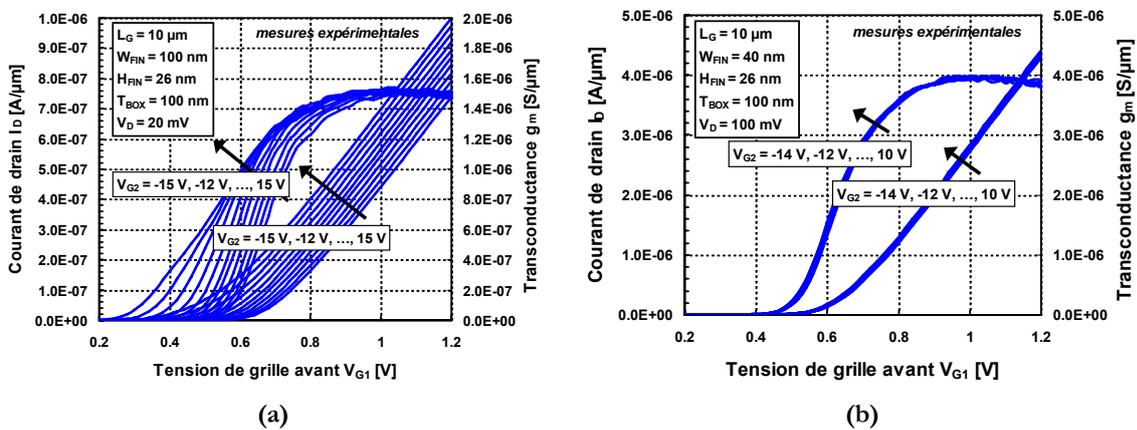


Fig. 90: Courant de drain en fonction de la tension de grille face avant pour différentes polarisations de face arrière pour $W_{FIN} = 100 \text{ nm}$ (a) et 40 nm (b).

On synthétise cette situation en traçant l'évolution de la tension de seuil en fonction de la tension de la grille arrière pour différentes largeurs de grille (Fig. 91). La zone où toutes les courbes se croisent sur la Fig. 91 correspond à la tension $V_{G2,INV}$ où la face arrière s'inverse. Cette tension n'est que peu influencée par le couplage des grilles latérales. Au dessous de $V_{G2,INV}$, on rentre en régime de désertion de la face arrière: on observe le faisceau de droite dont la pente illustre le couplage vertical bidimensionnel; ce couplage diminue bien quand la largeur de l'aileron est réduite. Le plateau correspondant à l'accumulation de la face arrière n'est pas très visible; la tension de seuil de la face avant en accumulation face arrière dépend de la largeur du transistor. La couche d'accumulation ne peut plus moduler la conduction dans l'aileron pour des transistors étroits. Notons que pour cette figure la chute rapide de la tension de seuil en inversion de la face arrière ($V_{G2} > 0 \text{ V}$) est due au fait que l'extraction a été faite avec la méthode 'à courant constant'.

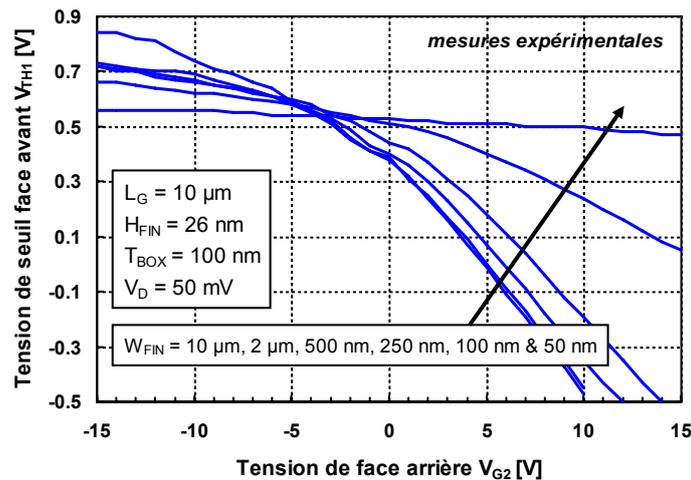


Fig. 91: Mesure expérimentales de l'évolution de la tension de seuil de la face avant en fonction de la tension de la face arrière pour différentes largeurs de grille. L'extraction de la tension de seuil est faite à courant constant.

On retrouve ces tendances sur les simulations numériques de la Fig. 92 (Les écarts entre les valeurs expérimentales et simulées viennent des différences de matériaux utilisés par rapport aux dispositifs expérimentaux (oxyde, grille). Cette figure synthétise la variation de la tension de seuil face avant en fonction de la polarisation de face arrière pour les configurations 'planaire complètement désertée' ($W_{FIN} = 80 \text{ nm}$ et $H_{FIN} = 20 \text{ nm}$), 'Triple-grille' ($W_{FIN} = H_{FIN} = 20 \text{ nm}$) et 'FinFET' ($W_{FIN} = 20 \text{ nm}$ et $H_{FIN} = 120 \text{ nm}$). Le fort couplage vertical (unidimensionnel) prévaut dans la configuration 'planaire complètement désertée' et conduit à une forte variation de la tension de seuil face avant avec la polarisation de la face arrière. Au contraire, en réduisant la largeur d'aileron (configurations 'Triple-grille' et 'FinFET' où $W_{FIN} = 20 \text{ nm}$) le couplage électrostatique devient bidimensionnel et l'influence des grilles latérales réduit le couplage entre les grilles avant et arrière.

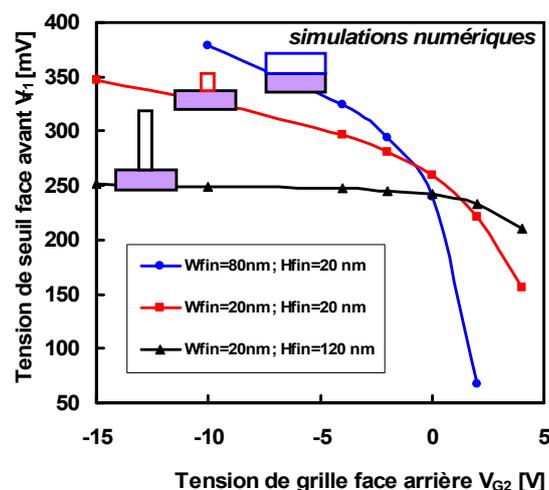


Fig. 92: Simulations numériques montrant l'évolution de la tension de seuil de la face avant en fonction de la tension de la face arrière pour différentes configurations: planaire (cercles, $W_{FIN} = 80 \text{ nm}$ et $H_{FIN} = 20 \text{ nm}$), Triple-grille (carrés, $W_{FIN} = H_{FIN} = 20 \text{ nm}$) et FinFET (triangles, $W_{FIN} = 20 \text{ nm}$ et $H_{FIN} = 120 \text{ nm}$). L'extraction de la tension de seuil est faite à courant constant.

La différence entre les configurations 'Triple-grille' ($W_{\text{FIN}} = H_{\text{FIN}} = 20 \text{ nm}$) et 'FinFET' ($W_{\text{FIN}} = 20 \text{ nm}$ et $H_{\text{FIN}} = 120 \text{ nm}$) est due à la modification du couplage électrostatique avec la hauteur de l'aileron. Sur la Fig. 93, on représente la coupe de la densité électronique simulée d'un transistor Triple-grille où le substrat est accumulé ($V_{G2} = -10 \text{ V}$) et la grille inversée ($V_G = 0.5 \text{ V}$). Les contours de la densité électronique montrent le rôle de la hauteur de l'aileron. Pour une largeur et une hauteur de 20 nm (on est alors en configuration Triple-grille, Fig. 93.a), une forte tension négative crée une couche d'accumulation bloquant l'inversion d'une partie des canaux latéraux. Le couplage de face arrière est fort et module la conduction dans l'aileron. Cependant, pour une faible largeur (ici, 20 nm) et un grand rapport de forme (c'est-à-dire le rapport hauteur/largeur, avec $H_{\text{FIN}} = 120 \text{ nm}$, Fig. 93.b), les effets de couplage du substrat sont restreints à la zone proche du BOX, le canal du haut étant fortement inversé. Le couplage de la face arrière est faible et le transistor est globalement insensible à la tension de face arrière.

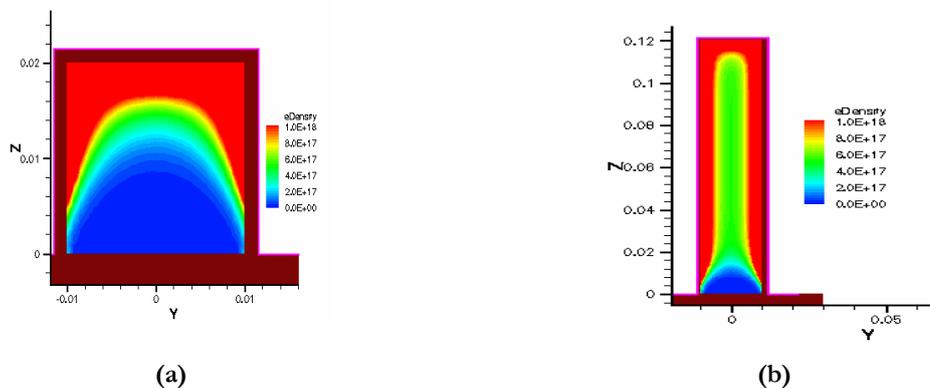


Fig. 93: (a): Coupe transversale du canal montrant la densité électronique pour $V_G = 0.5 \text{ V}$ et $V_{G2} = -10 \text{ V}$ pour $W_{\text{FIN}} = H_{\text{FIN}} = 20 \text{ nm}$. (b): Coupe transversale du canal montrant la densité électronique pour $V_G = 0.5 \text{ V}$ et $V_{G2} = -10 \text{ V}$ pour $W_{\text{FIN}} = 20 \text{ nm}$ et $H_{\text{FIN}} = 120 \text{ nm}$.

Avec la grille pénétrant dans l'oxyde enterré, les géométries de type PIFET et Ω FET vont également modifier le couplage électrostatique. L'examen des différences entre les courbes de couplage obtenues en simulations pour des transistors Triple-grille, PIFET et Ω FET permet d'évaluer quel paramètre est important dans le couplage latéral (Fig. 94). Pour des transistors larges ($W_{\text{FIN}} = 0.5 \mu\text{m}$), les courbes sont évidemment les mêmes. Pour une épaisseur de 100 nm ou 30 nm, les géométries PIFET et Ω FET isolent bien mieux le transistor de l'influence de la face arrière qu'une structure Triple-grille. Le gain en isolation peut donc être attribué au couplage des grilles latérales dans l'oxyde enterré. La différence de comportement entre PIFET et Ω FET va quant à lui dépendre de la pénétration de la grille sous le canal (pour ces simulations, il n'est ici que de 5 nm; le décalage est donc faible). Du point de vue électrostatique, on se situe quelque part entre trois et quatre grilles.

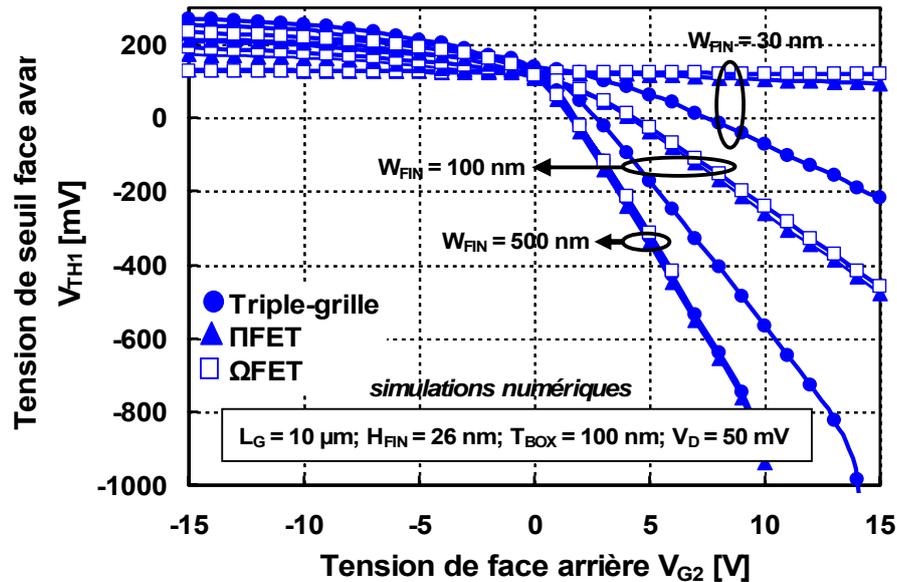


Fig. 94: Simulations numériques de l'évolution de la tension de seuil de la face avant en fonction de la tension de face arrière pour différentes largeurs de grille. Les transistors simulés sont de type Triple-grille (cercles), PFET (triangles) et Ω FET (carrés).

5.2.2.2 Comportement en accumulation face arrière

Sur la Fig. 95.a, les caractéristiques $I_D(V_G)$ pour des transistors Ω FETs de largeur variable ($W_{FIN} = 100$ nm à $10 \mu\text{m}$) sont mesurées. Le substrat est très fortement polarisé en accumulation ($V_{G2} = -30$ V). Pour les transistors larges, on observe deux régions distinctes en dessous du seuil; on n'en observe qu'une seule pour des transistors étroits. L'interprétation de cette différence tient dans le rôle joué par les grilles latérales dans les deux cas. Quelque soit la largeur du dispositif, des pentes sous le seuil comparables sont observables à faible tension de grille. Ceci montre que les courants sous le seuil sont confinés à faible tension de grille à proximité des grilles latérales ou dans les coins. En augmentant la tension de grille avant, on implique également la grille supérieure dans la conduction sous le seuil. Dans un transistor étroit, l'influence de la grille supérieure est faible et n'est donc pas visible. Pour un dispositif large, on peut par contre distinguer le changement d'allure dû au courant additionnel le long du canal supérieur.

Sur la Fig. 95.b, la pente sous le seuil est mesurée en fonction de la polarisation de la face arrière pour un transistor 'complètement déserté', un transistor Ω FET large ($W_{FIN} = 10 \mu\text{m}$) et un transistor Ω FET étroit ($W_{FIN} = 40$ nm). Un transistor 'complètement déserté' est strictement planaire; la dégradation de la pente sous le seuil observée en accumulation de la face arrière est donc due à la perte de contrôle de la grille supérieure. Par contre, les pentes sous le seuil des transistors Ω FET ne semblent pas dépendre de la polarisation de la face arrière; on conserve une pente sous le seuil quasi-idéale dans tous les cas (65 mV/décade à 300 K). La pente sous le seuil ne dépend également pas de la largeur de grille: ceci corrobore le fait que les conductions impliquées à faible V_G sous le seuil se font exclusivement le long des grilles verticales.

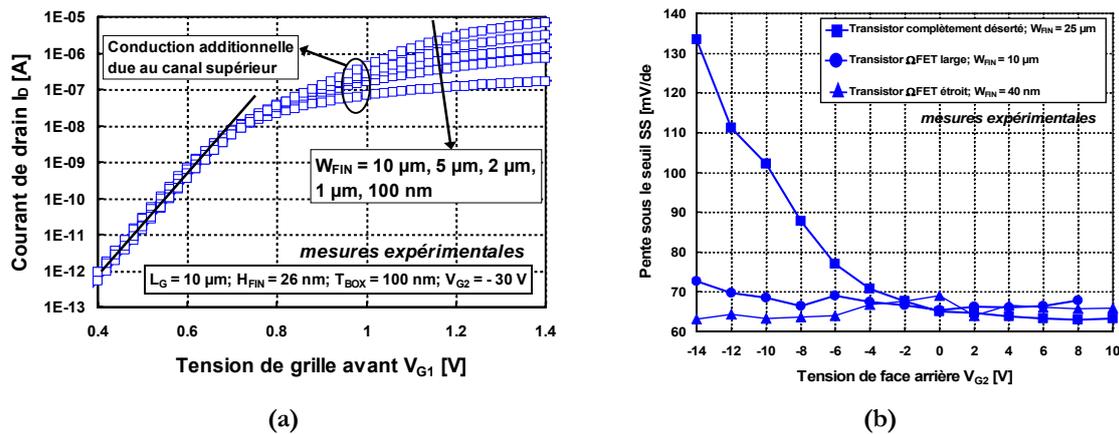


Fig. 95: (a): Courant de drain mesuré en fonction de la tension de face avant V_{G1} pour un transistor Ω FET ayant une largeur de l'aileron W_{FIN} de $10 \mu\text{m}$ à 100nm . Le substrat est fortement polarisé en accumulation ($V_{G2} = -30 \text{V}$). La polarisation de drain V_D est de 50mV , la longueur de grille L_G de $10 \mu\text{m}$ et la hauteur de l'aileron H_{FIN} de 26nm . (b): Pente sous le seuil expérimentale en fonction de la polarisation de la face arrière V_{G2} . Les dispositifs testés sont un transistor large (cercles), un transistor Ω FET étroit (triangles) et un transistor 'complètement déserté' large (carrés). La longueur de grille L_G est de $10 \mu\text{m}$ et la polarisation de drain V_D de 100mV .

5.2.3 Application: la résistance aux radiations

L'application naturelle de l'insensibilité des Ω FETs aux polarisations de face arrière est la tenue aux radiations. En effet, dans un environnement radiatif le mécanisme pouvant perturber le fonctionnement d'un transistor est en général la création de paires électrons/trous le long de la trace de la particule. L'effet principal est l'apparition d'un courant transitoire parasite. Ce phénomène est dépendant de l'architecture utilisée (Silicium massif, PDSOI, FDSOI, ...) et c'est notamment pour réduire la sensibilité à ces effets transitoires que le SOI fut développé. En effet, l'oxyde enterré permet l'isolation diélectrique totale entre deux transistors voisins, supprimant de ce fait le phénomène de 'latch-up' présent dans les transistors sur silicium massif par la mise en conduction d'un thyristor parasite NPNP⁺.

De plus, les particules chargées qui traversent un dispositif génèrent des paires électrons/trous dans le silicium ainsi que dans les oxydes. Dans les structures de type totalement désertées, ils modifient ainsi le potentiel de la face arrière, et donc la tension de seuil du canal avant par couplage électrostatique. En irradiant les transistors avec une certaine dose, on provoque ainsi un décalage de tension de seuil similaire à ce qui se passe en polarisant positivement la face arrière (Fig. 96.a).

Dans le cas des transistors de type Triple-grille, ces effets sont moins connus. Ils ont été mesurés sur les transistors Ω FETs du LETI (mesures M. Gaillardin). Pour un transistor étroit, le décalage obtenu est pratiquement nul. Sur la Fig. 96.b, le décalage de la tension de seuil ΔV_{TH} après irradiation est mesuré pour différentes largeurs d'aileron (de $W_{FIN} = 10 \mu\text{m}$ à $W_{FIN} = 40 \text{nm}$) et pour différentes doses d'irradiation (de 10 à 500krad ; l'unité de mesure de la dose radioactive est le rad, qui équivaut à l'énergie massique absorbée; $1 \text{rad} = 0.01 \text{J.kg}^{-1}$). En

augmentant la dose cumulée d'irradiation, le nombre de charges positives qui vont être piégées dans le BOX augmente. Ainsi, la tension de seuil d'un NMOS Ω FET diminue d'autant plus que la dose d'irradiation est importante. Cependant, cet effet est quasiment supprimé en réduisant la largeur des ailerons. Grâce au couplage latéral dans les transistors Ω FETs, l'effet de la dose cumulée est fortement atténué. Ces transistors étroits sont intrinsèquement très résistants aux radiations. A titre d'ordre de grandeur, la dose annuelle de radiation reçue sur la lune est de 46 rad [ESA], et une bombe à neutrons peut libérer instantanément une dose de 8000 rad [Astrosurf].

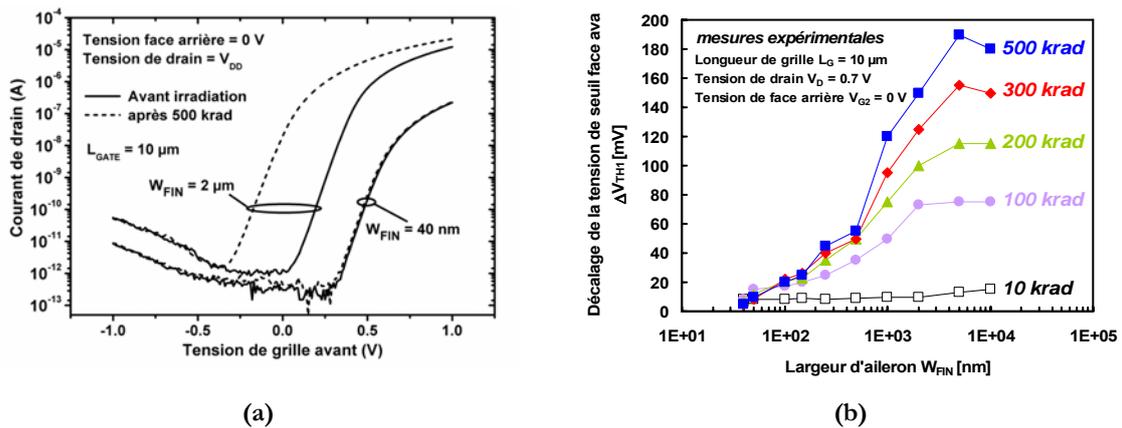


Fig. 96: Courant de drain en fonction de la tension de grille face avant de transistors Ω FET pour une largeur d'aileron de 2 μm et 40 nm, avant (lignes pleines) et après (lignes pointillées) irradiation (a). Décalage de la tension de seuil face avant en fonction de la largeur des ailerons pour différentes doses d'irradiations (b). (d'après [Gaillardin'06])

6 DIVSB (*Drain Induced Virtual Substrate Biasing*)

6.1 Effet DIVSB dans les structures Triple-grille

6.1.1 Définition

Pour un transistor long, la charge désertée dans le canal est essentiellement contrôlée par la grille, et le potentiel de surface le long du canal est pratiquement plat sur l'ensemble de la longueur de grille sauf au voisinage des jonctions où il remonte dans les zones de charges d'espace. Quand on diminue la longueur du canal, la fraction de charge désertée dans le canal contrôlée par la grille diminue au profit de celle contrôlée par les zones de charges d'espace de la source et du drain. Ainsi, le potentiel de surface est de plus en plus contrôlé par les zones de charges d'espace. Ceci conduit à l'abaissement de la barrière d'énergie entre la source et le drain. C'est le principe de l'effet de canal court (*SCE : Short Channel Effect*). De plus, lorsque l'on augmente la polarisation du drain, il y a extension de la zone de charge d'espace au niveau du drain. Ceci provoque une réduction supplémentaire de la barrière d'énergie, c'est l'effet DIBL (*Drain Induced Barrier lowering*).

Dans les structures SOI, le champ électrique issu du drain peut pénétrer dans l'oxyde enterré et le substrat silicium. Ces champs périphériques (*fringing fields*) augmentent le potentiel de surface à l'interface arrière (canal-BOX, Fig. 97). A cause du couplage existant entre les interfaces avant et arrière dans les dispositifs 'complètement désertés' (modèle de Lim et Fossum), la conduction dans le canal avant est dégradée. C'est l'effet DIVSB (*Drain Induced Virtual Substrate Biasing*) [Ernst'99][Ernst'00].

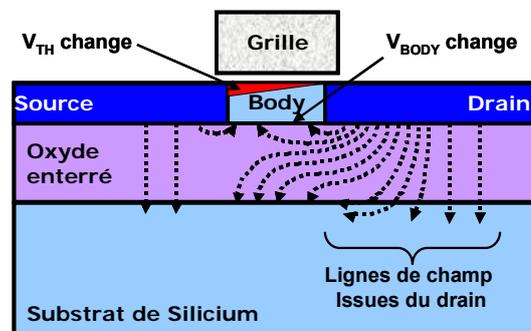


Fig. 97: Représentation schématique des lignes de champ issues du drain conduisant à l'effet DIVSB.

6.1.2 Mesure expérimentale

Expérimentalement, Il est impossible de mesurer le DIVSB seul puisqu'il agit conjointement avec le DIBL. Dans la Fig. 98, le décalage expérimental de la tension de seuil (c'est-à-dire DIBL + DIVSB) est mesuré pour un transistor court ($L_G = 30$ nm) en fonction de la largeur de l'aile et de la polarisation de la face arrière. Pour une polarisation donnée, le 'DIBL + DIVSB' augmente avec la largeur de l'aile à cause du moins bon couplage électrostatique latéral. Quand le substrat est polarisé en accumulation ($V_{G2} = -15$ V), ce décalage de la tension de seuil est minimisé même pour des dispositifs relativement larges

($W_{FIN} = 80 \text{ nm}$): $\Delta V_T/\Delta V_D < 50 \text{ mV/V}$. La couche d'accumulation agit donc en quelque sorte comme une couche protectrice isolant le dispositif de l'effet de la face arrière et du drain. Il joue le même rôle qu'un 'plan de masse' (couche métallique à l'interface substrat/BOX servant à bloquer les variations de potentiel dans cette zone [Ernst'00] [Oshima'04]) et réduit conjointement DIBL et DIVSB.

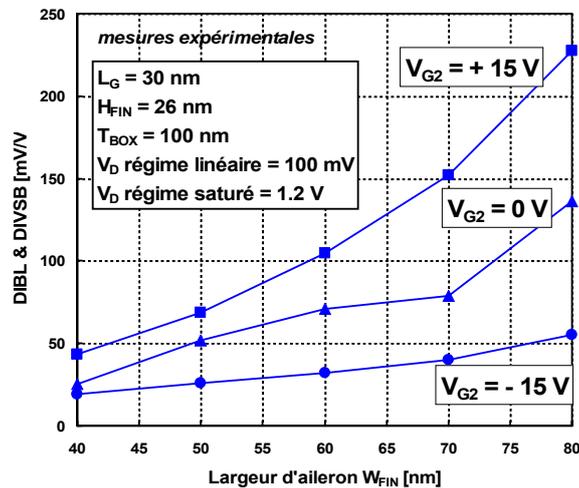


Fig. 98: Mesure expérimentale du décalage de la tension de seuil $\Delta V_T/\Delta V_D$ induite par DIBL et DIVSB en fonction de la largeur de l'aileron pour différentes polarisations de la face arrière (+ 15, 0 et - 15 V, c'est-à-dire de l'inversion à l'accumulation de la face arrière). La mesure a été faite sur un transistor court ($L_G = 30 \text{ nm}$); $V_D = 0.1$ et 1.2 V .

6.1.3 Simulations numériques: effet de la largeur de la grille

Les simulations numériques permettent d'explorer en profondeur le mécanisme du DIVSB. La Fig. 99 montre les résultats obtenus en simulation pour le potentiel dans l'oxyde enterré d'un transistor Triple-grille ($L_G = 100 \text{ nm}$, $W_{FIN} = 30 \text{ nm}$, $V_G = V_D = 1.2 \text{ V}$, $t_{BOX} = 100 \text{ nm}$). Le couplage source/drain dans l'oxyde enterré (Fig. 99.b) a la même allure que ce qu'on observe dans le cas des transistors simple-grille complètement désertés. Cependant, la coupe transversale (Fig. 99.c) effectuée dans l'axe de la grille montre également que le couplage des grilles latérales s'étend à travers l'oxyde enterré.

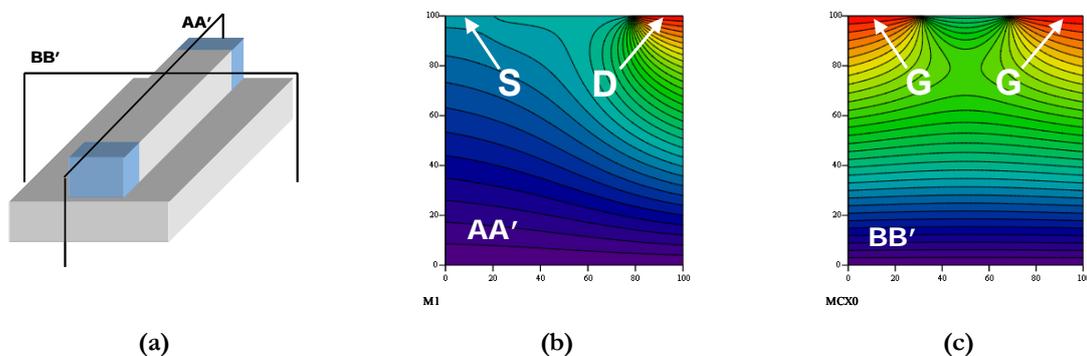


Fig. 99: Représentation d'un potentiel simulé dans l'oxyde enterré d'un transistor Triple-grille (a) selon une coupe longitudinale (b) et transversale (c). La couleur rouge (resp. violette) représente un potentiel élevé (resp. faible).

Dans le cas des dispositifs Triple-grille, les grilles latérales vont donc jouer un rôle supplémentaire en influant sur le potentiel de la face arrière. Sur un dispositif large ($W_{FIN} = 500$ nm), on est proche d'une configuration 'totalement désertée' (*fully depleted*) et l'influence du potentiel de drain est très visible dans l'oxyde enterré. En réduisant la longueur de grille, la zone soumise à l'influence des lignes de champ issues du drain devient proportionnellement plus importante (Fig. 100).

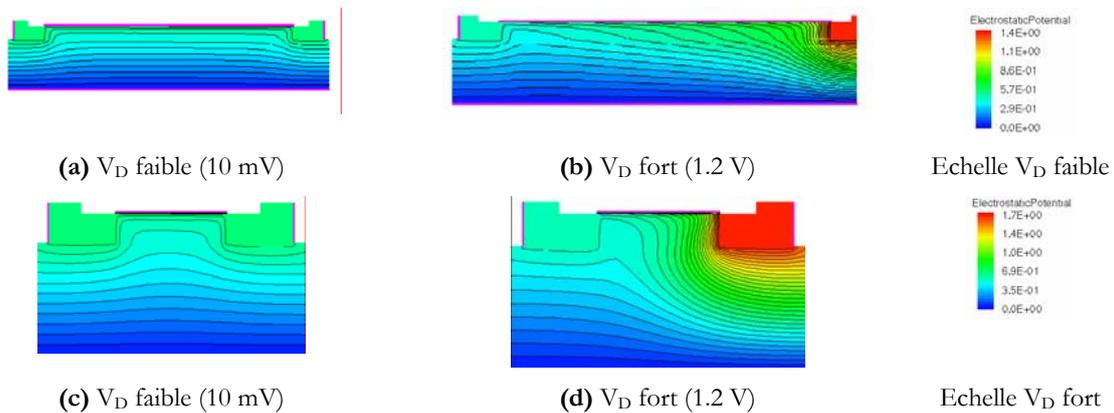


Fig. 100: Coupe de potentiel selon l'axe source/drain de la simulation d'un transistor Triple-grille large et long ($L_G = 500$ nm; (a) & (b)) et d'un transistor large et court ($L_G = 100$ nm; (c) & (d)). $V_G = 1.2$ V, $H_{FIN} = 26$ nm $t_{BOX} = 100$ nm et $W_{FIN} = 500$ nm.

Dans un dispositif Triple-grille étroit, le potentiel issu des grilles latérales a une influence notable sur le potentiel au voisinage de l'interface canal-BOX (Fig. 101). Ainsi, on peut attendre que le couplage des grilles latérales modifie le phénomène de DIVSB, et ce d'autant plus que le canal est étroit.

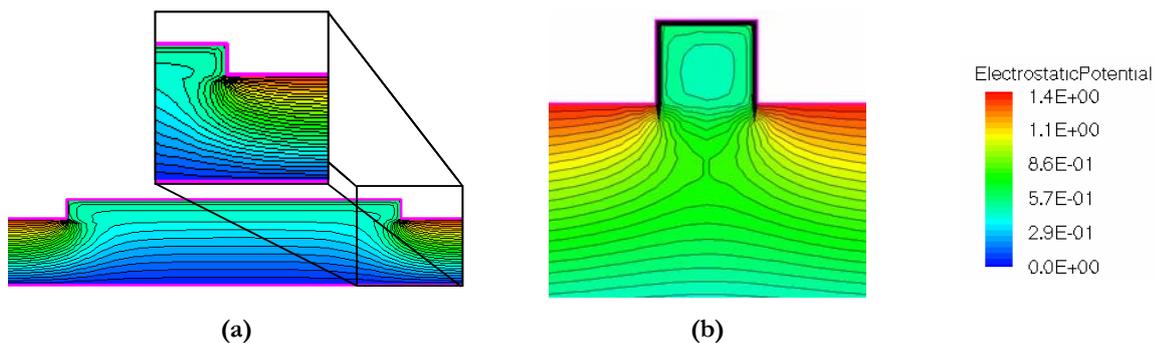


Fig. 101: Coupe transversale du potentiel pris au milieu du canal entre les grilles pour un transistor Triple-grille large ($W_{FIN} = 500$ nm) (a) et pour un transistor Triple-grille étroit ($W_{FIN} = 30$ nm) (b). $L_G = 500$ nm.

Sur la Fig. 102, le potentiel simulé pris au centre du dispositif (à $L_G/2$ et à $W_{FIN}/2$) est représenté en fonction de la profondeur verticale pour différentes configurations de longueurs et de largeur de grille. On polarise le drain à 10 mV et la grille à 1.2 V. De manière attendue, pour toutes les largeurs de grilles, l'écart entre le potentiel pris dans un canal court et dans un

canal long est faible. Cependant, en réduisant la largeur de la grille ($W_{\text{FIN}} = 500 \text{ nm}$, 100 nm puis 30 nm), on observe une remontée de potentiel dans le BOX due aux grilles latérales.

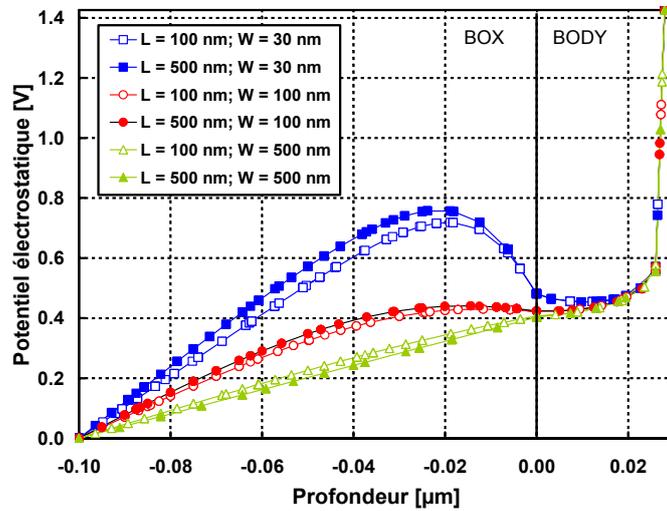


Fig. 102: Coupe de potentiel à V_D faible (10 mV) pour des transistors Triple-grille longs ou courts avec différentes largeurs de grille. $V_G - V_T = 1.2 \text{ V}$.

Sur la Fig. 103, on réalise les mêmes simulations pour une forte polarisation de drain ($V_D = 1.2 \text{ V}$). Pour une même largeur de grille, le potentiel dans le BOX est plus élevé pour un transistor court que pour un transistor long. Le DIVSB est considérablement réduit quand on diminue la largeur de la grille, à cause du couplage latéral.

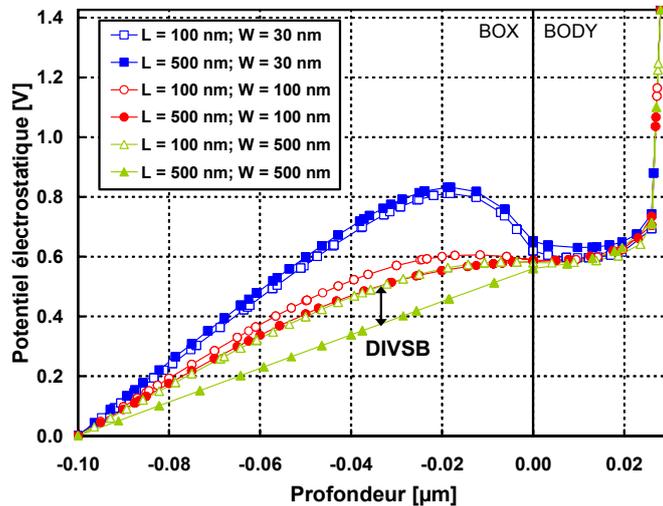


Fig. 103: Coupe de potentiel à V_D fort (1.2 V) pour des transistors Triple-grille longs ou courts avec différentes largeurs de grille. $V_G - V_T = 1.2 \text{ V}$.

En polarisant maintenant faiblement la grille, on réduit le couplage latéral et la modification du potentiel dans le BOX devient très dépendante du potentiel du drain. On observe ce phénomène en comparant la Fig. 104 (V_D faible) avec la Fig. 105 (V_D fort). En

réduisant la longueur de la grille, on arrive ainsi à des variations spectaculaires du potentiel dans l'oxyde enterré.

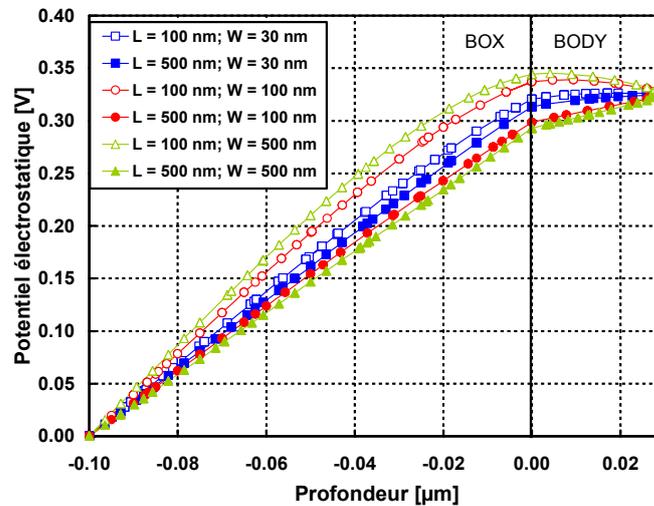


Fig. 104: Coupe de potentiel à V_D faible (10 mV) pour des transistors Triple-grille longs ou courts avec différentes largeurs de grille. $V_G - V_T = 0.1$ V.

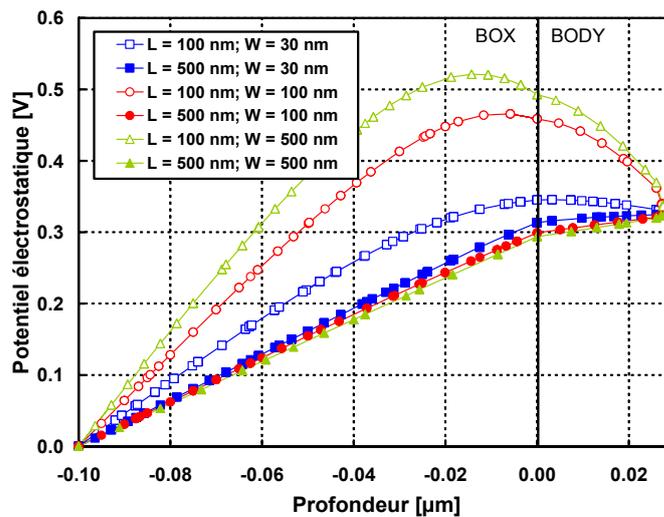


Fig. 105: Coupe de potentiel à V_D fort (1.2 V) pour des transistors longs/courts avec différentes largeurs de grille. $V_G - V_T = 0.1$ V.

On obtient ainsi un résultat important, dû au couplage des grilles latérales. Ce couplage latéral permet visiblement d'isoler la zone d'oxyde enterré située sous le canal. Cet effet est bien entendu amplifié si on utilise une structure Π - ou Ω -FET. Par conséquent, l'effet DIVSB est complètement supprimé dans un transistor Triple-grille. Un transistor Triple-grille sera donc naturellement immunisé contre tout effet de canal court dû au couplage drain/BOX/canal.

6.2 Modélisation analytique du DIVSB

6.2.1 Cas bidimensionnel: transistor 'complètement déserté'

Cette partie reprend les principaux résultats exposés dans la thèse de T. Ernst [Ernst'00] pour des transistors SOI planaires. Pour modéliser le comportement électrostatique dans l'oxyde enterré, il faut résoudre l'équation de Poisson (ou équation de Laplace en l'absence de charges) avec les conditions aux limites imposées par les tensions appliquées à la structure. Afin de résoudre l'équation de Laplace, on peut utiliser une transformation conforme qui est une transformation analytique d'un espace complexe dans lui-même. Toute fonction du plan complexe dans lui-même vérifiant l'équation de Laplace, une géométrie peut ainsi être simplifiée de manière à résoudre le problème électrostatique plus facilement (Fig. 106). Le potentiel de déformation complexe V créé par l'électrode V_1 est alors de la forme:

$$V(x + iy) = \frac{V_1}{i\pi} \left[\ln\left(1 + \exp\left(\pi \frac{x + iy}{t_{BOX}}\right)\right) \right] \quad \text{Eq. (49)}$$

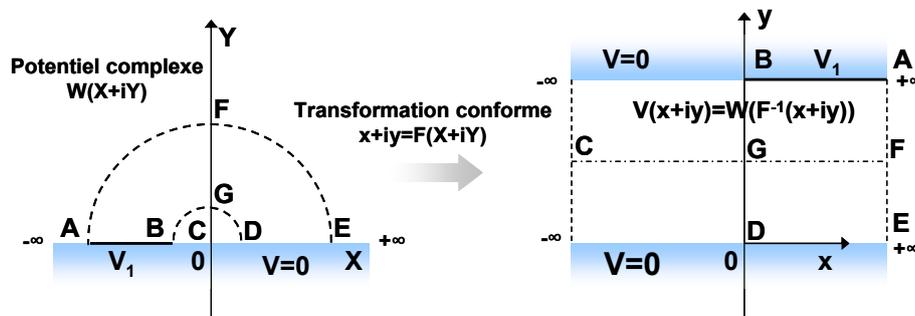


Fig. 106: Représentation simplifiée d'une transformation conforme pour le transistor *FDSOI* planaire.

Le théorème de superposition peut être ensuite utilisé pour décomposer la résolution de l'équation de Laplace en trois problèmes simples (Fig. 107):

$$\psi_{TOT}(x, y, V_S, V_D, V_B, V_{SUB}) = \psi_{SUB}(y, V_B, V_{SUB}) + \psi_S(x, y, V_B, V_S) + \psi_D(x, y, V_B, V_D) \quad \text{Eq. (50)}$$

où V_S , V_D , V_{SUB} et V_B sont respectivement les polarisations de source, de drain, de face arrière et à l'interface silicium/BOX.

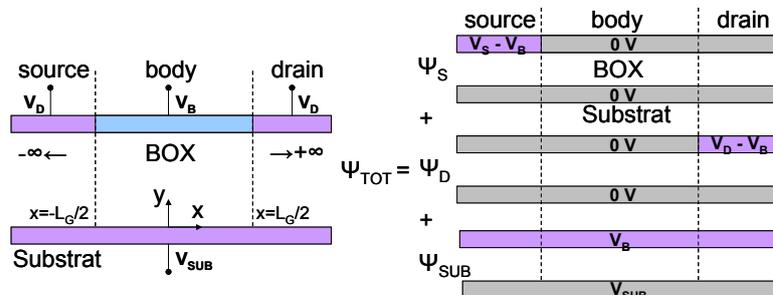


Fig. 107: Théorème de superposition décomposant le problème électrostatique en trois composantes: couplage unidimensionnel entre le canal et le substrat, et deux potentiels de déformation issus des capacités de bord venant de la source et du drain.

En utilisant le formalisme de l'Eq. (49), on peut exprimer les contributions des différentes électrodes dans le plan complexe:

$$\begin{aligned}\psi_D^C(x, y) &= \frac{V_D - V_B}{\pi \cdot i} \ln \left[1 + \exp \left[\frac{\pi}{t_{BOX}} \left((x + iy) - \frac{L}{2} \right) \right] \right] \\ \psi_S^C(x, y) &= \frac{V_S - V_B}{\pi \cdot i} \ln \left[1 + \exp \left[\frac{\pi}{t_{BOX}} \left(-(x - iy) - \frac{L}{2} \right) \right] \right] \\ \psi_{SUB}^C(x, y) &= \frac{V_B - V_{SUB}}{it_{BOX}} (x + iy) + V_{SUB}\end{aligned}\tag{Eq. (51)}$$

L'expression du potentiel de déformation dû au drain se calcule en prenant la partie réelle de l'expression précédente:

$$\begin{aligned}\Psi_D(x, y) &= \operatorname{Re} \left\{ \frac{V_D - V_B}{\pi \cdot i} \ln \left[1 + \exp \left[\frac{\pi}{t_{BOX}} \left((x + iy) - \frac{L}{2} \right) \right] \right] \right\} \\ &= \frac{V_D - V_B}{\pi} \cdot \arctan \left[\frac{\left[\sin \left(\frac{\pi}{t_{BOX}} y \right) \exp \left(\frac{\pi}{t_{BOX}} \left(x - \frac{L}{2} \right) \right) \right]}{\left[1 + \cos \left(\frac{\pi}{t_{BOX}} y \right) \exp \left(\frac{\pi}{t_{BOX}} \left(x - \frac{L}{2} \right) \right) \right]} \right]\end{aligned}\tag{Eq. (52)}$$

Les termes d'influence de la source et de la face arrière peuvent se calculer de la même façon. L'expression en réel n'est cependant valable que sous le canal. On peut obtenir une solution générale en réel en décomposant la fonction par segment. Les capacités associées au drain C_{BD} et à la source C_{BS} (Fig. 108.a) se calculent en dérivant les potentiels obtenus et en appliquant le théorème de Gauss:

$$\begin{aligned}C_{BD} &= \frac{\varepsilon_{BOX}}{t_{BOX}} \frac{1}{\left[\exp \left[\frac{-\pi}{t_{BOX}} \left(x - \frac{L}{2} \right) \right] - 1 \right]} \\ C_{BS} &= \frac{\varepsilon_{BOX}}{t_{BOX}} \frac{1}{\left[\exp \left[\frac{\pi}{t_{BOX}} \left(x + \frac{L}{2} \right) \right] - 1 \right]}\end{aligned}\tag{Eq. (53)}$$

Un des résultats très intéressants de ce formalisme original vient du fait que ces capacités se comportent comme si elles étaient créées par des électrodes virtuelles parallèles à la face arrière (Fig. 108.b). C'est cette propriété qui a donné son nom à l'effet DIVSB. Ces électrodes fictives sont situées à une distance effective t_{BD} (resp. t_{BS}) pour le drain (resp. la source):

$$\begin{aligned}t_{BD} &= t_{BOX} \left[\exp \left[\frac{-\pi}{t_{BOX}} \left(x - \frac{L}{2} \right) \right] - 1 \right] \\ t_{BS} &= t_{BOX} \left[\exp \left[\frac{\pi}{t_{BOX}} \left(x + \frac{L}{2} \right) \right] - 1 \right]\end{aligned}\tag{Eq. (54)}$$

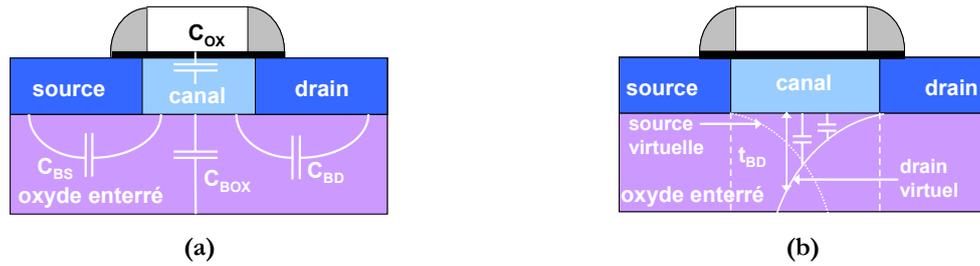


Fig. 108: Représentations schématiques des couplages capacitifs dans l'oxyde enterré (a) et des électrodes virtuelles (b).

6.2.2 Cas tridimensionnel: transistors Triple-grilles

Ce paragraphe propose un élargissement de ce formalisme au cas des transistors Triple-grille. Dans un transistor Triple-grille, la grille contrôle en plus du canal horizontal deux canaux latéraux. Par conséquent, un modèle décrivant le comportement électrostatique dans le BOX doit être tridimensionnel puisqu'il doit prendre en compte simultanément les couplages longitudinaux (selon l'axe source/drain) et verticaux (couplage entre le canal et le BOX) comme dans la section précédente, mais également l'influence des grilles latérales.

La structure simplifiée à modéliser consiste donc en une superposition du problème bidimensionnel (zones de source, drain et canal) avec les grilles latérales perpendiculaires, en contact avec le BOX de chaque côté du canal (Fig. 109).

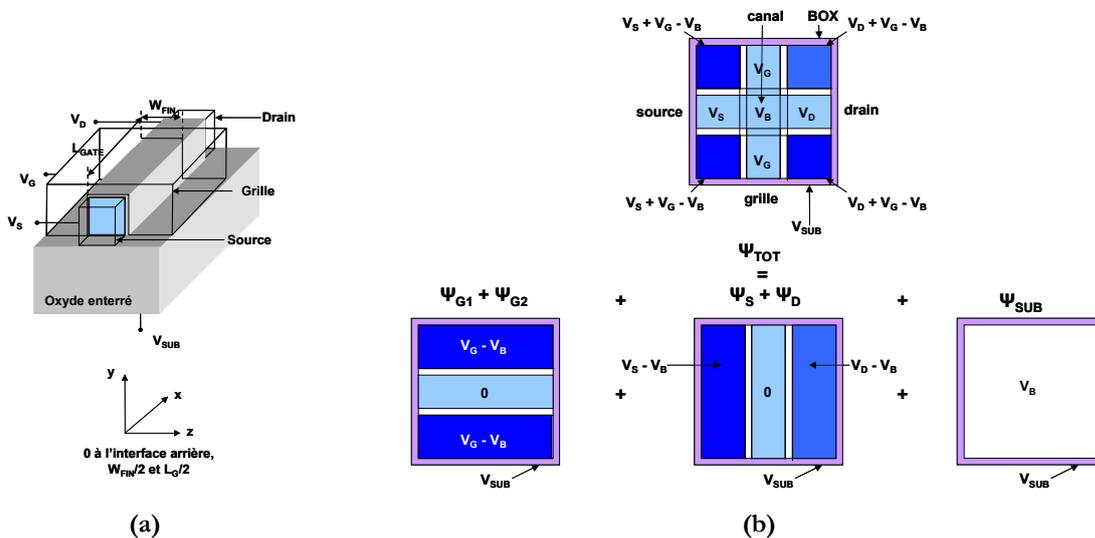


Fig. 109: Représentation schématique d'un transistor Triple-grille (a) et de la structure de modélisation (b).

De la même manière que dans la section précédente, on peut utiliser le théorème de superposition (Fig. 109.b) afin de modéliser le potentiel dans l'oxyde enterré. Il revient alors à prendre en compte les deux termes du potentiel de la grille correspondant aux deux zones où la grille est en contact avec l'oxyde enterré:

$$\begin{aligned} \psi_{TOT}(x, y, z, V_S, V_D, V_G, V_B, V_{SUB}) = & \psi_{SUB}(y, V_B, V_{SUB}) + \psi_S(x, y, V_S, V_B) + \\ & \psi_D(x, y, V_D, V_B) + \psi_{G1}(y, z, V_G, V_B) + \psi_{G2}(y, z, V_G, V_B) \end{aligned} \quad \text{Eq. (55)}$$

Les expressions de ψ_{SUB} , ψ_S et ψ_D sont directement données par l'Eq. (51). Les deux potentiels de grille s'obtiennent de la même manière que précédemment:

$$\begin{aligned} \psi_{G1}^C(y, z) = & \frac{V_G - V_B}{\pi \cdot i} \ln \left[1 + \exp \left[\frac{\pi}{t_{BOX}} \left((y + iz) - \frac{W}{2} \right) \right] \right] \\ \psi_{G2}^C(y, z) = & \frac{V_G - V_B}{\pi \cdot i} \ln \left[1 + \exp \left[\frac{\pi}{t_{BOX}} \left(-(y - iz) - \frac{W}{2} \right) \right] \right] \end{aligned} \quad \text{Eq. (56)}$$

où $q = z + iy$ est la coordonnée complexe et W la largeur de l'aileron.

Les capacités de bord associées à ce couplage canal/oxyde enterré/grille s'expriment comme dans l'Eq. (53):

$$\begin{aligned} C_{BG1} = & \frac{\epsilon_{BOX}}{t_{BOX}} \frac{1}{\left[\exp \left[\frac{-\pi}{t_{BOX}} \left(z - \frac{W}{2} \right) \right] - 1 \right]} \\ C_{BG2} = & \frac{\epsilon_{BOX}}{t_{BOX}} \frac{1}{\left[\exp \left[\frac{\pi}{t_{BOX}} \left(z + \frac{W}{2} \right) \right] - 1 \right]} \end{aligned} \quad \text{Eq. (57)}$$

Ce modèle analytique a été comparé aux simulations numériques réalisées avec DESSIS 3D. Il est tout d'abord appliqué à des configurations 'quasi bidimensionnelles': dispositif long ($L_G = 500$ nm) et étroit ($W_{FIN} = 30$ nm) ou court ($L_G = 100$ nm) et large ($W_{FIN} = 500$ nm). Le potentiel électrostatique est extrait au centre du dispositif (à $L_G/2$ et $W_{FIN}/2$) et est représenté en fonction de la profondeur dans le canal dans la Fig. 110. Pour une configuration 'longue et étroite' (carrés de la Fig. 110), la bosse de potentiel dans le BOX apparaissant n'est pas due au couplage de drain mais au couplage des grilles latérales. Pour une configuration 'courte et large' (cercles de la Fig. 110), les grilles latérales ne jouent pas de rôle. La bosse observée dans l'oxyde enterré est donc due au potentiel de déformation du drain; c'est l'effet *DIV/SB*. Dans les deux cas 'quasi 2D', les potentiels sont parfaitement modélisés.

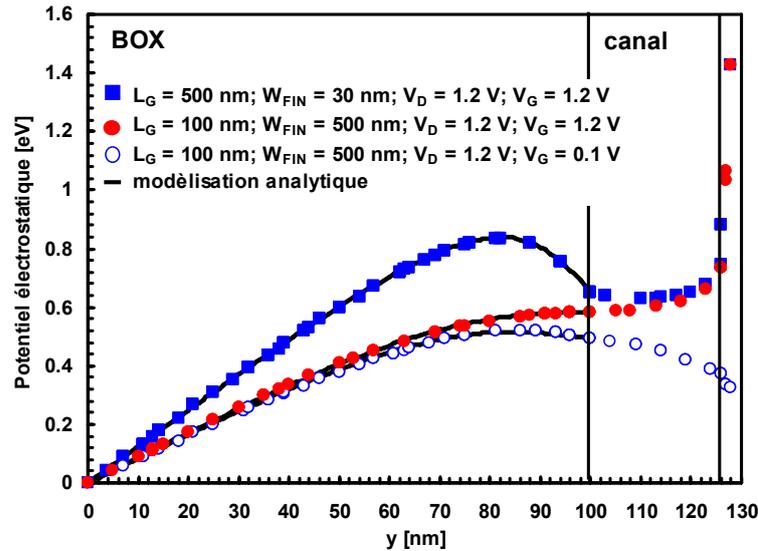


Fig. 110: Comparaison entre le potentiel électrostatique au centre du dispositif (i.e. à $L_G/2$ et $W_{FIN}/2$) obtenu avec le modèle analytique et les simulations numériques dans le cas d'une configuration "quasi-bidimensionnelle". Les carrés représentent la simulation numérique obtenue pour un transistor long ($L_G = 500$ nm) et étroit ($W_{FIN} = 30$ nm) avec $V_D = 1.2$ V et $V_G = 1.2$ V. Les cercles correspondent à un transistor court ($L_G = 100$ nm) et large ($W_{FIN} = 500$ nm) avec $V_D = 1.2$ V et $V_G = 1.2$ V (cercles pleins) et $V_D = 1.2$ V et $V_G = 0.1$ V (cercles ouverts). Les traits pleins représentent les modèles analytiques correspondant à chaque situation.

Dans le cas d'une configuration réellement tridimensionnelle ($L_G = 100$ nm, $W_{FIN} = 30$ nm, $t_{si} = 25$ nm, Fig. 111), le potentiel sous l'interface canal/oxyde enterré est bien contrôlé par les grilles latérales. Par conséquent, le potentiel présente une bosse pour une forte tension de grille ($V_G = 1.2$ V) et est à peu près plat pour une faible tension de grille ($V_G = 0.1$ V). Les simulations montrent que le profil de potentiel n'est pas altéré de manière significative en augmentant la tension de drain de 10 mV à 1.2 V. Cela confirme que l'effet DIVSB est bien supprimé par le couplage issu des grilles latérales dans le cas d'un transistor Triple-grille étroit.

Le modèle analytique correspondant (Fig. 111) reproduit assez fidèlement la bosse de potentiel dans l'oxyde enterré. Cependant, un désaccord quantitatif entre le modèle et les simulations est observé, en particulier quand les tensions de drain et de grille sont élevées (cercles pleins de la Fig. 111). Systématiquement, on observe que le modèle prédit un potentiel plus élevé que ce qui est observé. Ce désaccord est causé par le théorème de superposition (Eq. (55)). En effet, le passage d'un modèle 2D à un modèle 3D provoque la génération de quatre électrodes parasites lorsque l'on applique la transformation conforme. Ces électrodes parasites ont un potentiel plus élevé que le potentiel de drain ou de grille. Par conséquent, le modèle conduit à surestimer le potentiel dans le canal.

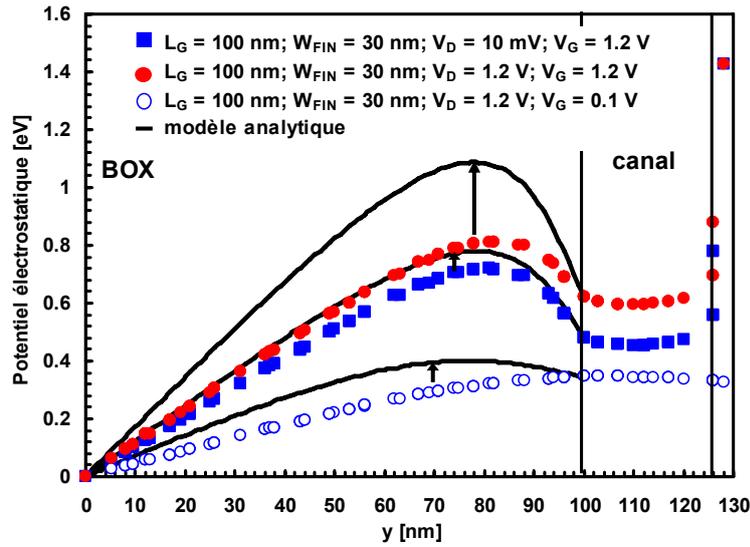


Fig. 111: Comparaison entre le potentiel électrostatique au centre du dispositif (i.e. à $L_G/2$ et $W_{FIN}/2$) obtenu avec le modèle analytique et les simulations numériques dans le cas d'une configuration tridimensionnelle ($L_G = 100$ nm, $W_{FIN} = 30$ nm). Les carrés représentent la simulation numérique obtenue pour $V_D = 10$ mV et $V_G = 1.2$ V, les cercles pleins correspondent à $V_D = 1.2$ V et $V_G = 1.2$ V et les cercles ouverts à $V_D = 1.2$ V et $V_G = 0.1$ V. Les traits pleins représentent les modèles analytiques correspondant à chaque situation.

La solution pour contrer ce phénomène d'électrodes parasites consiste à évaluer le rayon d'action de chaque grille à l'aide de paramètres correspondant à la configuration géométrique (Fig. 112.a). Des coefficients de pondération peuvent ainsi être appliqués aux potentiels de l'Eq. (55):

$$\psi_{TOT}(x, y, z, V_S, V_D, V_G, V_B, V_{SUB}) = \psi_{SUB}(y, V_B, V_{SUB}) + \frac{\alpha}{\pi} \left[2 \arctan\left(\frac{W}{L}\right) [\psi_S(x, y, V_S, V_B) + \psi_D(x, y, V_D, V_B)] + 2 \arctan\left(\frac{L}{W}\right) [\psi_{G1}(y, z, V_G, V_B) + \psi_{G2}(y, z, V_G, V_B)] \right] \quad \text{Eq. (58)}$$

avec α un paramètre d'ajustement.

Pour les cas quasi 2D exposés précédemment (Fig. 110), le paramètre d'ajustement géométrique est égal à 1 et l'Eq. (57) se réduit à:

$$\begin{aligned} \psi_{TOT} &= \psi_{SUB} + \psi_S + \psi_D \quad \text{si } W \gg L \\ \psi_{TOT} &= \psi_{SUB} + \psi_{G1} + \psi_{G2} \quad \text{si } L \gg W \end{aligned} \quad \text{Eq. (59)}$$

On retrouve ainsi les deux problèmes bidimensionnels exposés plus haut. Dans le cas du transistor court et étroit où le problème des électrodes parasites se posait, l'application de l'Eq. (58) avec $\alpha = 1$ permet d'avoir un bon accord entre le modèle analytique et les simulations numériques (Fig. 112.b). Avec $\alpha = 1.15$, on obtient un accord parfait.

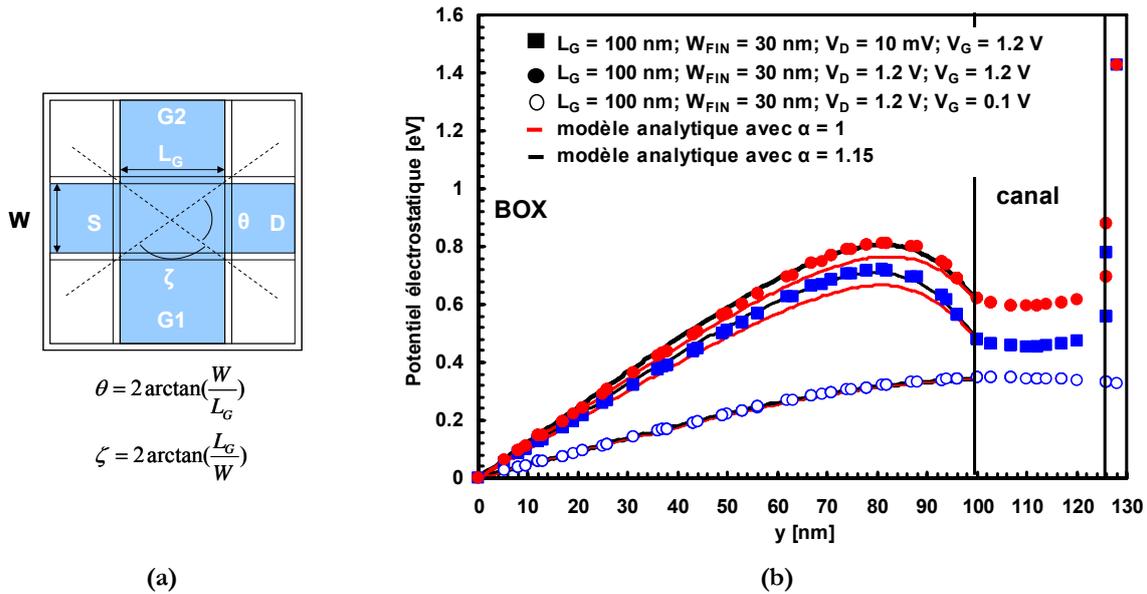


Fig. 112: (a): Détermination géométrique des facteurs de pondération. (b): Comparaison entre le potentiel électrostatique au centre du dispositif (i.e. à $L_G/2$ et $W_{FIN}/2$) obtenu avec le modèle analytique et les simulations numériques dans le cas d'une configuration tridimensionnelle ($L_G = 100$ nm, $W_{FIN} = 30$ nm). Les carrés représentent la simulation numérique obtenue pour $V_D = 10$ mV et $V_G = 1.2$ V, les cercles pleins correspondent à $V_D = 1.2$ V et $V_G = 1.2$ V et les cercles ouverts à $V_D = 1.2$ V et $V_G = 0.1$ V. Les traits pleins représentent les modèles analytiques correspondant à chaque situation avec l'application des facteurs de pondération ($\alpha = 1.15$ et $\alpha = 1.15$).

6.3 Discussion

On peut se poser la question de l'impact d'un phénomène comme le DIVSB sur les performances électriques d'un transistor de type complètement déserté ultra-court. Cependant, il est impossible lors d'une mesure électrique de décorréler le DIBL du DIVSB. En simulant une structure Triple-grille avec une épaisseur de film variable, on peut cependant extrapoler la valeur du DIVSB (Fig. 113). En effet, une réduction de l'épaisseur du film revient à réduire le *DIBL* tout en augmentant le *DIVSB*. En extrapolant la valeur extraite à H_{FIN} nul, on obtient une valeur proche de 20 mV/V. Cette valeur est beaucoup plus faible que le DIBL de manière générale. Le couplage latéral issu des grilles verticales (pour $L_G = 100$ nm et $W_{FIN} = 30$ nm) permet également de réduire conjointement DIBL et DIVSB, conduisant à une quasi-insensibilité à ces effets électrostatiques à condition que l'épaisseur du film ne soit pas trop faible. Encore une fois, on arrive donc à la relaxation des contraintes d'échelle pour les transistors verticaux de la famille des FinFETs.

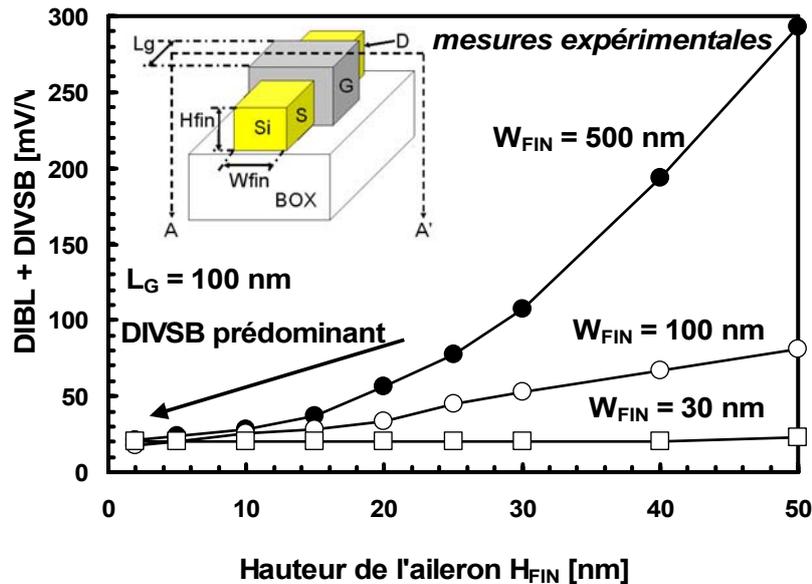


Fig. 113: Simulation pour une structure Triple-grille de la somme DIBL + DIVSB en fonction de la hauteur d'aileron pour différentes largeurs d'aileron ($L_G = 100$ nm).

La manière de réduire le DIVSB consiste à atténuer au maximum le couplage drain/BOX/canal. Un BOX fin semble la meilleure manière d'y arriver. On peut également penser à utiliser un oxyde enterré avec une faible permittivité diélectrique, ou encore une architecture de type 'plan de masse' qui permet d'évacuer les lignes de champ issu du drain vers l'interface BOX-substrat [Oshima'04] [Bresson'05].

Par contre, l'autoéchauffement (ou *self heating*) du dispositif est également à considérer. En effet, si la chaleur générée par effet Joule lors du fonctionnement statique du transistor ne se dissipe pas la température dans le canal augmente. Cette augmentation de température conduit à une chute de mobilité, réduisant ainsi le courant de sortie du transistor. Pour de faibles sections de silicium et la forte résistance thermique des oxydes, cela pourrait s'avérer très gênant. C'est ainsi que de nouveaux types d'oxydes enterrés à faible résistance thermique et de faible épaisseur sont actuellement en cours d'élaboration [DeBeaumont'05]. On peut citer par exemple l'alumine Al_2O_3 . Par contre, ces matériaux '*high-k thermique*' sont également '*high-k électriques*' [Oshima'04]. Une forte diminution de l'autoéchauffement pourrait ainsi aller de pair avec une augmentation du DIVSB ! A condition que la dissipation de la chaleur à travers l'oxyde enterré reste acceptable, une solution consisterait donc à utiliser des BOX très fins ou à coupler cette solution 'BOX alternatif' avec une architecture multigrille qui supprime naturellement le DIVSB.

7 Conclusions

Dans ce chapitre, nous avons abordé la question des couplages capacitifs dans les structures Triple-grille. Ces couplages sont relativement novateurs à cause de la structure géométrique des Triple-grille. Des mesures électriques et simulations effectuées, quelques conclusions peuvent être tirées:

- ✓ Un bon respect des règles d'échelle permet de continuer à réduire les dimensions des dispositifs sans dégradation des performances. Evidemment, il faut pour cela que ce soit techniquement possible et reproductible.
- ✓ Un bon Triple-grille est un Triple-grille dont les flancs sont verticaux et la largeur étroite.
- ✓ Les effets de coin induisent un décalage de la tension de seuil assez faible même en utilisant un canal intrinsèque.
- ✓ Les effets de couplage latéraux, déjà mis à contribution dans le maintien des performances électriques, isolent complètement le canal du substrat. Par conséquent:
 - Les effets de face arrière sont très limités dans les transistors Triple-grille. Ces transistors sont ainsi naturellement très résistants aux radiations (confirmé par des mesures expérimentales).
 - L'effet du couplage du drain (DIVSB) est également supprimé pour des canaux étroits. Cet effet a été démontré par des simulations numériques, et modélisé analytiquement.

Références du chapitre 4

- [Akarvardar'04] K. Akarvardar, S. Cristoloveanu, P. Gentil, "Threshold Voltage Model of the SOI 4-Gate Transistor", *IEEE Int. SOI conf.* 2004, pp. 89-90, 2004.
- [Ancona'87] M. G. Ancona, H. F. Tiersen, "Macroscopic physics in the silicon inversion layer", *Phys. Rev. B*, vol. 35, no. 15, pp. 7959-7965, 1987.
- [Ancona'89] M. G. Ancona, G. J. Iafrate, "Quantum correction to the equation of state of an electron gas in a semiconductor", *Phys. Rev. B*, vol. 39, no. 13, pp. 9536-9540, 1989.
- [Astrosurf] <http://www.astrosurf.com>, site dédié à l'astronomie amateur.
- [Balestra'87] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini and T. Elewa, "Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance", *IEEE Electron Device Letters*, vol. EDL-8, pages 410-412, 1987.
- [Bresson'05] N. Bresson, S. Cristoloveanu, C. Mazuré, F. Letertre, H. Iwai, "Integration of buried insulators with high thermal conductivity in SOI MOSFETs: Thermal properties and short channel effects", *Solid-State Electronics*, vol. 49, no. 9, pp. 1522-1528, 2005.
- [Burenkov'03] A. Burenkov, J. Lorenz, "Corner Effect in Double and Triple Gate FinFETs", proceedings of the 33th European Solid-State Device Research conference (*ESSDERC'03*), pp. 135-138, 2003.
- [Colinge'04] J.-P. Colinge, "Multiple-gate SOI MOSFETs", *Solid-State Electronics*, vol. 48, no. 6, pp. 897-905, 2004.
- [Colinge'06] J.-P. Colinge, J.C. Alderman, W. Xiong, C.R. Cleavelin, "Quantum-mechanical effects in trigate SOI MOSFETs", *Electron Devices, IEEE Transactions on*, vol. 53, no. 5, pp. 1131-1136, 2006.
- [Cristoloveanu'95] S. Cristoloveanu and S.S. Li, "electrical characterization of silicon-on-insulator materials and devices", Kluwer Academic Publishers, p. 240, 1995.
- [Cristoloveanu'04] S. Cristoloveanu, V. Ferlet-Cavrois, "Introduction to SOI MOSFETs: Context, Radiation Effects, and Future Trends", *International Journal of High Speed Electronics and Systems*, vol. 14, no. 2, pp. 465-487, 2004.

- [Daugé'04] F. Daugé, J. Pretet, S. Cristoloveanu, A. Vandooren, L. Mathew, J. Jomaah, B.-Y. Nguyen, "Coupling effects and channels separation in FinFETs", *Solid-State Electronics*, vol. 48 no. 4, pp. 535-542, 2004.
- [DeBeaumont'05] C. de Beaumont, H. Moriceau, O. Rayssac, N. Bresson, S. Cristoloveanu, A.M. Charvet, "ALD alumina films as buried dielectric layers for SOI structures", 207th meeting of the Electrochemical Society, *Silicon-on-Insulator technology and devices session XII*, pp. 231-236, 2005.
- [Ernst'99] T. Ernst, S. Cristoloveanu, "Buried oxide fringing capacitance: a new physical model and its implication on SOI device scaling and architecture", *IEEE Int. SOI Conf.*, pp. 38-39, 1999.
- [Ernst'00] T. Ernst, "Etude des structures MOSFET avancées sur SOI pour les applications basse consommation", *thèse de doctorat*, INPG, 2000.
- [ESA] <http://sci.esa.int>, site de l'agence spatiale européenne.
- [Gaillardin'06] M. Gaillardin, P. Paillet, V. Ferlet-Cavrois, S. Cristoloveanu, O. Faynot, and C. Jahan, "High tolerance to total ionizing dose of Ω -shaped gate field-effect transistors", *Appl. Phys. Lett.* 88, 223511, 2006.
- [Greiner'94] W. Greiner, "Mécanique quantique: une introduction", 3^{ème} édition, *Ed. Springer*, ISBN 3-540-64347-8, 1994.
- [ISE] "ISE TCAD Release 9.0, DESSIS user documentation", *Device Simulation*, vol. 4.a.
- [Lim'83] H.K. Lim and J.G. Fossum, "Threshold voltage of thin-film silicon-on-insulator (SOI) MOSFETs", *IEEE Transactions on Electron Devices*, vol. 30, p1244, 1983.
- [Mathieu'96] H. Mathieu, "Physique des semiconducteurs et des composants électroniques", 3^{ème} édition, *Ed. Masson*, ISBN 2-225-85124-7, 1996.
- [Oshima'04] K. Oshima, "Solutions technologiques avancées pour CMOS ultime: grille métal damascène, diélectriques *high-k*, SOI avec film mince, alumine enterré et plan de masse", *thèse de doctorat*, INPG, 2004.
- [Poiroux'05] T. Poiroux, M. Vinet, O. Faynot, J. Widiez, J. Lolivier, T. Ernst, B. Previtali, S. Deleonibus, "Multiple gate devices: advantages and challenges", *Microelectronic Engineering*, vol. 80, pp. 378-385, 2005.

- [Prétet'03] J. Pretet, F. Daugé, A. Vandooren, L. Mathew, B. Y. Nguyen, J. Jomaah, S. Cristoloveanu, “Substrate effects in SOI FinFETs,” in *Proc. Electrochem. Soc.*, vol. 2003-05, pp. 231–236, 2003.
- [Städele'04] M. Städele, R. J. Luyken, M. Roosz, M. Specht, W. Rösner, L. Dreeskornfeld, J. Hartwich, F. Hofmann, J. Kretz, E. Landgraf, L. Risch, “A comprehensive study of corner effects in tri-gate transistors”, proceedings of the *34th European Solid-State Device Research conference (ESSDERC'04)*, pp. 165-168,2004.
- [Sze'81] S. M. Sze, “Physics of Semiconductor Devices (2nd ed.)”, *John Wiley and Sons (WIE)*. ISBN 0-47-105661-8, 1981.
- [Toyabe'04] T. Toyabe, “Two and three dimensional MOSFETs simulation with density gradient model”, *Proceedings of the fourth international workshop on junction technology IWJT'04*, pp. 317-320.

Chapitre 5

Mesures de mobilité dans les transistors Ω FETs – influence des plans cristallins

Dans ce chapitre, les aspects propres aux phénomènes de transport dans les structures FinFET sont abordés. De par leur structure non planaire, les transistors Ω FETs ont des flancs verticaux dont l'orientation cristalline dépend de la direction cristalline du transport.

Des mesures systématiques de mobilité pour différentes directions cristallines, orientations cristallines, largeurs et longueurs de grille sont effectuées. On observe une chute de la mobilité avec la réduction de la longueur de grille. Les mesures expérimentales montrent aussi des résultats parfois inattendus: on ne retrouve pas les effets prévus théoriquement en modulant la largeur de l'aile ou en changeant la direction cristalline du dispositif. On observe également une amélioration de la mobilité pour les transistors étroits dans le cas des NMOS et PMOS, ce qui va à l'encontre des effets théoriquement attendus. Nous proposons une explication basée sur la contrainte induite par le procédé, qui permet d'expliquer en partie ces phénomènes. Des mesures en température sont conduites, et montrent que les collisions coulombiennes sur les flancs verticaux dans les canaux étroits ne dégradent pas significativement la mobilité. Un taux de balisticité de 75 % a également été mesuré à une longueur de grille de 50 nm.

Sommaire du chapitre 5

1	Introduction du chapitre	149
2	Définition de la mobilité	149
	2.1 Concept de masse effective.....	149
	2.2 Relation entre vitesse des porteurs et champ électrique	150
	2.3 Mécanismes de dégradation de la mobilité.....	151
3	Méthodes d'extraction de la mobilité	154
	3.1 Coefficients d'atténuation de la mobilité	154
	3.2 Méthode du maximum de transconductance $g_{m,max}$	154
	3.2.1 Transistors planaires	154
	3.2.2 Transistors Triple-grille	155
	3.3 Méthode de la fonction Y	155
	3.4 Méthode de la dérivée seconde de l'inverse du courant de drain	156
	3.5 Méthode Split C-V	157
	3.5.1 Méthode Split CV classique	157
	3.5.2 Cas des canaux courts.....	158
4	Mesures de mobilité sur les transistors Ω FET'	160
	4.1 Effet de la longueur de grille	160
	4.2 Effet de la largeur du canal	160
	4.2.1 Généralités sur les plans cristallins.....	160
	4.2.2 Effets du changement de direction cristalline ou de plan cristallin.....	162
	4.2.3 Mesures sur les transistors Ω FET's – effet de la largeur	165
	4.2.4 Effet du changement d'orientation cristalline.....	169
	4.2.5 Etude de l'effet de la contrainte	170
5	Mesures en température	173
	5.1 Variation des caractéristiques avec la température.....	173
	5.2 Mesures sur des transistors longs et larges	174
	5.3 Mesures sur des transistors longs et étroits.....	175
6	Mesure du taux de balisticité à basse température	176
7	Conclusions	178

1 Introduction du chapitre

Nous avons vu dans le chapitre précédent les avantages électrostatiques d'une structure multigrille de type Ω FET. L'objectif de ce chapitre est de présenter les mesures de transport réalisées dans le cadre de cette thèse. L'optimisation de la mobilité des dispositifs est un autre moyen d'améliorer le courant de sortie sans dégrader les courants de fuite. En particulier, l'orientation cristalline des dispositifs Ω FET dépend de leur orientation sur la plaque. Après avoir présenté les effets théoriques attendus, les mesures expérimentales seront exposées. En particulier, l'effet du changement d'orientation et de plan cristallin sera mis en lumière par des mesures en fonction de l'orientation sur la plaque, de la largeur de l'aileron et de la température. Enfin, une mesure expérimentale du taux de balisticité sera faite à la fin du chapitre.

2 Définition de la mobilité

2.1 Concept de masse effective

Un électron dans la bande de conduction est décrit en physique du solide par une fonction d'onde obtenue après résolution du théorème de Bloch. En termes corpusculaires, il s'agit d'une particule chargée dans un potentiel cristallin; l'électron subit donc l'influence du potentiel du réseau et n'est pas libre. On suppose alors que l'électron est une *quasi-particule* de charge $q = -e$ et de masse m_e^* appelée *masse effective* de l'électron. On prend ainsi en compte l'influence globale du potentiel cristallin sur l'électron dans la masse effective.

En combinant l'énergie quantique d'un électron avec le principe fondamental de la dynamique classique, on peut montrer [Mathieu'96] que la masse effective est reliée à la courbure des bandes d'énergie dans l'espace réciproque par la relation:

$$m_e^* = \frac{\hbar^2}{\partial^2 E / \partial k^2} \quad \text{Eq. (60)}$$

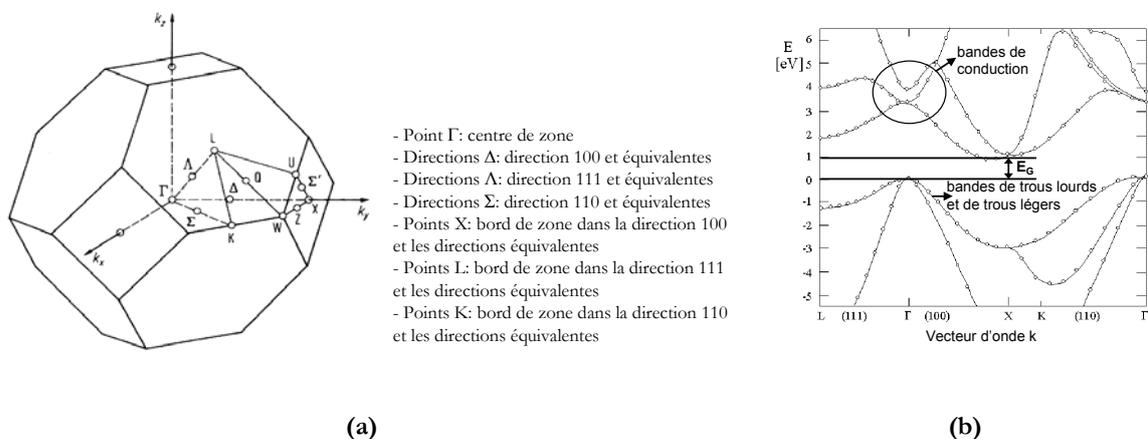


Fig. 114: Première zone de Brillouin d'un cristal cubique (a) et structure des bandes d'énergie dans le cas du silicium (b).

Dans le cas du silicium, les minima de la bande de conduction sont situés au point X (vallées Δ) et sont au nombre de 6 (Fig. 114.a et Fig. 114.b). A proximité d'un extremum, on peut utiliser une approximation parabolique pour une bande d'énergie (développement en série de Taylor à l'ordre 2). L'énergie est alors proportionnelle au carré du vecteur d'onde. Les surfaces iso-énergétiques sont donc 6 ellipsoïdes centrées en X (Fig. 115.a). En ce qui concerne la bande de valence, on peut appliquer les mêmes considérations aux trous. Cependant, la bande de valence est dégénérée en bande des trous lourds et des trous légers avec un maximum en Γ (Fig. 114.b). En raison de la dégénérescence se produisant en $k = 0$, ces bandes ne sont pas isotropes (Fig. 115.b).

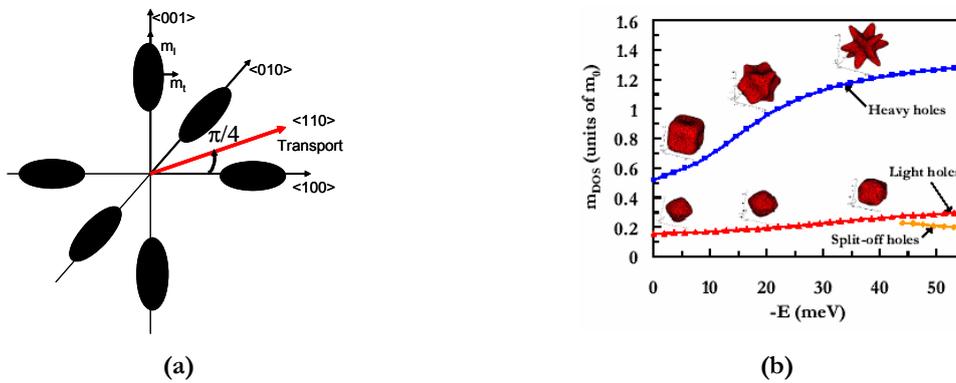


Fig. 115: (a): Surfaces d'énergie constante des bandes de conduction du silicium (d'après [Andrieu'05]). (b): Variation de la masse effective des trous lourds et légers en fonction de l'énergie (la référence est l'énergie en bord de bande pour les trous lourds et légers) (d'après [Guillaume'05]).

2.2 Relation entre vitesse des porteurs et champ électrique

En présence d'un champ électrique \vec{E} appliqué à l'instant $t = t_0$, un porteur d'une charge q est soumis à une force $\vec{F} = q\vec{E}$. En appliquant le principe fondamental de la dynamique à ce porteur de masse effective m^* [Mathieu'96], il vient:

$$m^* \frac{d\vec{v}_i(t)}{dt} = q\vec{E} \quad \text{Eq. (61)}$$

En intégrant cette équation de t_0 à $t_0 + t$, on obtient alors:

$$\vec{v}_i(t) = \frac{q\vec{E}}{m^*} t \quad \text{Eq. (62)}$$

La vitesse du porteur augmente donc linéairement entre deux collisions. En appelant τ_c la valeur moyenne du temps entre chaque collision, la valeur moyenne de la vitesse s'écrit alors:

$$v = \frac{1}{\tau_c} \int_0^{\tau_c} \vec{v}_i(t) dt = \frac{1}{\tau_c} \frac{q\vec{E}}{m^*} \frac{\tau_c^2}{2} \quad \text{Eq. (63)}$$

En posant alors $\tau = \tau_c/2$, Eq. (63) devient:

$$\bar{v} = \frac{q\bar{E}}{m^*} \tau \quad \text{Eq. (64)}$$

La mobilité μ d'un porteur de charge est définie comme la magnitude entre la vitesse de dérive d'un porteur et le champ électrique appliqué [Kittel'96]:

$$\bar{v} = \mu \bar{E} \quad \text{Eq. (65)}$$

En combinant l'Eq. (64) et l'Eq. (65), on obtient alors la formule générale de la mobilité:

$$\mu = \frac{q\tau}{m^*} \quad \text{Eq. (66)}$$

La mobilité est une grandeur définie positive quelque soit le type de porteur. A cause de la dégénérescence de la bande de valence au point Γ , des transitions interbandes pour les trous sont possibles et limitent considérablement la mobilité. En conséquence, la mobilité des trous est typiquement plus faible que la mobilité des électrons. Le Tab. 4 présente les valeurs des mobilités à température ambiante pour quelques matériaux semiconducteurs massifs.

Cristal	$\mu_{\text{électrons}}$ [cm ² .V ⁻¹ .s ⁻¹]	μ_{trous} [cm ² .V ⁻¹ .s ⁻¹]	Cristal	$\mu_{\text{électrons}}$ [cm ² .V ⁻¹ .s ⁻¹]	μ_{trous} [cm ² .V ⁻¹ .s ⁻¹]
Diamant	1800	1200	InAs	30000	450
Si	1350	480	GaAs	8000	300
Ge	3600	1800	SiC	100	10 à 20

Tab. 4: Mobilités des porteurs à température ambiante pour quelques matériaux semiconducteurs massifs [Kittel'96].

2.3 Mécanismes de dégradation de la mobilité

E_{EFF} est le champ transverse électrique effectif (perpendiculaire à la direction du transport) associé à la polarisation de la grille V_G . Il dépend du dopage du substrat (intervenant dans le terme de charge de désertion Q_{DEP}) et de la charge d'inversion $Q_{\text{INV}} = C_{\text{OX}} (V_G - V_T)$ selon la relation :

$$E_{\text{EFF}} = \frac{\eta Q_{\text{INV}} + Q_{\text{DEP}}}{\epsilon_{\text{Si}}} \quad \text{Eq. (67)}$$

où η est un facteur tenant compte du fait que l'on considère un champ moyen à l'intérieur de la couche d'inversion. Cette constante est égale à 1/2 pour les électrons et à 1/3 pour les trous dans un plan (100). Elle vaut également 1/3 pour les électrons dans un plan (110) ou (111) [Takagi'94-2]. Dans le canal d'un transistor MOS, un porteur est soumis principalement à trois types d'interactions [Takagi'94-1]:

- ✓ Interactions coulombiennes

Ce sont les interactions entre le porteur et les particules électriquement chargées. Elles sont induites par toute charge mobile ou fixe comme les dopants ou les charges dans l'oxyde de grille. A fort champ, elles sont écartées par la forte densité de porteurs. Le terme coulombien d'atténuation de la mobilité est de la forme:

$$\mu_{\text{COULOMB}} \propto E_{\text{EFF}} T \quad \text{Eq. (68)}$$

où E_{EFF} est le champ transverse effectif (perpendiculaire à la direction du transport) et T la température.

✓ Interactions avec les phonons

Les phonons sont des quasi-particules associées aux vibrations du réseau. A basse température, l'agitation thermique est faible et la dégradation par les phonons est moins importante. Le terme phonique d'atténuation de la mobilité est de la forme:

$$\mu_{PHONON} \propto E_{EFF}^{-1/3} T^{\beta} \quad \text{Eq. (69)}$$

où β dépend principalement de l'orientation cristalline [Takagi'94-2]. Ce terme est compris entre -1 et -2.

✓ Interactions avec l'interface silicium/oxyde de grille

Ce terme représente les collisions dues à la rugosité à l'interface silicium/oxyde de grille. A fort champ, les porteurs se rapprochent de l'interface silicium/oxyde de grille et ce terme devient prépondérant. Il ne dépend normalement pas de la température et on peut l'exprimer sous la forme:

$$\mu_{RUGOSITE} \propto E_{EFF}^{-2} \quad \text{Eq. (70)}$$

Ces collisions vont diminuer la mobilité des porteurs de manière significative. Ces différents mécanismes n'interagissant pas entre eux et la dégradation de la mobilité étant gouvernée par le temps le plus court entre chaque collision, la mobilité totale μ_{EFF} s'exprime usuellement selon la relation (ou loi de Mathiessen):

$$\frac{1}{\mu_{EFF}} = \frac{1}{\mu_{COULOMB}} + \frac{1}{\mu_{PHONON}} + \frac{1}{\mu_{RUGOSITE}} \quad \text{Eq. (71)}$$

La contribution relative de ces trois mécanismes dépend du champ transverse effectif et de la température. A une température donnée, la mobilité est dégradée par les interactions coulombiennes à faible champ et par les interactions de rugosité à fort champ. La Fig. 116.a montre l'allure de la mobilité en fonction du champ effectif vu par un porteur, résultante de toutes ces interactions.

A faible champ, la mobilité dépend peu de la rugosité de l'interface silicium/oxyde de grille. A basse température ($T < 100$ K), la mobilité effective diminue si les interactions coulombiennes sont importantes. Si elles sont négligeables devant les interactions phoniques, on observe un plateau de mobilité (Fig. 116.b).

Dans une plage de température intermédiaire, les interactions prédominantes sont les interactions avec les phonons. La mobilité en fonction de la température suit alors une loi de puissance. La position du 'coude' de mobilité dépend de l'importance des interactions phoniques (Fig. 116.b). De même, le terme de dépendance β des interactions phoniques décale le 'coude' (Fig. 116.c).

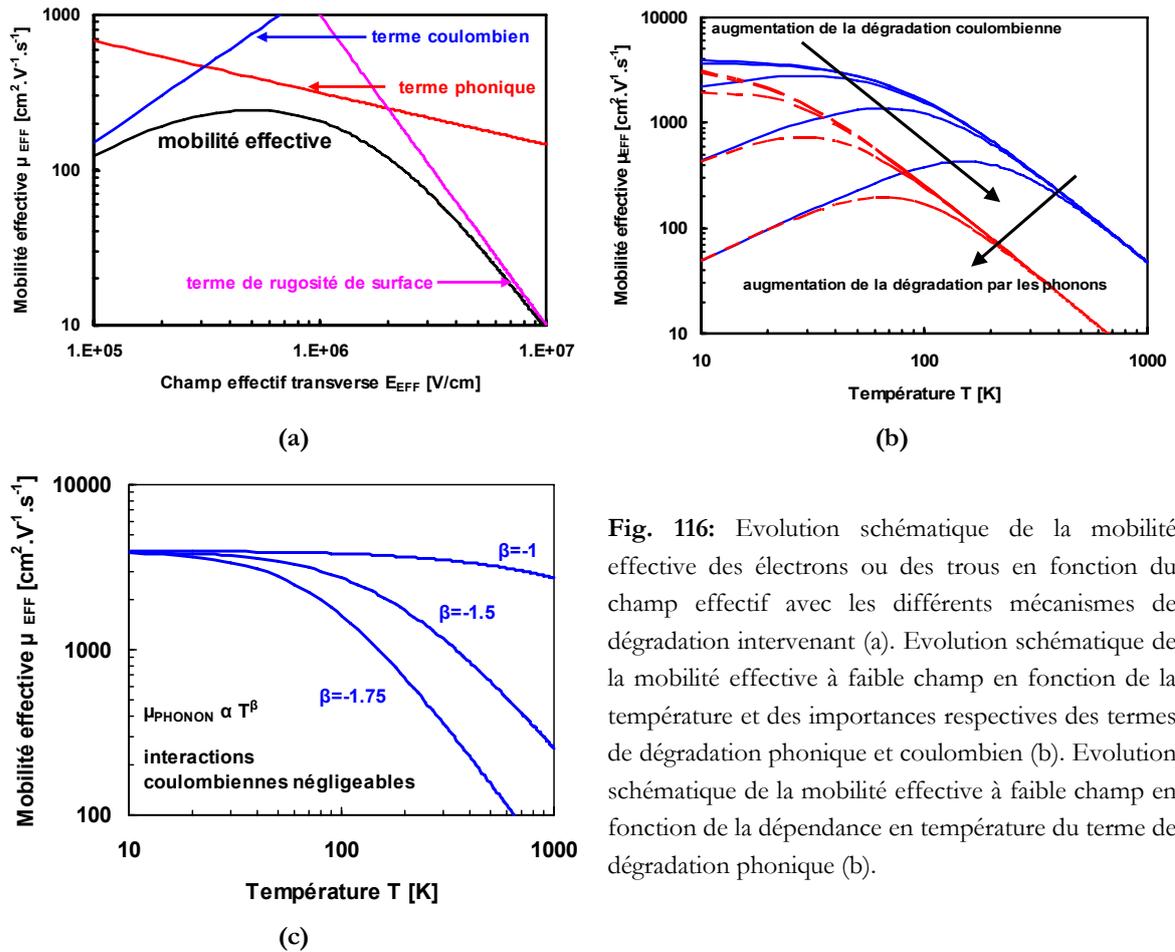


Fig. 116: Evolution schématique de la mobilité effective des électrons ou des trous en fonction du champ effectif avec les différents mécanismes de dégradation intervenant (a). Evolution schématique de la mobilité effective à faible champ en fonction de la température et des importances respectives des termes de dégradation phonique et coulombien (b). Evolution schématique de la mobilité effective à faible champ en fonction de la dépendance en température du terme de dégradation phonique (b).

A cause de tous ces phénomènes de dégradation, la mobilité mesurée dans le canal d'un transistor est bien plus faible que celle mesurée dans un matériau massif. Le maximum de mobilité effective est ainsi d'environ 600-800 cm².V⁻¹.s⁻¹ pour les électrons et de 200 cm².V⁻¹.s⁻¹ pour les trous dans le cas d'un empilement de grille SiO₂/Polysilicium. Avec un oxyde *high-k* comme HfO₂, la mobilité mesurée est encore plus faible car des collisions supplémentaires interviennent (Fig. 117).

Le plan cristallin dans lequel s'effectue la conduction est également important. La mobilité des électrons est maximale dans un plan (100) (Fig. 117.a). Des mobilités pour les trous 2.5 fois plus élevées dans un plan (110) que dans un plan (100) ont été mesurées (Fig. 117.b). C'est un effet que l'on peut utiliser pour augmenter les performances d'un transistor vertical (voir section 4.2.2.2).

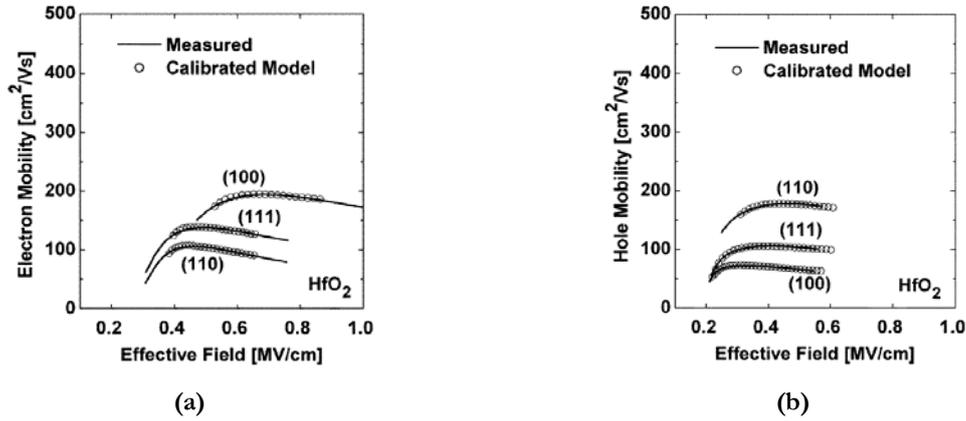


Fig. 117: Mobilité sur HfO₂ selon différents plans cristallins pour les électrons (a) et les trous (b) (d'après [Chang'04]).

3 Méthodes d'extraction de la mobilité

3.1 Coefficients d'atténuation de la mobilité

A température donnée, la dégradation de la mobilité avec le champ effectif (reliée à la tension de grille par Q_{INV}) peut être exprimée plus simplement à l'aide de la relation empirique [Reichert'98]:

$$\mu_{EFF} = \frac{\mu_0}{1 + \theta_1(V_G - V_T) + \theta_2(V_G - V_T)^2} \quad \text{Eq. (72)}$$

Cette relation n'est valable qu'en inversion forte. Le terme θ_1 traduit la dégradation de la mobilité par les résistances d'accès R_{SD} . Le terme quadratique θ_2 traduit quand à lui l'effet de rugosité de surface. Il est important surtout quand l'oxyde de grille est très fin. La prise en compte des coefficients d'atténuation conduit à des méthodes d'extraction de mobilité décrites dans les sections suivantes.

3.2 Méthode du maximum de transconductance $g_{m,max}$

3.2.1 Transistors planaires

En prenant la relation simple donnant le courant passant dans un MOS en régime linéaire, il vient :

$$I_D = \mu_{EFF} C_{OX} \frac{W}{L} \left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right] \quad \text{Eq. (73)}$$

où W est la largeur du canal et L sa longueur.

En utilisant la relation $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$ et en dérivant l'Eq. (73) pour obtenir la transconductance maximale $g_{m,max}$, il vient pour le canal la relation :

$$g_{m,max} = \mu \frac{\epsilon_{ox}}{t_{ox}} \frac{W}{L} V_{DS} \quad \text{Eq. (74)}$$

On dispose donc d'une mesure directe de la mobilité. Cependant, cette mesure suppose que la mobilité ne dépend pas de la longueur de grille. On ne s'affranchit donc pas de l'effet des résistances séries et de la rugosité des flancs.

3.2.2 Transistors Triple-grille

En considérant un transistor Triple-grille comme 3 transistors planaires (le canal du dessus plus deux canaux latéraux) indépendants, c'est-à-dire en ne prenant pas en compte les effets de couplage entre les grilles, il a été développé une technique d'extraction des mobilités planaires et latérales du transistor [Pretet'03].

En reprenant Eq. (73) pour obtenir la transconductance maximale $g_{m,\max}$, il vient pour le canal avant la relation :

$$g_{m,\max}^{\text{avant}} = \mu_{\text{avant}} \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}^{\text{avant}}} \frac{W}{L} V_{DS} \quad \text{Eq. (75)}$$

et pour un canal latéral :

$$g_{m,\max}^{\text{latéral}} = \mu_{\text{latéral}} \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}^{\text{latéral}}} \frac{H}{L} V_{DS} \quad \text{Eq. (76)}$$

La transconductance totale $g_{m,\max}$ (celle qui peut être calculée à partir de la courbe $I_{DS}(V_G)$ mesurée) s'exprime donc comme:

$$g_{m,\max}^{\text{total}} = g_{m,\max}^{\text{avant}} + 2g_{m,\max}^{\text{latéral}} = \mu_{\text{avant}} \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}^{\text{avant}}} \frac{W}{L} V_{DS} + 2\mu_{\text{latéral}} \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}^{\text{latéral}}} \frac{H}{L} V_{DS} \quad \text{Eq. (77)}$$

En posant $A = \frac{1}{2} \frac{L}{H} \frac{t_{\text{ox}}^{\text{latéral}}}{\varepsilon_{\text{ox}}} \frac{1}{V_D}$, on peut exprimer l'Eq. (77) sous la forme:

$$A g_{m,\max}^{\text{total}} = \mu_{\text{latéral}} + \frac{t_{\text{ox}}^{\text{latéral}}}{t_{\text{ox}}^{\text{avant}}} \frac{\mu_{\text{avant}}}{2H} W \quad \text{Eq. (78)}$$

La courbe $A g_{m,\max}^{\text{total}} = f(W)$ est donc une droite dont la pente permet de calculer μ_{avant} et l'ordonnée à l'origine $\mu_{\text{latéral}}$.

Cette méthode suppose cependant que le transistor Triple-Grille se comporte comme l'assemblage de trois transistors planaires n'ayant pas d'interactions entre eux. Pour les canaux larges (W_{FIN} grand), on se rapproche d'un comportement de type planaire et l'influence des canaux latéraux devient très faible. L'extraction de la mobilité avant sera juste, et celle de la mobilité latérale entachée d'une incertitude.

3.3 Méthode de la fonction Y

En ne tenant compte que du premier coefficient d'atténuation, le courant de drain en régime linéaire au dessus du seuil est de la forme:

$$I_{DS} = \mu_0 C_{\text{ox}} \frac{W}{L} V_{DS} \frac{V_G - V_T}{1 + \theta_1 (V_G - V_T)} \quad \text{Eq. (79)}$$

Afin de pouvoir extraire la mobilité, la méthode la plus simple consiste à chercher une fonction linéaire par rapport à $(V_G - V_T)$. On élimine le terme θ_1 en dérivant l'inverse du courant de drain:

$$\frac{\partial(1/I_D)}{\partial V_G} = \frac{-L}{\mu_0 C_{ox} W V_D} \frac{1}{(V_G - V_T)^2} \quad \text{Eq. (80)}$$

Ce résultat se linéarise afin de pouvoir extraire la mobilité μ_0 en utilisant la pente de cette droite:

$$F(V_G) = \sqrt{\frac{-1}{\partial(1/I_D)/\partial V_G}} = \sqrt{\frac{\mu_0 C_{ox} W V_D}{L}} (V_G - V_T) \quad \text{Eq. (81)}$$

La fonction est mathématiquement équivalente à la grandeur $Y(V_G) = \frac{I_D}{\sqrt{g_m}}$ [Ghibaudo'88].

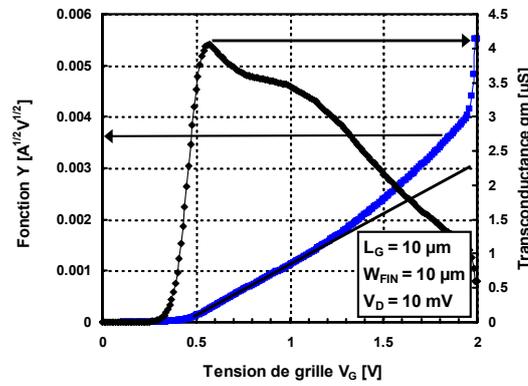


Fig. 118: Transconductance et fonction Y obtenues pour un transistor large complètement déserté ($W_{FIN} = 10 \mu\text{m}$).

La Fig. 118 montre l'allure typique d'une fonction Y. Cependant, la zone linéaire n'est pas toujours parfaite, en particulier pour les forts champs. Cet effet de champ fort est supposé être dû au confinement des porteurs à proximité de l'interface diélectrique/canal; la rugosité de cette interface conduit à une non-linéarité due au coefficient θ_2 .

3.4 Méthode de la dérivée seconde de l'inverse du courant de drain

Cette méthode est l'extension à l'ordre deux de la méthode précédente. En effet, si on tient compte de l'influence d'un deuxième terme d'atténuation θ_2 , la fonction Y devient:

$$Y = \sqrt{\frac{\mu_0 C_{ox} W V_D}{L}} \frac{V_G - V_T}{\sqrt{\theta_2 (V_G - V_T)^2 - 1}} \quad \text{Eq. (82)}$$

Pour de fortes tensions de grille, la fonction Y devient fortement non linéaire. Comme précédemment, on peut supprimer les coefficients d'atténuation en dérivant deux fois l'inverse du courant de drain [Faynot'94][McLarty'95]:

$$\frac{\partial^2(1/I_D)}{\partial V_G^2} = \frac{2L}{\mu_0 C_{OX} W V_D} \frac{1}{(V_G - V_T)^3} \quad \text{Eq. (83)}$$

En linéarisant, il vient:

$$H(V_G) = \sqrt[3]{\frac{\mu_0 C_{OX} W V_D}{2L}} (V_G - V_T) \quad \text{Eq. (84)}$$

On élimine ainsi l'incertitude sur la zone linéaire à choisir que l'on peut rencontrer en utilisant la fonction Y. Cependant, comme il s'agit d'une double dérivée et que la mobilité est proportionnelle au cube de la pente extraite, la barre d'erreur sur le résultat est plus importante.

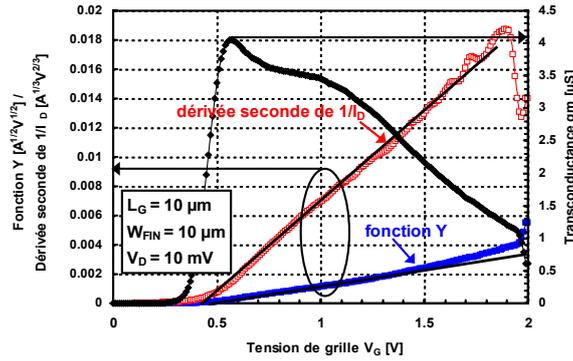


Fig. 119: Transconductance, fonction Y et deuxième dérivée de l'inverse du courant de drain obtenues pour un transistor large complètement déserté ($W_{FIN} = 10 \mu\text{m}$).

3.5 Méthode Split C-V

La méthode la plus couramment employée pour mesurer la mobilité effective est la méthode dite 'Split CV' [Koomen'73]. Elle se base sur une mesure capacitive et une mesure courant-tension afin d'obtenir la variation de la mobilité effective avec le champ électrique effectif.

Il est nécessaire d'utiliser un transistor assez large, afin d'avoir des capacités de l'ordre du pF. C'est pourquoi on utilise généralement des transistors très longs et les plus larges possibles. Cependant, dans le cas des Triple-grille où on souhaite mesurer les mobilités des flancs latéraux, on ne peut plus utiliser de transistors larges. La solution consiste alors à utiliser des motifs en réseaux. Les motifs en réseaux Triple-grille sont ainsi constitués de 50 ailerons en parallèle.

3.5.1 Méthode Split CV classique

On mesure tout d'abord la capacité grille-canal (source et drain connectés, substrat à la masse). On mesure ainsi la charge d'inversion et de désertion avec la relation :

$$Q_{INV}(V_G) = \int_{V_{GACC}}^{V_G} C_{GRILLE-CANAL}(u) du \quad \text{Eq. (85)}$$

où V_{GACC} est une tension de grille prise en accumulation.

On peut remonter à l'expression du champ effectif avec la relation :

$$E_{EFF} = \frac{\eta Q_{INV} + Q_{DEP}}{\epsilon_{Si}} \quad \text{Eq. (86)}$$

où η vaut $\frac{1}{2}$ pour les électrons et $\frac{1}{3}$ pour les trous dans un plan (100). Q_{DEP} peut être supposé négligeable devant Q_{INV} pour des dispositifs complètement désertés.

Maintenant, en considérant un transistor opérant en régime ohmique, on peut relier la mobilité effective μ_{EFF} avec le courant de drain mesuré par la relation :

$$\mu_{EFF} = \frac{L}{W} \frac{I_D}{(C_{OX}(V_G - V_T))V_D} = \frac{L}{W} \frac{I_D}{Q_{INV}V_D} \quad \text{Eq. (87)}$$

En forte inversion, le champ effectif est complètement contrôlé par la charge en forte inversion. Si c'est juste le transport dans le canal qui nous intéresse, on peut donc juste tracer la variation de la mobilité effective en fonction de la charge d'inversion (Eq. (87)). Un des intérêts de cette méthode est qu'elle normalise la largeur du dispositif.

3.5.2 Cas des canaux courts

Cette partie reprend les résultats exposés dans la thèse de K. Romanjek [Romanjek'04-2]. Dans le cas d'un transistor court planaire, les différentes capacités parasites prennent beaucoup d'importance dans le signal et peuvent considérablement perturber la mesure de la mobilité. De précédentes études ont permis de mieux comprendre ces phénomènes [Prégaldiny'02][Romanjek'04]. La mesure de la capacité grille-canal va être surévaluée par une capacité parasite dépendante de la tension de grille, mais indépendante de la longueur de grille. L'origine physique de cette capacité parasite peut se comprendre en analysant les différentes capacités de recouvrement (*overlap*) dans le transistor (Fig. 120.a).

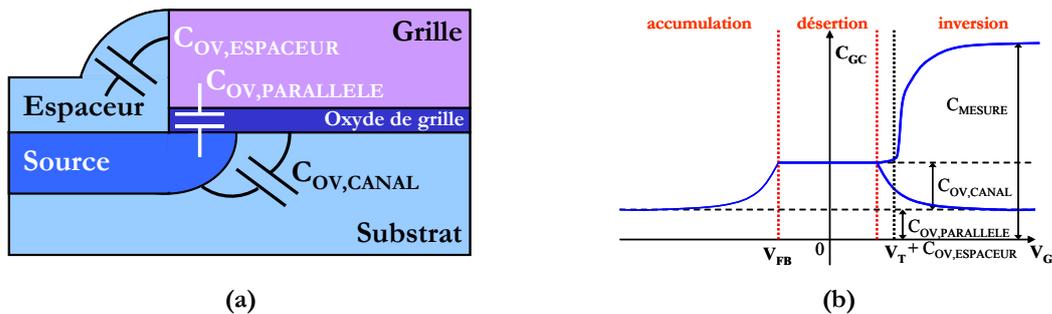


Fig. 120: Capacités de recouvrement dans la structure côté source (a) et allure de la courbe capacité grille-canal résultante (b).

Une capacité étant une zone isolante entre deux zones à comportement métallique, on peut en distinguer trois composantes dans le transistor. $C_{OV,ESPACEUR}$ est la capacité reliant la source et la grille via l'espaceur, $C_{OV,PARALLELE}$ est la capacité reliant la source et la grille via l'oxyde de grille et $C_{OV,CANAL}$ est la capacité reliant la source et la grille via le substrat. Les deux capacités $C_{OV,ESPACEUR}$ et $C_{OV,PARALLELE}$ n'entraînent qu'un décalage constant. $C_{OV,CANAL}$ existe uniquement s'il n'y a pas de charges mobiles dans le substrat. Par contre, la capacité $C_{OV,CANAL}$ est maximale en désertion. En accumulation, il se crée des charges mobiles (porteurs

majoritaires du substrat) qui vont écranter peu à peu cette capacité $C_{OV,CANAL}$. De façon symétrique, en inversion cette capacité sera écrantée par les charges mobiles du canal d'inversion (porteurs minoritaires du substrat). En forte inversion, si le plateau dû aux capacités $C_{OV,ESPACEUR}$ et $C_{OV,PARALLELE}$ est corrigé, on retrouve également la capacité d'oxyde effective (Fig. 120.b).

La méthode pour s'affranchir des capacités parasites consiste à mesurer un jeu de capacités pour différentes longueurs de grille. Après correction de $C_{OV,ESPACEUR}$ et $C_{OV,PARALLELE}$, la capacité parasite restante est $C_{OV,CANAL}$ (Fig. 121.a). Cette capacité parasite étant fonction de la tension de grille et non de la longueur de grille, on retranche alors les capacités mesurées une à une afin d'obtenir la capacité corrigée (Fig. 121.b). A cause de l'intégration de la capacité parasite $C_{OV,CANAL}$, la charge d'inversion calculée à partir de l'intégration de la capacité grille-canal est sur-évaluée si on n'applique pas cette correction (Fig. 122.a). La mobilité obtenue avec la méthode Split CV sera alors considérablement sous-évaluée (Fig. 122.b).

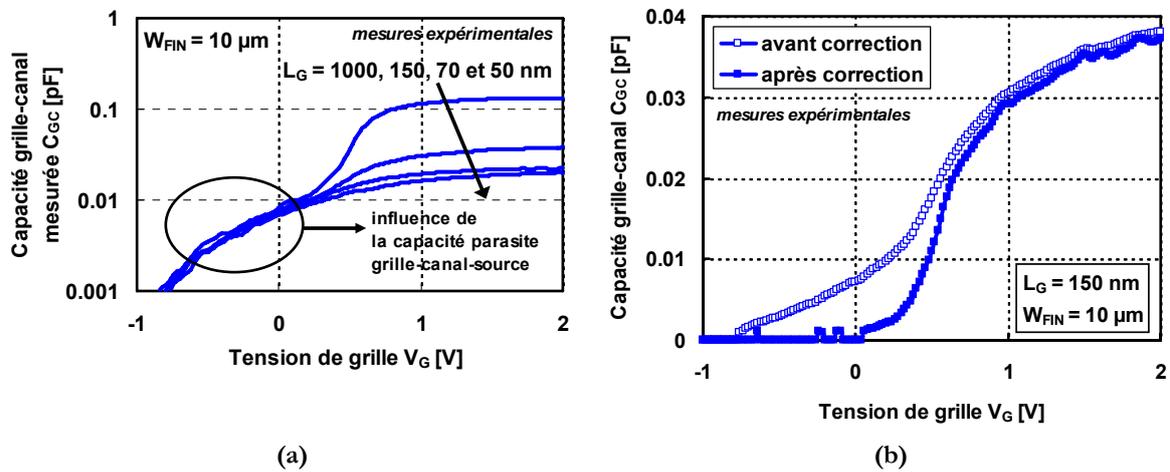


Fig. 121: Capacité grille-canal en fonction de la longueur de grille (échelle logarithmique) (a) et effet de la correction de $C_{OV,CANAL}$ sur la capacité totale (b).

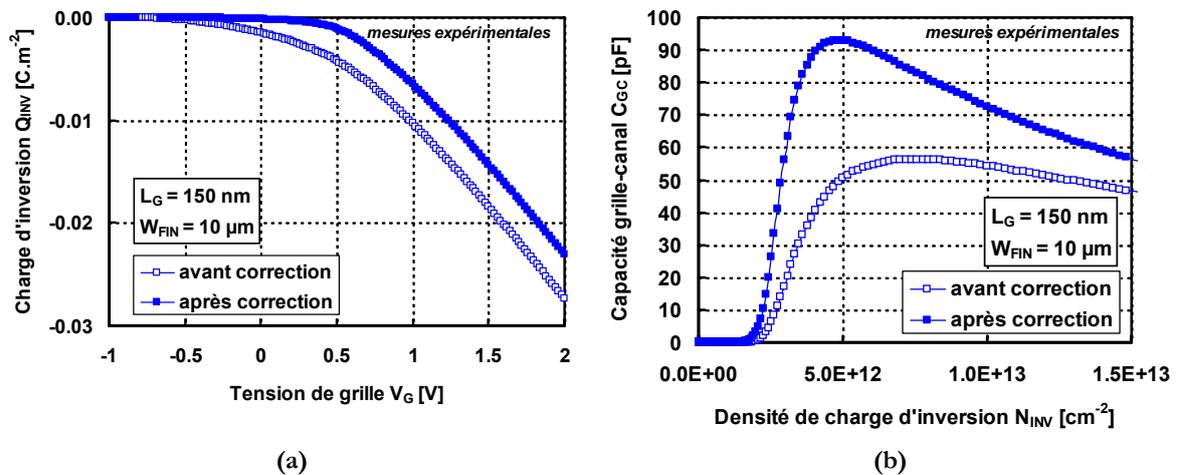


Fig. 122: Effet de la correction de la capacité $C_{OV,CANAL}$ sur la charge d'inversion (a) et sur la mobilité extraite par la méthode Split CV (b).

4 Mesures de mobilité sur les transistors Ω FET

4.1 Effet de la longueur de grille

Nous avons vu que les capacités parasites dans les structures MOS devenaient importantes pour les courtes longueurs de grille, conduisant à l'extraction d'une charge d'inversion en fonction de la tension de grille considérablement sur-évaluée. L'application d'une correction (voir section 3.5.2) permet de s'affranchir des effets de canaux courts. Cette méthode a été employée sur des transistors larges ($W_{\text{FIN}} = 10 \mu\text{m}$, afin d'avoir une surface suffisante à la mesure d'une capacité) dont la longueur de grille L_G varie de $1 \mu\text{m}$ à 50 nm (Fig. 123).

On observe une chute spectaculaire (de $200 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$ à $50 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$, soit un facteur 4) de la mobilité entre $L_G = 1 \mu\text{m}$ et $L_G = 50 \text{ nm}$. Cette dégradation a déjà été observée sur des dispositifs utilisant le même empilement de grille [Andrieu'05][Widiez'05][Widiez'05-2]. Dans ces références, des mesures ont montré une forte chute de mobilité pour les transistors courts à basse température (en dessous de 100 K). Ceci indique que la dégradation de mobilité est majoritairement due aux interactions coulombiennes et/ou aux interactions avec des défauts neutres. Ces charges ou défauts additionnels pourraient être dus à l'implantation source/drain, aux charges aux jonctions induites par les gravures, voire aux matériaux utilisés (oxyde de grille ou grille). L'explication physique de ces phénomènes est actuellement en débat.

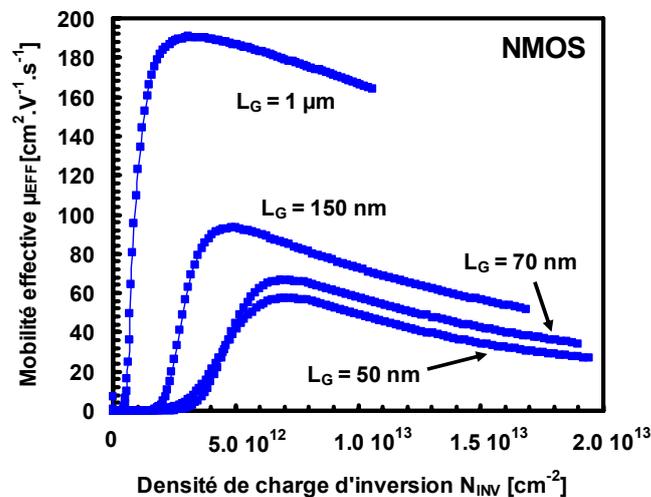


Fig. 123: Mobilité en fonction de la densité de charge d'inversion pour différentes longueurs de grille (de $1 \mu\text{m}$ à 50 nm). La mobilité est extraite avec la méthode Split CV (effet de canal court corrigé). La largeur de l'aileron W_{FIN} est de $10 \mu\text{m}$ et sa hauteur H_{FIN} de 26 nm .

4.2 Effet de la largeur du canal

4.2.1 Généralités sur les plans cristallins

Les cristaux covalents sont construits avec des éléments de la quatrième colonne du tableau périodique de Mendeleïev. Ces éléments ont quatre électrons périphériques qu'ils mettent en commun avec quatre voisins pour établir des liaisons covalentes [Mathieu'96]. Le

silicium, le germanium et le diamant sont des cristaux covalents. Grossièrement, la maille élémentaire de ces cristaux est constituée de deux structures 'cubique face centrée' décalées d'un quart de maille dans toutes les directions [Rousseau'95] (Fig. 124.a). Un plan passant par trois nœuds et donc contenant une infinité de nœuds est un plan réticulaire. L'ensemble des plans réticulaires parallèles constitue une famille de plans contenant l'ensemble des nœuds du réseau. On définit ces plans par des indices, dits de Miller [Rousseau'95]. On distingue ainsi dans le cristal de silicium les plans cristallins (100), (110) et (111) (correspondant respectivement aux Fig. 124.b, c et d).

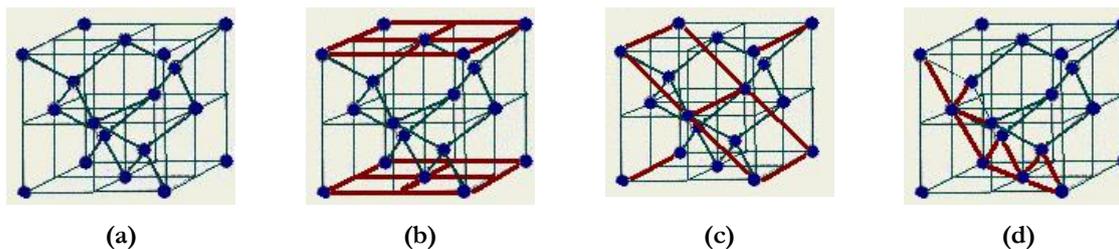


Fig. 124: Maille élémentaire du cristal de silicium (a). Visualisation des plans cristallins (100) (b), (110) (c) et (111) (d). (d'après [Byu])

Suivant le plan de découpe, une plaquette de silicium peut avoir une de ces trois orientations. Dans chaque plan cristallin, le transport des porteurs peut se faire selon certains vecteurs que l'on nomme directions cristallographiques (Fig. 125).

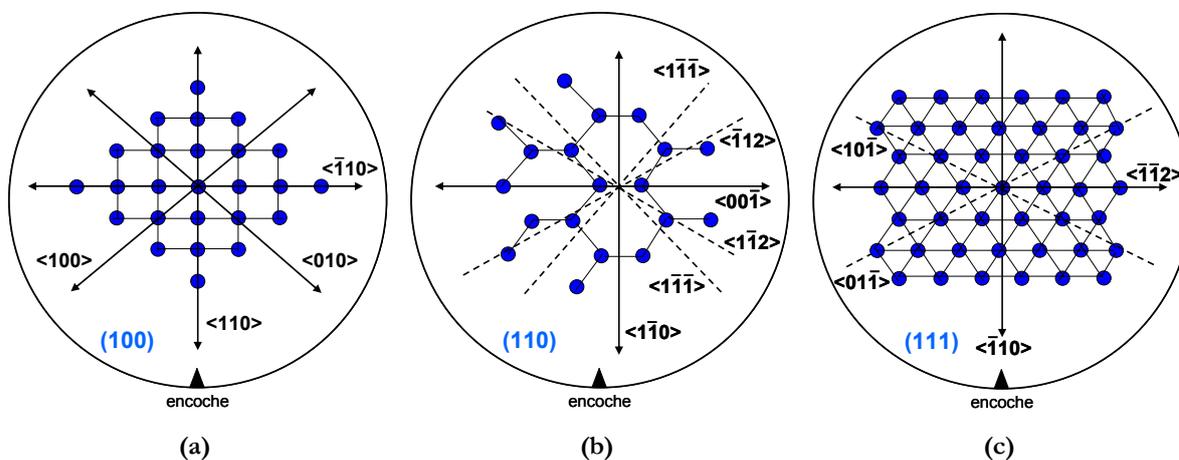


Fig. 125: Directions cristallines possibles pour un plan (100) (a), (110) (b) ou (111) (c).

Les plaques utilisées pour la fabrication de nos transistors sont des plaques dont le plan cristallin est (100). L'encoche montre la direction cristalline $\langle 110 \rangle$ (Fig. 125.a). Cette disposition est très classique. Dans les dispositifs fabriqués, nous disposons de dispositifs orientés à 0° ou à 45° par rapport à l'encoche, c'est-à-dire ayant respectivement une direction cristalline $\langle 110 \rangle$ et $\langle 100 \rangle$.

4.2.2 Effets du changement de direction cristalline ou de plan cristallin

4.2.2.1 Effet du changement de direction cristalline

L'effet du changement de direction cristalline est connu depuis de nombreuses années. Rappelons que les dispositifs sont généralement orientés dans l'axe de l'encoche, c'est-à-dire dans la direction cristalline $\langle 110 \rangle$. En orientant les dispositifs dans la direction $\langle 100 \rangle$, il a été montré que la mobilité des trous est améliorée d'environ 7% sans que la mobilité des électrons ne soit sensiblement dégradée [Sayama'99] (Fig. 126.a). Ainsi, les performances d'une technologie CMOS sont très simplement améliorées en tournant la plaque à 45° par rapport à l'encoche. Le mécanisme physique proposé pour expliquer ce phénomène était la plus petite masse effective des trous lourds dans la direction $\langle 100 \rangle$ et la quasi-isotropie des masses effectives pour les électrons (Fig. 126.b).

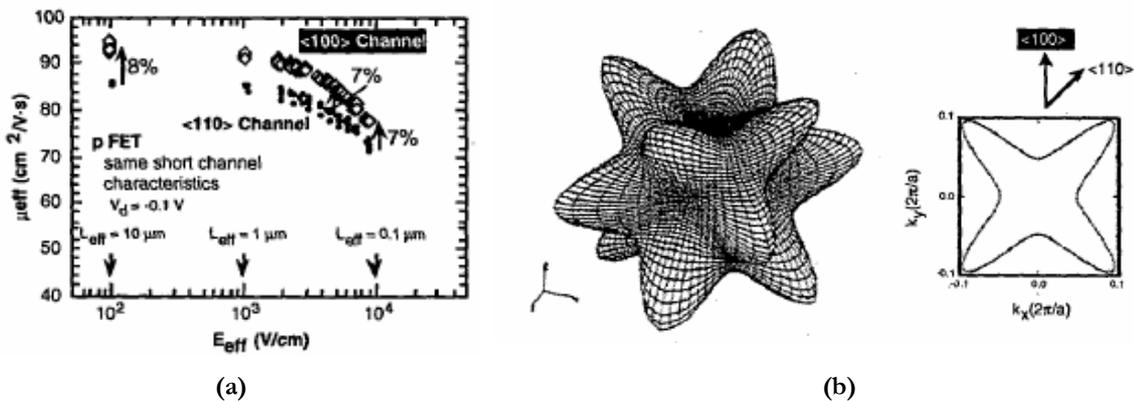


Fig. 126: Mobilité des trous en fonction du champ effectif (a). Surface iso-énergie et coupe en $k_z=0$ pour les trous lourds dans le silicium à 40 mV du bord de bande (b) (d'après [Sayama'99]).

Cependant, l'amélioration observée pour les transistors PMOS semble diminuer pour les grandes longueurs de grille (Fig. 127.a). Or la variation des masses effectives ne dépend pas à priori de la longueur de grille. Récemment, il a été proposé [Saito'06] que pour des transistors courts (donc également étroits) la contrainte biaxiale compressive induite par les isolations *STI* était responsable de l'amélioration de la mobilité dans les transistors PMOS (Fig. 127.b). Ce phénomène reste donc en débat. Nos transistors Ω FETs ne possèdent pas d'isolations *STI*, mais d'autres parties du transistor comme la grille pourraient également induire des contraintes sur le canal.

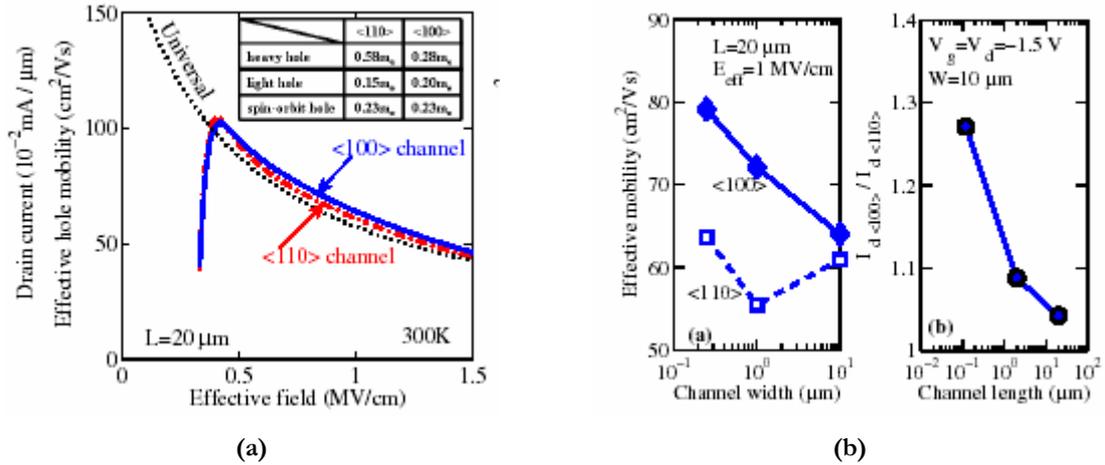


Fig. 127: Mobilité selon les directions cristallines $\langle 100 \rangle$ et $\langle 110 \rangle$ dans un transistor long ($L_G = 20 \mu\text{m}$) (a). Effet de canal court et de canal étroit sur le gain en mobilité/en courant (b) (d'après [Saito'06]).

4.2.2.2 Effet du changement de plan cristallin

L'examen de la première zone de Brillouin dans l'espace réciproque montre que pour des dispositifs Triple-grille, les flancs verticaux et horizontaux peuvent avoir des orientations tout à fait différentes. De même, l'orientation cristalline d'un flanc vertical dépend de la direction cristalline du dispositif (Fig. 128.a). Ainsi, si le dispositif est dans la direction $\langle 110 \rangle$ sur une plaque (100), les flancs verticaux sont dans un plan (110); s'il est dans la direction $\langle 100 \rangle$, les flancs verticaux sont dans un plan (100) (Fig. 128.b).

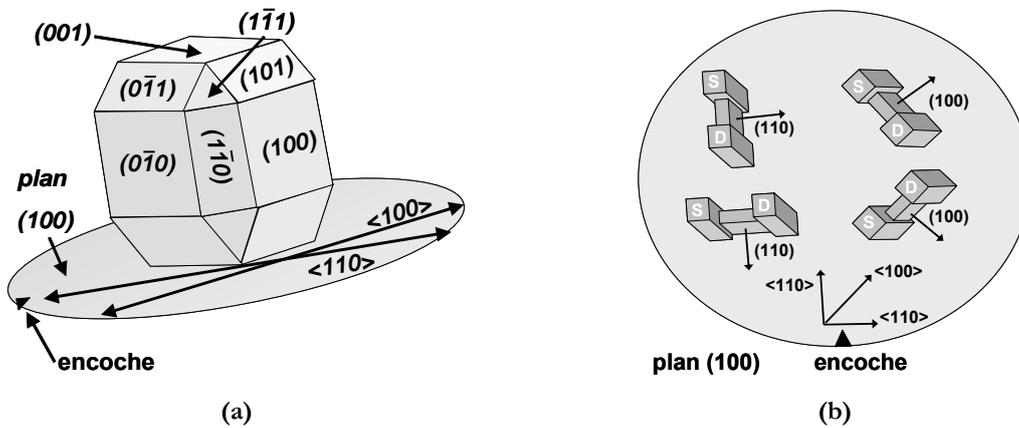


Fig. 128: Plans et directions cristallines dans le silicium (a) et plans cristallins des flancs d'un transistor FinFET selon leur orientation (b).

Cet effet d'orientation cristalline pour les flancs verticaux peut être utilisé pour améliorer les performances en disposant les NMOS dans un plan (100) et les PMOS dans un plan (110) simplement en les tournant par rapport à l'encoche. Un inverseur à transistors FinFETs peut être alors optimisé en utilisant un NMOS dans un plan (100) et un PMOS dans un plan (110). Cependant, NMOS et PMOS sont alors décalés de 45° , engendrant une plus grande surface

occupée sur la plaque (Fig. 129.a). Ce problème de surface est levé si on réussit à fabriquer différents plans cristallins sur une même plaque SOI [Yang'03][Yang'06].

De même, un transistor Triple-grille verra en théorie sa mobilité varier selon le rapport entre la hauteur et la largeur de l'aileron en raison de la combinaison des canaux horizontaux et verticaux. Par exemple, si le dispositif est orienté à 0° par rapport à l'encoche le plan (100) (respectivement (110)) va prédominer avec un rapport hauteur/largeur faible (respectivement élevé) (Fig. 129.b).

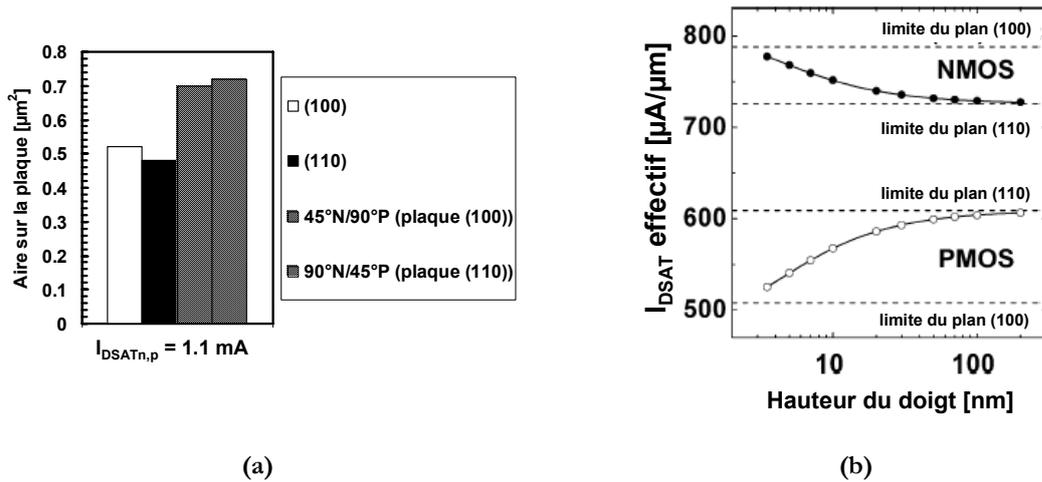


Fig. 129: (a): Aire occupée sur la plaque à courant de saturation donné (simulations) pour un inverseur à transistors FinFETs pour différentes orientations cristallines. L'espacement entre chaque aileron est de 100 nm, et la hauteur due l'aileron est de 100 nm. (b): Evolution du courant de saturation simulé en fonction de la hauteur de l'aileron pour un transistor Triple-grille orienté à 0° par rapport à l'encoche (d'après [Chang'04]).

4.2.2.3 Résumé

Nous avons remarqué des chanfreins sur la structure expérimentale (voir chapitre 3, Fig. 46.a). A priori, ils sont dans un plan (111) si la direction cristalline du dispositif est $\langle 110 \rangle$ et dans un plan (110) si la direction est $\langle 100 \rangle$. Par souci de simplicité, nous ne nous sommes pas occupés de cette conduction dans l'étude qui suit.

Dans la section précédente, nous avons vu que:

- ✓ En diminuant la largeur de l'aileron, on privilégie les flancs verticaux (mobilité μ_{LAT} , Fig. 130) dont l'orientation cristalline peut être différente de celle du canal supérieur (mobilité μ_{TOP}). La mobilité totale μ_{TOTAL} mesurée est alors de la forme:

$$\mu_{\text{TOTAL}} = \frac{W_{\text{FIN}}\mu_{\text{TOP}} + 2H_{\text{FIN}}\mu_{\text{LAT}}}{W_{\text{FIN}} + 2H_{\text{FIN}}} \quad \text{Eq. (88)}$$

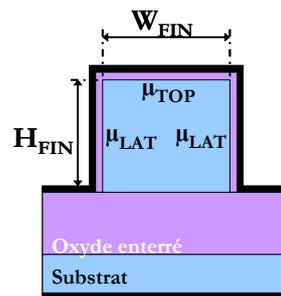


Fig. 130: Les différents canaux de conceptions intervenant dans un transistor Triple-grille.

- ✓ En changeant la position du transistor par rapport à l'encoche, on modifie l'orientation cristalline du transport; la mobilité peut ainsi varier.

Le Tab. 5 résume les différents effets attendus dans un transistor Ω FET.

Position	Expérience	Mobilité	Effet attendu - NMOS	Effet attendu - PMOS
0° / l'encoche	Diminution de la largeur de l'aileron	$\mu_{((001) \langle 110 \rangle)} \rightarrow \mu_{((110) \langle 110 \rangle)}$	diminution de la mobilité	augmentation de la mobilité
45° / l'encoche	Diminution de la largeur de l'aileron	$\mu_{((001) \langle 100 \rangle)} \rightarrow \mu_{((010) \langle 100 \rangle)}$	mobilité constante	mobilité constante
Transistor large	changement d'orientation cristalline	$\mu_{((001) \langle 110 \rangle)} \rightarrow \mu_{((001) \langle 100 \rangle)}$	mobilité constante	augmentation de la mobilité
Transistor étroit	changement d'orientation cristalline	canaux latéraux: $\mu_{((110) \langle 110 \rangle)} \rightarrow \mu_{((010) \langle 100 \rangle)}$ canal supérieur: $\mu_{((001) \langle 110 \rangle)} \rightarrow \mu_{((001) \langle 100 \rangle)}$	augmentation de la mobilité	diminution de la mobilité

Tab. 5: Synthèse des effets attendus sur les dispositifs Ω FETs en fonction de la direction et de l'orientation cristalline.

4.2.3 Mesures sur les transistors Ω FETs – effet de la largeur

4.2.3.1 Transistors isolés

Les transconductances normalisées par la largeur effective du canal ($W_{EFF} = W_{FIN} + 2H_{FIN}$) obtenues pour différentes largeurs d'aileron sont montrées sur la Fig. 131 (Fig. 131.a pour les NMOS et Fig. 131.b pour les PMOS). Les transistors sont choisis longs ($L_G = 10 \mu\text{m}$) et sont isolés.

L'allure de ces courbes de transconductance soulève un certain nombre de questions. Pour les transistors NMOS larges (Fig. 131.a), on observe en plus du pic principal de transconductance un replat à environ 400 mV au dessus du seuil. Ce pic semble coïncider avec le décalage de tension de seuil observé quand on passe des canaux larges aux canaux étroits. On pourrait donc penser qu'il s'agit de la tension de seuil d'activation des canaux latéraux, visible également pour des canaux larges. Cependant, cette tendance n'est pas observée pour des transistors PMOS (Fig. 131.b). On note même un décalage positif de la tension de seuil

quand la largeur de l'aileron est réduite. Le rôle joué par les résistances source/drain pourrait expliquer ces différences entre NMOS et PMOS et entre différentes largeurs d'aileron.

La valeur du pic de mobilité en fonction de la largeur de l'aileron a également une allure surprenante; le pic de transconductance semble ainsi diminuer quand on réduit la largeur de l'aileron, puis fortement augmenter pour les petites largeurs. Ainsi, on observe sur les transistors étroits une mobilité meilleure que sur les transistors larges. Nous reviendrons sur ce point dans la section 4.2.3.3.

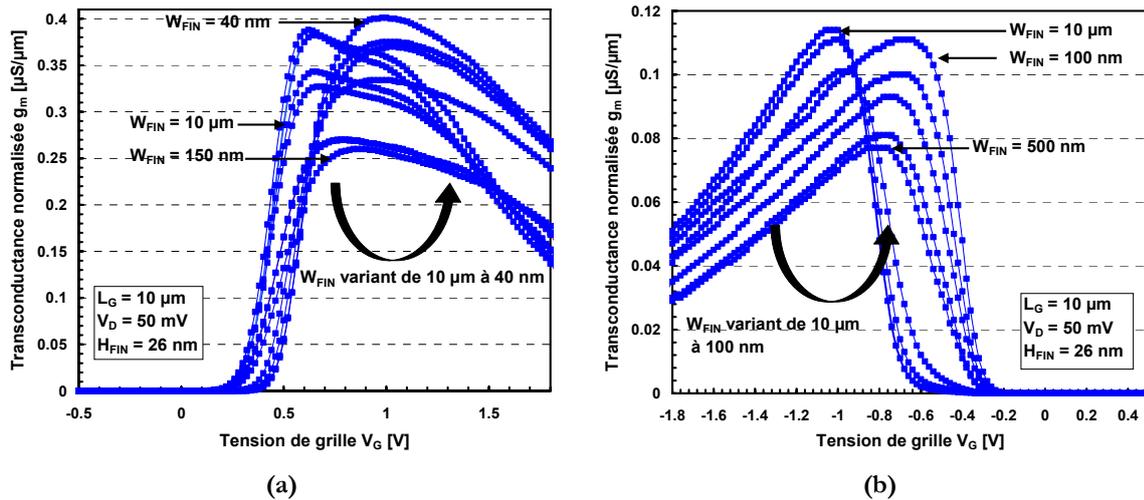


Fig. 131: Transconductance normalisée par la largeur effective du canal ($W_{\text{EFF}} = W_{\text{FIN}} + 2H_{\text{FIN}}$). La largeur de l'aileron W_{FIN} varie de 10 μm à 40 nm pour un NMOS (a) et de 10 μm à 100 nm pour un PMOS (b). $L_G = 10 \mu\text{m}$ et $H_{\text{FIN}} = 26 \text{ nm}$.

4.2.3.2 Transistors en réseaux

Des mesures similaires à celles de la section précédente ont été réalisées sur des transistors en réseau (Fig. 132.a pour les NMOS et Fig. 132.b pour les PMOS). Par rapport aux mesures sur des transistors isolés, on constate que le pic principal de transconductance se décale peu et que le replat bien que moins marqué est toujours visible. L'argument de l'activation des flancs verticaux ne tient donc plus puisque le pic de transconductance ne se décale pas.

De manière similaire à ce qui a été observé sur les transistors isolés, la mobilité pour les transistors étroits est meilleure que pour un transistor large dans le cas des NMOS et des PMOS. On contredit ainsi les prévisions s'appuyant sur les flancs cristallographiques. Pour résumer, le replat s'observe dans le cas des transistors NMOS larges et dans une moindre mesure pour les transistors NMOS en réseau. On peut donc raisonnablement penser que ce comportement est dû aux zones de source et drain.

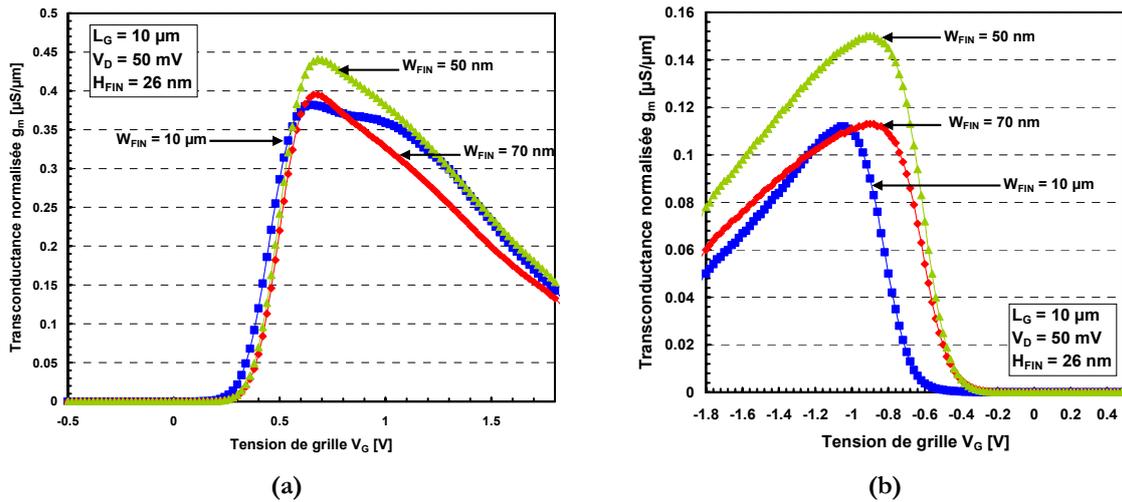


Fig. 132: Transconductance normalisée pour des NMOS (a) et PMOS (b) en réseaux. $W_{FIN} = 10 \mu\text{m}$ représente la référence obtenue avec un transistor planaire. La hauteur des ailerons H_{FIN} est de 26 nm; avec 50 ailerons en parallèle, la largeur effective totale est respectivement de $6.1 \mu\text{m}$ pour $W_{EFF} = 70 \text{ nm}$ et de $5.1 \mu\text{m}$ pour $W_{EFF} = 50 \text{ nm}$. La longueur de grille L_G est de $10 \mu\text{m}$.

4.2.3.3 Mesures de mobilité

Au vu des surprenantes courbes de transconductance observées, nous avons extrait les mobilités pour les MOS longs ($L_G = 10 \mu\text{m}$) en fonction de la largeur de l'aileron avec les méthodes décrites au début de ce chapitre. On retrouve la tendance obtenue en traçant les transconductances. Les méthodes de maximum de transconductance, de fonction Y et de dérivée seconde de $1/I_D$ présentent la même allure de courbe en 'V'. A partir de $W_{FIN} = 10 \mu\text{m}$, la mobilité diminue légèrement jusqu'à une plage comprise entre 0.1 et 0.5 μm ; ensuite, la mobilité augmente fortement jusqu'à $W_{FIN} = 50 \text{ nm}$.

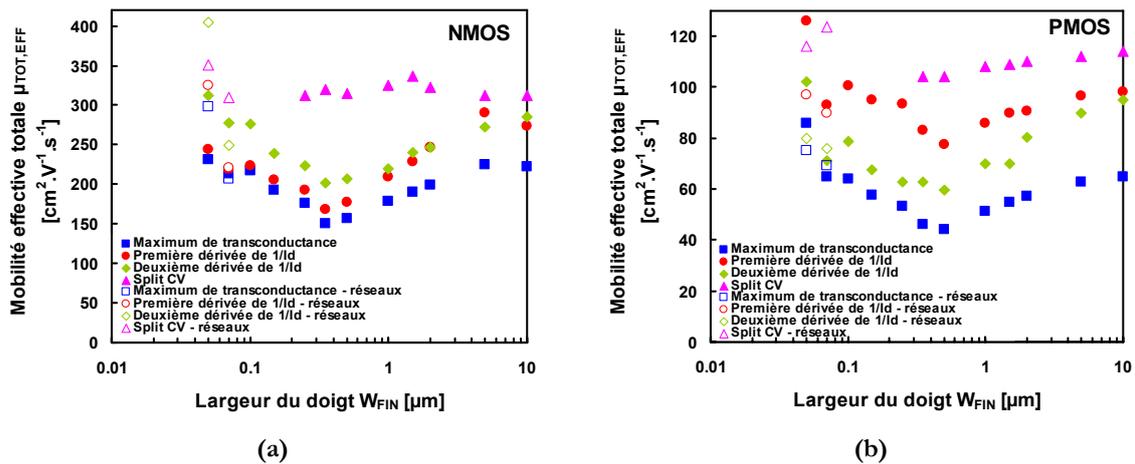


Fig. 133: Mobilités effectives totales en fonction de la largeur W_{FIN} de l'aileron pour des transistors NMOS (a) et PMOS (b). Les transistors sont isolés (symboles fermés) ou en réseaux (symboles ouverts). La hauteur de l'aileron H_{FIN} est de 26 nm dans tous les cas; la longueur de grille L_G est de $10 \mu\text{m}$. La direction cristalline est $\langle 110 \rangle$ (orientation à 0° par rapport à l'encoche).

Les méthodes les unes par rapport aux autres sont cohérentes (avec l'exception notable de la méthode split-CV). Les mesures sur des transistors isolés (symboles ouverts) ou en réseau (symboles fermés) donnent sensiblement les mêmes résultats.

On peut penser que la dégradation de mobilité en diminuant la largeur de l'aileron pourrait être due à l'influence grandissante d'une zone plus rugueuse en bord de zone active. L'incertitude sur les largeurs des ailerons peut également induire ce comportement surprenant. D'ailleurs, le changement de tendance au niveau des largeurs correspond au changement de lithographie (*'deep-UV'/'e-beam'*). La méthode Split CV permet intrinsèquement de s'affranchir de la largeur du transistor:

$$\mu_{EFF} = \frac{L}{W} \frac{I_D}{Q_{INV} V_D} = \frac{L}{W} \frac{I_D}{\int C_{OX}(V_G) dV_G} = L^2 \frac{I_D}{\int C(V_G) dV_G} \quad \text{Eq. (89)}$$

En utilisant la méthode Split-CV, la surface de capacité devient cependant trop faible en dessous de $W_{FIN} = 150$ nm. C'est là que les transistors en réseaux trouvent tout leur intérêt, conservant une grande surface effective avec une largeur d'aileron faible. Les valeurs de mobilité obtenues avec la méthode Split CV sont montrées sur la Fig. 133 (triangles). On observe bien une disparition de la tendance en 'V': la mobilité en fonction de la largeur de la grille est globalement plate, avec une légère augmentation pour les petites largeurs.

On semble donc assez loin des tendances anticipées en considérant les plans cristallins prévus sur les flancs verticaux. Les effets attendus avec un dispositif dont la direction cristalline est $\langle 110 \rangle$ sont montrés sur la Fig. 134 (cercles) à partir de la formule décrite dans l'Eq. (88). Dans la plage de largeur étudiée, on attend ainsi une dégradation de la mobilité des NMOS de 20 % environ et une augmentation de la mobilité des PMOS d'environ 66 %.

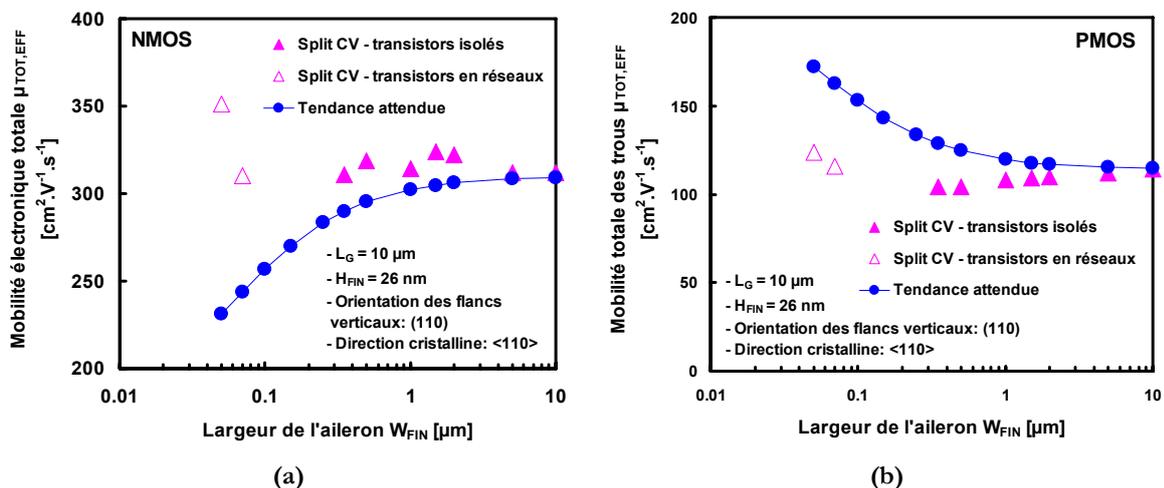


Fig. 134: Comparaison entre les valeurs obtenues expérimentalement et la tendance attendue. La courbe de tendance est calculée avec Eq. (88) à partir de la mesure effectuée à $W_{FIN} = 10 \mu\text{m}$ et du rapport de mobilité montré dans [Chang'04] et dans la Fig. 117.

4.2.4 Effet du changement d'orientation cristalline

La largeur des dispositifs étudiés, orientés à 45° par rapport à l'encoche, varie de 250 nm à 50 nm. A cause de leur faible surface, on ne peut pas utiliser la méthode Split-CV. La Fig. 135 (Fig. 135.a pour les électrons et Fig. 135.b pour les trous) présente les mobilités extraites pour ces transistors en fonction de la largeur d'aileron avec trois méthodes d'extraction. Les courbes obtenues avec des transistors orientés à 0° par rapport à l'encoche sont également présentées à titre de comparaison.

Du fait du changement de plan cristallin en passant de 0 à 45° (les flancs verticaux passent de l'orientation (110) à (100)), on attend une variation de mobilité pour les transistors étroits. On prévoit ainsi une forte augmentation de la mobilité des NMOS et une diminution dans le cas des PMOS; de manière quantitative, on attend environ + 50 % de variation pour les NMOS et environ - 25 % pour les PMOS par rapport aux dispositifs larges.

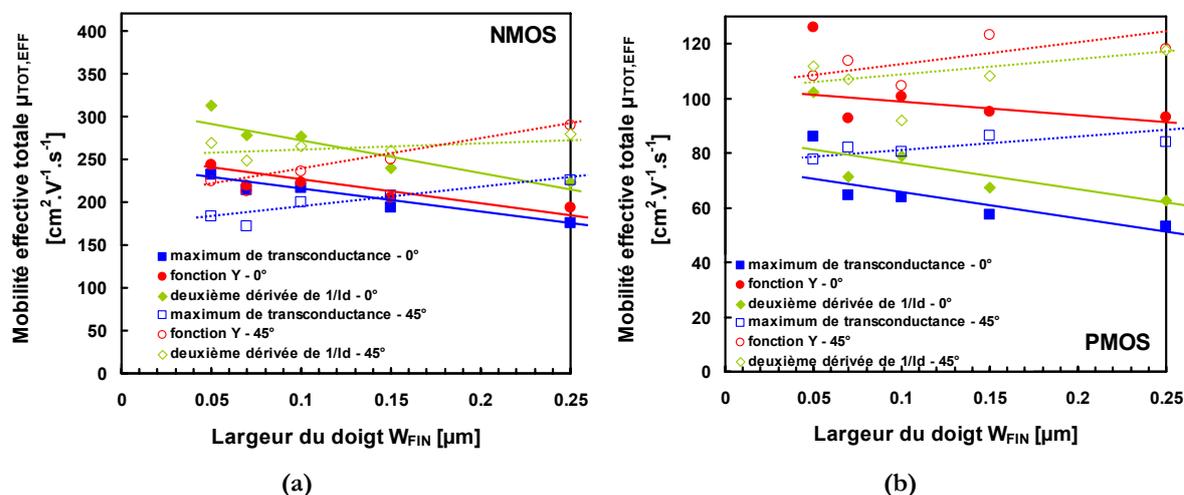


Fig. 135: Comparaison des mobilités obtenues en fonction de la largeur de grille pour un changement d'orientation cristalline dans le cas des électrons (a) et des trous (b). Les symboles fermés représentent les transistors orientés dans la direction $\langle 110 \rangle$ et les symboles ouverts les transistors orientés dans la direction $\langle 100 \rangle$.

Sur la Fig. 135.a, la tendance pour les dispositifs NMOS orientés dans la direction $\langle 100 \rangle$ (à 45° par rapport à l'encoche) va dans le sens d'une dégradation des performances pour les dispositifs étroits. Dans la direction $\langle 110 \rangle$ (à 0° par rapport à l'encoche), c'est l'inverse. Pour les trous (Fig. 135.b), la mobilité est plus élevée avec une orientation cristalline $\langle 100 \rangle$ par rapport à une orientation $\langle 110 \rangle$, mais les tendances vont vers un rapprochement des mobilités pour les dispositifs étroits.

Malgré la forte dispersion des courbes, il est donc clair que les variations attendues ne sont pas observées et que les effets de changement de plans cristallins ne sont pas visibles. La rugosité des flancs verticaux ne peut être l'explication puisqu'elle diminuerait systématiquement les mobilités mesurées, ce qui n'est pas le cas. D'après la photo de la Fig. 46.a, l'épaisseur d'oxyde t_{OX} déposé sur les flancs et sur le sommet du canal est identique et n'engendre donc pas de variation de la capacité d'oxyde.

Ces variations de mobilité ne peuvent donc pas être expliquées simplement. La prochaine section s'intéressera à l'influence de la contrainte sur la mobilité, et l'influence de défauts chargés (collisions coulombiennes) sera étudiée dans la section 5.

4.2.5 Etude de l'effet de la contrainte induite par la grille

La contrainte sur le canal induite par la grille en TiN dans les dispositifs étroits pourrait être une explication de l'amélioration de la mobilité, conjointement pour NMOS et PMOS, observée entre les dispositifs larges et les dispositifs étroits.

La contrainte induite par la grille peut être calculée optiquement en mesurant la différence de rayon de courbure d'une plaque avant et après le dépôt de TiN, et est de l'ordre de 1 GPa. Des simulations numériques 3D (utilisant CASTEM, outil développé au CEA) permettent d'avoir un ordre d'idée du champ des contraintes dans la structure. Pour les dispositifs étroits, on obtient une contrainte quasi-nulle dans le sens de la longueur (Fig. 136.a) et une contrainte globalement en tension dans le sens de la largeur (Fig. 136.b).

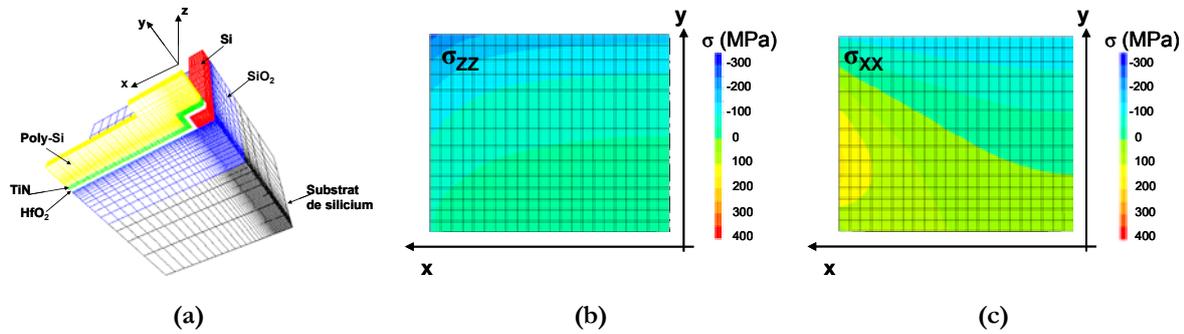


Fig. 136: (a): Structure de simulation utilisée montrant les coordonnées. (b): Champ des contraintes σ_{zz} (le long de l'axe source/drain) en milieu de grille ($\hat{a} L_G/2$). (c): Champ des contraintes σ_{xx} (le long de l'axe de la grille) en milieu de grille ($\hat{a} L_G/2$). $L_G = 30$ nm, $W_{FIN} = 30$ nm, $H_{FIN} = 26$ nm.

La variation de mobilité induite par la contrainte σ s'exprime avec la relation:

$$\frac{\Delta\mu}{\mu} = -\Pi\sigma \quad \text{Eq. (90)}$$

où Π est un tenseur dit de piézorésistivité.

Le tenseur de piézorésistivité Π et les coefficients piézoélectriques du silicium [Wortman'65] sont montrés respectivement dans la Fig. 137 et le Tab. 6.

$$\begin{pmatrix} \delta\rho_1 \\ \delta\rho_2 \\ \delta\rho_3 \\ \delta\rho_4 \\ \delta\rho_5 \\ \delta\rho_6 \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\pi_{11} + \pi_{12} + \pi_{44}}{2} & \frac{\pi_{11} + \pi_{12} - \pi_{44}}{2} & \pi_{12} & 0 & 0 & 0 \\ \frac{\pi_{11} + \pi_{12} - \pi_{44}}{2} & \frac{\pi_{11} + \pi_{12} + \pi_{44}}{2} & \pi_{12} & 0 & 0 & 0 \\ 2 & 2 & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{11} - \pi_{12} \end{pmatrix}}_{\Pi} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{pmatrix}$$

Fig. 137: Expression du tenseur de piézorésistivité Π .

Transport	Direction <100>		Direction <110>		
	<100>	<010>, <001> ou <110>	<110>	<-110>	<001>
Terme(s)	$-\pi_{11}$	$-\pi_{12}$	$-(\pi_{11} + \pi_{12} + \pi_{44})/2$	$-(\pi_{11} + \pi_{12} - \pi_{44})/2$	$-\pi_{12}$
NMOS	102	-53.4	31.6	17.6	-53.4
PMOS	-6.6	1.1	-71.8	66.3	1.1

Tab. 6: Coefficients $-\pi$ intervenant dans le cas d'une contrainte uniaxiale. Une contrainte en tension (resp. compression) est bénéfique si le coefficient correspondant est positif (resp. négatif) (d'après [Andrieu'05]).

Une contrainte en tension ($\sigma > 0$) sera d'autant plus bénéfique que le coefficient $-\pi$ sera positif et grand. Par conséquent, pour un dispositif orienté dans la direction <110>, une contrainte en tension dans le sens de la largeur (contrainte dans la direction <-110>, voir Fig. 125.a) améliore la mobilité pour les NMOS et les PMOS.

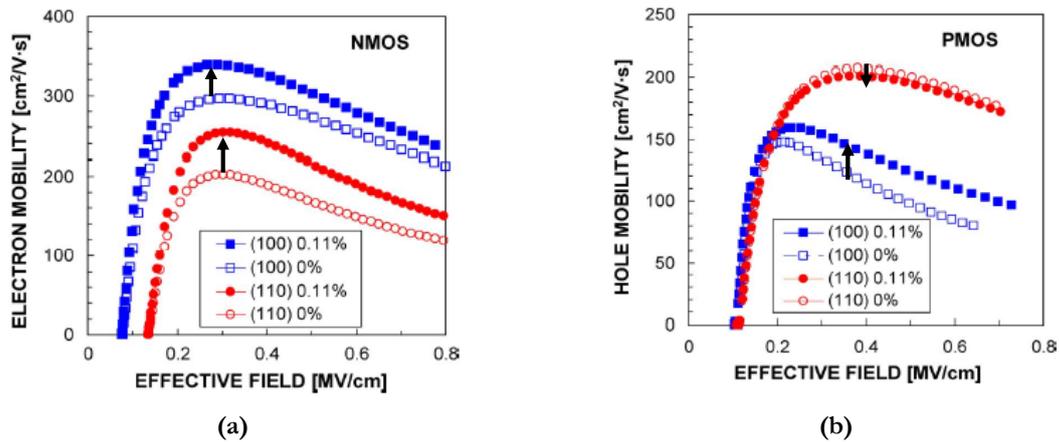
Pour un dispositif dans la direction <100>, une contrainte en tension dans le sens de la largeur (contrainte dans la direction <010>, voir Fig. 125.a) entraîne une dégradation de mobilité pour les NMOS et n'a quasiment pas d'influence pour les PMOS.

Quantitativement, la variation de mobilité simulée entre un dispositif large (supposé sans contrainte) et étroit ($W_{\text{FIN}} = 30$ nm, contrainte induite par la grille) est montrée dans le Tab. 7. Ces dispositifs sont orientés dans la direction <110>; on obtient ainsi une amélioration de mobilité pour les dispositifs étroits dans le cas des NMOS et des PMOS. Pour un NMOS, la variation calculée est respectivement de 4% et 13% pour des plans (100) et (110). Pour des PMOS, la variation est respectivement de 1% et 14% pour des plans (100) et (110). Par rapport aux mesures de la Fig. 133 et de la Fig. 134, la contrainte en tension induite par la grille est donc une explication plausible pour expliquer l'augmentation légère de mobilité observée expérimentalement.

Transport	Direction <110>			
	NMOS		PMOS	
Type	(100)	(110)	(100)	(110)
$\Delta\mu/\mu$	+ 4 %	+ 13 %	+ 1 %	+ 14 %

Tab. 7: Variation de mobilité simulée entre un dispositif large (supposé sans contrainte) et étroit ($W_{\text{FIN}} = 30$ nm, contrainte induite par la grille). $L_G = 10$ μm , $W_{\text{FIN}} = 30$ nm, $H_{\text{FIN}} = 26$ nm (simulations C. Dupré).

Des données expérimentales publiées récemment vont dans le même sens [Shin'06]. Pour ces travaux, la contrainte en tension est induite par une courbure imposée à la plaque, et n'est pas engendrée par la grille comme dans notre cas. L'accord mesure-théorie obtenu est excellent (Fig. 138.c) et montre une amélioration de mobilité entre 8 et 25 %.



	STRESS (MPa)			AMELIORATION DE LA MOBILITE (%)			
	σ_{XX}	σ_{YY}	σ_{ZZ}	(100) NMOS	(110) NMOS	(100) PMOS	(110) PMOS
Expérience				14.4	25.9	8.1	-2.6
Théorie ^(*)	290	350	-160	14.4	25.6	1.7	-2.6

(*) d'après [Smith'54]

(c)

Fig. 138: (a)-(b): Impact d'une contrainte sur la mobilité pour un NMOS (a) et PMOS (b). L'effet pour des plans (100) et (110) est présenté. La contrainte de surface est d'environ 200 MPa (~ 0.11 %). (c): Comparaison des décalages de mobilité théoriques et expérimentaux (d'après [Shin'06]).

Il faut maintenant examiner si la contrainte permet d'expliquer les variations de la Fig. 135. Dans la Fig. 135.a (cas des NMOS), la contrainte uniaxiale dans le sens de la largeur devrait donc améliorer la mobilité des transistors orientés dans la direction $\langle 110 \rangle$ et dégrader la mobilité des transistors orientés dans la direction $\langle 100 \rangle$. Pour les PMOS (Fig. 135.b), la contrainte uniaxiale dans le sens de la largeur devrait également améliorer la mobilité des transistors orientés dans la direction $\langle 110 \rangle$ et ne pas avoir de grande influence sur la mobilité des transistors orientés dans la direction $\langle 100 \rangle$.

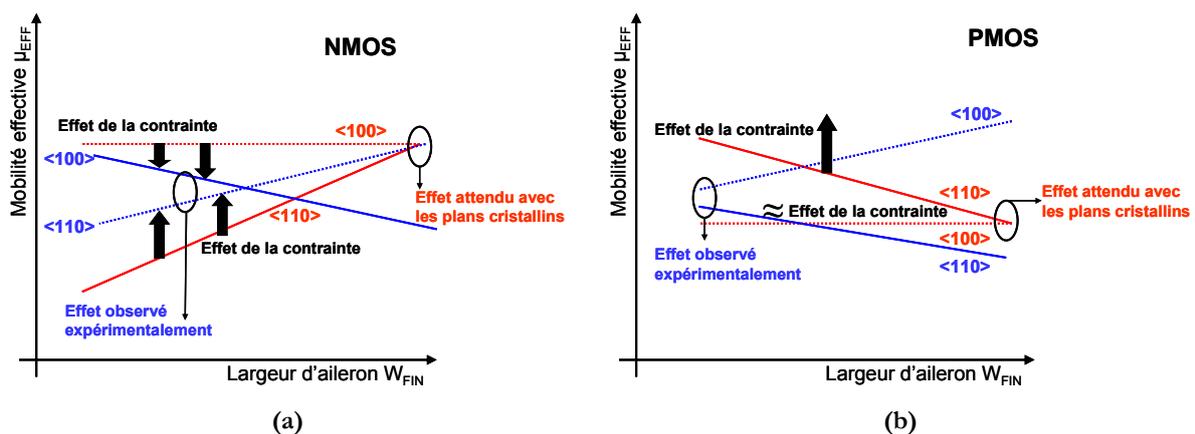


Fig. 139: Comparatif des mesures expérimentales avec l'influence attendue des effets de plan cristallin et de la contrainte pour les NMOS (a) et les PMOS (b).

La Fig. 139 synthétise les effets observés expérimentalement avec les effets attendus de plan cristallin et de contrainte pour les NMOS (Fig. 139.a) et les PMOS (Fig. 139.b). Pour les électrons, l'influence croisée de la contrainte et de la mobilité des différents plans cristallins est

telle qu'elle pourrait expliquer la faible différence de mobilité observée entre les directions cristallines $\langle 110 \rangle$ et $\langle 100 \rangle$. La contrainte induite par la grille étant supposée de plus en plus importante pour les dispositifs étroits, elle n'explique pas par contre le croisement de mobilité observé vers $W_{\text{FIN}} = 150$ nm.

Pour les dispositifs PMOS orientés dans la direction $\langle 110 \rangle$, l'effet de plan cristallin et de la contrainte vont dans le sens d'une augmentation de mobilité pour les faibles largeurs. C'est ce qu'on observe expérimentalement. Par contre, la dégradation de mobilité dans la direction $\langle 100 \rangle$ observée expérimentalement ne s'explique ni avec la contrainte ni avec l'influence des plans cristallins.

Bien que séduisante, la contrainte ne permet donc pas d'expliquer totalement les tendances observées expérimentalement.

5 Mesures en température

Les mesures à basses températures ont pour intérêt de décorréler les différents mécanismes de dégradation de mobilité. En particulier, si la gravure des canaux étroits induit des charges, la mobilité à basse température devrait en être d'autant affectée puisque la mobilité à faible champ et à basse température est limitée par les collisions coulombiennes.

5.1 Variation des caractéristiques avec la température

La tension de seuil d'un transistor est reliée à la température par le biais du potentiel de Fermi, que ce soit dans le cas d'un transistor sur silicium massif [Sze'81] ou 'complètement déserté' [Poiroux'05].

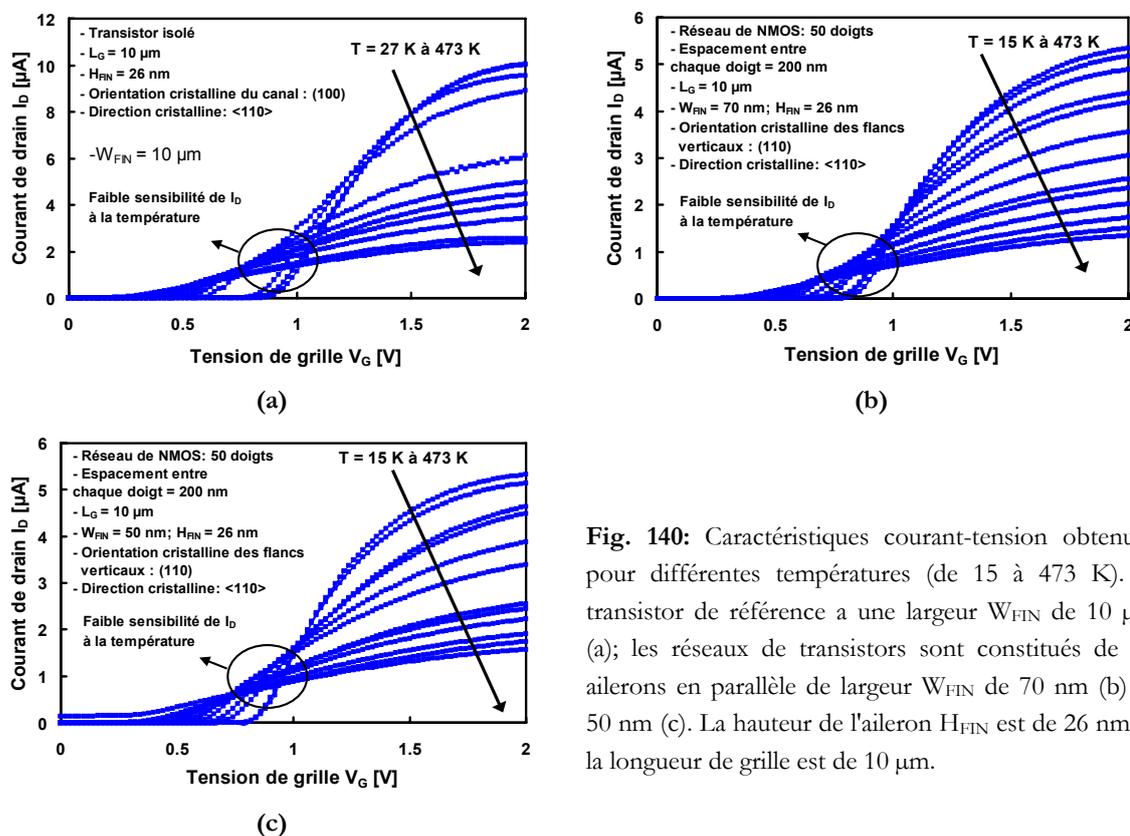


Fig. 140: Caractéristiques courant-tension obtenues pour différentes températures (de 15 à 473 K). le transistor de référence a une largeur W_{FIN} de 10 μm (a); les réseaux de transistors sont constitués de 50 ailerons en parallèle de largeur W_{FIN} de 70 nm (b) et 50 nm (c). La hauteur de l'aileron H_{FIN} est de 26 nm et la longueur de grille est de 10 μm .

Quand la température diminue, la concentration intrinsèque de porteurs décroît fortement et le potentiel de Fermi croît. Par conséquent, la tension de seuil d'un transistor augmente quand la température diminue. Nous avons également vu que la mobilité augmentait quand la température diminuait. Légèrement au-dessus du seuil, ces deux effets se contrebalancent pour former une zone où le courant de drain varie peu quelque soit la température (*ZTC*, *Zero Temperature Coefficient* [Osman'95]). C'est exactement ce qui est observable sur la Fig. 140 pour un transistor large (Fig. 140.a) ou pour des transistors en réseaux (Fig. 140.b et Fig. 140.c).

5.2 Mesures sur des transistors longs et larges

Sur des transistors longs ($L_G = 10 \mu\text{m}$ pour ce paragraphe), on peut négliger l'effet des R_{SD} . La méthode d'extraction de la mobilité choisie pour les études en température est celle du pic de transconductance. Ça n'est pas la plus précise, mais elle permet de minimiser les incertitudes de mesure.

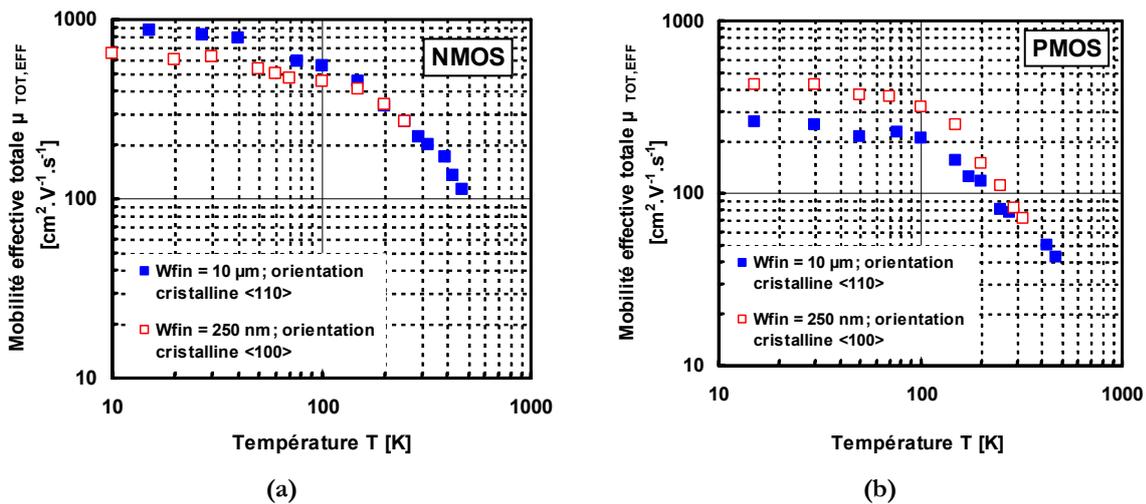


Fig. 141: Mobilité en fonction de la température pour des transistors NMOS (a) et PMOS (b) orientés à 0° ou 45° par rapport à l'encoche. La largeur des ailerons W_{FIN} vaut $10 \mu\text{m}$ pour les dispositifs dans la direction $\langle 110 \rangle$ et 250 nm dans la direction $\langle 100 \rangle$. Les valeurs ont été extraites avec le maximum de transconductance. La longueur de grille est $L_G = 10 \mu\text{m}$.

Sur la Fig. 141, on compare les mobilités obtenues en changeant la direction cristalline ($\langle 110 \rangle$ dans l'axe de l'encoche, ou $\langle 100 \rangle$ à 45° par rapport à l'encoche) pour des transistors larges. Les résultats obtenus pour les électrons sont montrés dans la Fig. 141.a et ceux obtenus pour les trous sont montrés dans la Fig. 141.b. Expérimentalement, on observe une meilleure mobilité à basse température dans la direction $\langle 110 \rangle$ pour les NMOS et dans la direction $\langle 100 \rangle$ dans les PMOS. Or, sur ces transistors longs et larges, on peut supposer que la contrainte induite par la grille est faible; par conséquent, les courbes de mobilité devraient se superposer.

Malheureusement, les largeurs des dispositifs ne sont pas les mêmes ($W_{FIN} = 10 \mu\text{m}$ pour la direction $\langle 110 \rangle$ et $W_{FIN} = 250 \text{ nm}$ dans la direction $\langle 100 \rangle$); or nous avons vu

qu'avec la méthode du pic de transconductance la mobilité diminuait fortement entre 10 μm et 250 nm; une certaine précaution dans l'interprétation des résultats s'impose donc.

5.3 Mesures sur des transistors longs et étroits

Dans la Fig. 142, on compare les mobilités obtenues pour des dispositifs dans la direction cristalline $\langle 110 \rangle$ et ayant des largeurs d'aileron différentes ($W_{\text{FIN}} = 50 \text{ nm}$, 70 nm et 10 μm). Les résultats obtenus pour les électrons sont montrés dans la Fig. 142.a et ceux obtenus pour les trous sont montrés dans la Fig. 142.b.

Dans la direction cristalline $\langle 110 \rangle$, les flancs verticaux sont dans un plan (110) alors que le canal horizontal est dans un plan (100). A basse température, les plans cristallins conservent leurs propriétés, c'est-à-dire que la mobilité des électrons est optimisée dans un plan (100), et celle des trous dans un plan (110) [Aoki'89].

Que ce soit pour les électrons (Fig. 142.a) ou pour les trous (Fig. 142.b), la variation de mobilité à basse température avec la largeur de l'aileron est faible pour les NMOS et pour les PMOS. Les effets de flancs cristallins ne sont donc pas visibles.

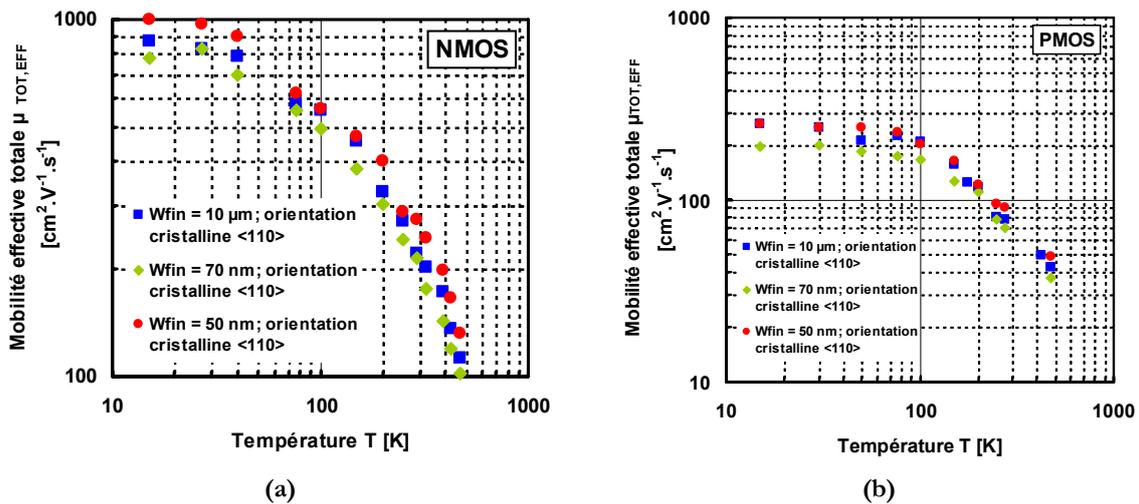


Fig. 142: Mobilité en fonction de la température pour des transistors NMOS (a) et PMOS (b) orientés dans la direction $\langle 110 \rangle$. La largeur des ailerons W_{FIN} est de 10 μm , 70 nm et 50 nm. Les valeurs ont été extraites avec le maximum de transconductance. La longueur de grille est $L_G = 10 \mu\text{m}$.

La contrainte appliquée sur le canal induit et le silicium non contraint sera alors faible. Autrement dit, réduire la température revient à minimiser l'influence de la contrainte.

Pour des NMOS étroits, l'effet de l'orientation cristalline et/ou des éventuelles charges sur les flancs verticaux (110) conduirait à une forte dégradation de la mobilité à basse température. Pour des PMOS étroits, l'effet de l'orientation cristalline des flancs verticaux (110) conduirait à une augmentation de mobilité; dans ce cas-là, l'effet conjugué des flancs cristallins et des charges éventuelles pourrait se compenser.

De l'absence de variation de mobilité entre les dispositifs larges et étroits à basse température, on peut donc déduire que les collisions coulombiennes ne sont pas beaucoup plus importantes dans les canaux étroits que dans les canaux larges. Les charges éventuelles

induites par le procédé sur les flancs verticaux ne peuvent pas expliquer les variations de mobilité observées à température ambiante.

6 Mesure du taux de balisticité

Un porteur est dit en régime balistique s'il passe de la source vers le drain sans collision dans le canal. Au fur et à mesure que les dimensions des transistors sont réduites, cette probabilité de passage direct entre source et drain augmente, et conduit à repenser le concept classique de mobilité qui s'appuie sur les collisions lors du déplacement d'un porteur. Le régime balistique fournit donc le maximum de courant possible pour une technologie donnée.

Le coefficient de rétrodiffusion r (Fig. 143.a) est défini comme étant la proportion de porteurs ayant suffisamment d'énergie pour retourner à la source après une interaction; ces électrons ne participent donc pas au courant. Le taux de balisticité $BR = (1-r)/(1+r)$ est défini quant à lui comme étant le rapport entre le courant mesuré et le courant qu'on aurait en régime totalement balistique. Les transistors fortement sub-microniques pourraient ainsi fonctionner avec un taux de balisticité proche de 1.

En mesurant le courant de sortie en saturation pour un jeu de températures proches (ici 50, 60 et 70 K), on peut déterminer expérimentalement le coefficient de rétrodiffusion et le taux de balisticité [Barral'06][Chen'02].

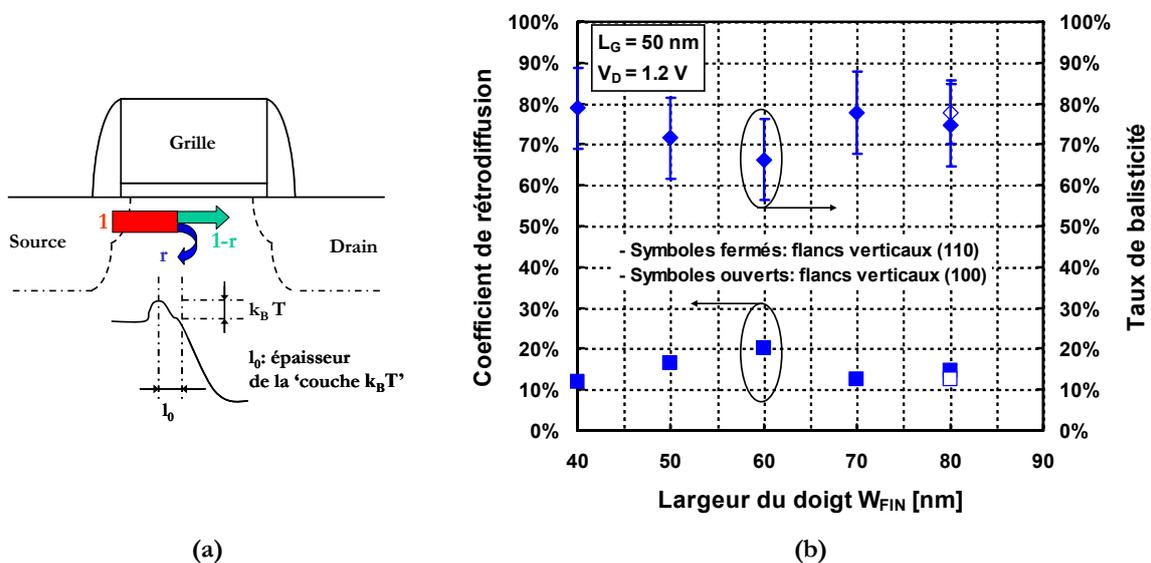


Fig. 143: Illustration schématique du coefficient de rétrodiffusion (a); calcul du coefficient de rétrodiffusion et du taux de balisticité en fonction de la largeur de l'aileron pour une longueur de grille de 50 nm à partir des mesures expérimentales à 50, 60 et 70 K (b).

Sur la Fig. 143.b, on extrait le coefficient de rétrodiffusion et le taux de balisticité pour des transistors Ω FETs. La longueur de grille est égale à 50 nm et la largeur de l'aileron varie de 40 à 80 nm. La largeur de l'aileron modifie le contrôle électrostatique de la grille et pourrait donc améliorer le taux de balisticité pour des transistors étroits. Cependant, on n'observe pas de variation importante du taux de balisticité en réduisant la largeur de l'aileron.

Un changement d'orientation cristalline des flancs est susceptible de modifier la rugosité de la surface des flancs, donc le taux de balisticité. Cependant, une mesure effectuée avec un transistor orienté à 45° par rapport à l'encoche ($W_{\text{FIN}} = 80$ nm, symbole ouvert) donne une valeur très proche de ce qu'on mesure avec un transistor orienté à 90° par rapport à l'encoche ($W_{\text{FIN}} = 80$ nm, symbole fermé). Les paragraphes précédents ayant montré que le changement de direction cristalline ne s'accompagnait pas d'un changement observable de direction cristalline des flancs, il faut cependant interpréter ces résultats avec circonspection.

La valeur calculée du taux de balisticité est proche de 75 % qui peut paraître un peu élevée, surtout au vu de la longueur de grille ($L_G = 50$ nm). Cette valeur vient de la relation entre le libre parcours moyen et le coefficient de rétrodiffusion. Cette approche est actuellement en débat.

7 Conclusions

Les mesures effectuées dans ce chapitre portent sur les phénomènes de transport dans les transistors Ω FETs. Elles permettent de dégager plusieurs phénomènes:

- ✓ La mobilité chute quand la longueur de grille diminue. Cette dégradation peut être attribuée aux charges additionnelles dans les régions de source et de drain ou bien encore aux collisions avec des défauts neutres.
- ✓ La mobilité en fonction de la largeur de l'aileron présente une forme caractéristique en 'V' pour certaines méthodes d'extraction. En utilisant la méthode Split CV, on normalise l'influence de la largeur de l'aileron et on n'obtient plus cette forme.
- ✓ En changeant plans et directions cristallines, on n'observe pas les variations attendues. La contrainte en tension induite par la grille permet d'expliquer l'amélioration de mobilité pour les transistors étroits orientés dans la direction $\langle 110 \rangle$ observée à la fois pour les NMOS et les PMOS. La contrainte induite par la grille ne permet cependant pas d'expliquer toutes les variations observées.
- ✓ Le comportement observé à basse température montre que les dispositifs étroits ont une mobilité comparable à celle de dispositifs larges. Les interactions coulombiennes sont comparables quelque soit la largeur du dispositif. Les charges induites par le procédé en bord de canal ne sont donc pas responsables des tendances observées. Au final, tout se passe comme si les flancs verticaux n'avaient pas les orientations cristallines attendues et que les variations de mobilité étaient contrôlées par la contrainte imposée par la grille.
- ✓ Le taux de balisticité extrait est élevé (75 %) pour une longueur de grille de 50 nm. Le taux de balisticité ne semble pas dépendre de la largeur de l'aileron.

Références du chapitre 5

- [Andrieu'05] F. Andrieu, "Transistor CMOS décanométrique à canaux contraints sur Si massif ou sur SOI – fabrication, caractérisation et étude du transport", *thèse de doctorat*, INPG, 2005.
- [Aoki'89] M. Aoki, K. Yano, T. Masuhara, K. Shimohigashi, "Fully symmetric cooled CMOS on (110) plane", *Electron Devices, IEEE Transactions on*, vol. 36, no. 8, pp. 1429-1433, 1989.
- [Barral'06] V. Barral, T. Poiroux, M. Vinet, J. Widiez, B. Previtali, P. Grosgeorges, G. Le Carval, S. Barraud, S. Deleonibus, D. Munteanu, "Experimental determination of the channel backscattering coefficient on 10nm-metal-gate Double-Gate transistors", proceedings of the *7th European Workshop on Ultimate Integration of Silicon (ULIS 2006)*, pp. 13-16, 2006.
- [Byu] http://www.ece.byu.edu/cleanroom/EW_orientation.phtml, site de l'Integrated Microfabrication Lab (IML) de la Brigham Young University (USA, Utah).
- [Chang'04] L. Chang, M. Jeong, M. Yang, "CMOS Circuit Performance Enhancement by Surface Orientation Optimization", *Electron Devices, IEEE Transactions on*, vol. 51 no. 10, pp. 1621-1627, 2004.
- [Chen'02] M.-J. Chen, H.-T. Huang, K.-C. Huang, P.-N. Chen, C.-S. Chang, C. Diaz, "Temperature Dependant Channel Backscattering Coefficients in Nanoscale MOSFETs", *IEDM'02 Technical Digest*, pp. 39-42, 2002.
- [Collaert'06] N. Collaert, R. Rooyackers, F. Clemente, P. Zimmerman, I. Cayrefourcq, B. Ghyselen, K.T. San, B. Eyckens, M. Jurczak, S. Biesemans, "Performance enhancement of MUGFET devices using Super Critical Strained-SOI (SC-SSOI) and CESL", *VLSI Tech. Dig.*, 2006. (à paraître)
- [Dupré'05] C. Dupré, "Evaluation de l'impact des déformations mécaniques et des défauts cristallins sur la mobilité lors de la fabrication de transistors MOSFET trois grilles à canaux contraints", *rapport stage DEA*, CEA-LETI, 2005.
- [Faynot'94] O. Faynot, S. Cristoloveanu, P. McLarty, C. Raynaud, J. Gautier, "A New Parameter Extraction Method for Ultra-Thin Oxide SOI MOSFETs", *IEEE Int. SOI conf.*, pp. 17-18, 1994.

- [Ghibaudo'88] G. Ghibaudo, "New method for the extraction of MOSFET parameters", *Electronic Letters*, vol. 24, no. 9, pp. 543-545, 1988.
- [Guillaume'05] T. Guillaume, M. Mouis, "Variations of hole mass in p-MOSFETs under process-induced mechanical stress", proceedings of the *35th European Solid-State Device Research conference (ESSDERC'05)*, pp. 289-292, 2005.
- [Jahan'05] C. Jahan, O. Faynot, M. Cassé, R. Ritzenthaler, L. Brévard, L. Tosti, X. Garros, C. Vizios, F. Allain, A.M. Papon, H. Dansas, F. Martin, M. Vinet, B. Guillaumot, A. Toffoli, B. Giffard and S. Deleonibus, "10nm Ω FETs transistors with TiN metal gate and HfO₂", *VLSI Tech. Dig.*, pp. 112-113, 2005.
- [Kavalieros'06] J. Kavalieros, B. Doyle, S. Datta, G. Dewey, M. Doczy, B. Jin, D. Lionberger, M. Metz, W. Rachmady, M. Radosavljevic, U. Shah, N. Zelick, R. Chau, "Tri-Gate Transistor Architecture with High-k Gate Dielectrics, Metal Gates and Strain Engineering", *VLSI Tech. Dig.*, 2006. (à paraître)
- [Kittel'96] C. Kittel, "Introduction to Solid State Physics", 7th edition, *John Wiley and Sons (WIE)*, ISBN 0471111813, 1996.
- [Koomen'73] J.Koomen, "Investigation of the MOST channel conductance in weak inversion", *Solid-State Electronics*, vol. 16 no. 7, pp. 801-810, 1973.
- [Mathieu'96] H. Mathieu, "Physique des semiconducteurs et des composants électroniques", 3^{ème} édition, *Ed. Masson*, ISBN 2-225-85124-7, 1996.
- [McLarty'95] P. K. McLarty, S. Cristoloveanu, O. Faynot, V. Misra, J. R. Hauser, J. J. Wortman, "A Simple Parameter Extraction Method for Ultra-Thin Oxide MOSFETs", *Solid-State Electronics*, vol. 38 no. 6, pp. 1175-1177, 1995.
- [Osman'95] A. A. Osman, M. A. Osman, N. S. Dogan, M. A. Imam, "Zero-temperature-coefficient biasing point of partially depleted SOI MOSFETs", *Electron Devices, IEEE Transactions on*, vol. 42, no. 9, pp. 1709-1711, 1995.
- [Poiroux'05] T. Poiroux, M. Vinet, O. Faynot, J. Widiez, J. Lolivier, T. Ernst, B. Previtali, S. Deleonibus, "Multiple gate devices: advantages and challenges", *Microelectronic Engineering*, vol. 80, pp. 378-385, 2005.
- [Prégaldiny'02] F. Prégaldiny, C. Lallement, D. Mathiot, "A simple efficient model of parasitic capacitances of deep-submicron LDD MOSFETs," *Solid-State Electronics*, vol. 46, no. 12, pp. 2191–2198, 2002.
- [Pretet'03] J. Pretet, 'Analyse de transistors SOI à dimensions critiques et architectures multi-grilles', thèse INPG, p. 146, 2003.

- [Reichert'98] G. Reichert, "Etude en haute température des transistors MOS submicroniques fabriqués sur silicium sur isolant", *thèse de doctorat*, INPG, 1998.
- [Romanjek'04] K. Romanjek, F. Andrieu, T. Ernst, G. Ghibaudo, "Improved Split C–V Method for Effective Mobility Extraction in sub-0.1- μ m Si MOSFETs", *IEEE Electron Device Letters*, vol. 25, no. 8, pp. 583-585, 2004.
- [Romanjek'04-2] K. Romanjek, "Caractérisation et modélisation des transistors de Lg 50nm et en deçà", *thèse de doctorat*, INPG, p. 154, 2004.
- [Rousseau'95] J.-J. Rousseau, "Cristallographie géométrique et radiocristallographie", *Ed. Masson*, ISBN 2-225-84990-0, 1995.
- [Saito'06] S. Saito, D. Hisamoto, Y. Kimura, N. Sugii, R. Tsuchiya, K. Torii, S. Kimura, "Origin of Drivability Enhancement in Scaled pMOSFETs with 45° Rotated <100> channels", *VLSI Tech. Dig.*, 2006. (à paraître)
- [Sayama'99] H. Sayama, Y. Nishida, H. Oda, T. Oishi, S. Shimizu, T. Kunikiyo, K. Sonoda, Y. Inoue and M. Inuishi, "Effect of <100> Channel Direction for High Performance SCE Immune pMOSFET with Less Than 0.15 μ m Gate Length", *IEDM'99 Technical Digest*, pp. 657-660, 1999.
- [Shin'06] K. Shin, W. Xiong, C. Y. Cho, C. R. Cleavelin, T. Schulz, K. Schuefer, P. Patruno, Lee Smith, T.-J. King, " Study of Bending-Induced Strain Effects on MuGFET Performance", *IEEE Electron Device Letters*, vol. 27, no. 8, pp. 671-673, 2006.
- [Smith'54] C. Smith, "Piezoresistance effect in germanium and silicon", *Physical Review*, vol. 94, no. 1, pp. 42-49, 1954.
- [Sze'81] S. M. Sze, "Physics of Semiconductor Devices (2nd ed.)", *John Wiley and Sons (WIE)*. ISBN 0-47-105661-8, 1981.
- [Takagi'94-1] S.-I. Takagi, A. Toriumi, M. Iwase, H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFETs: Part 1-Effects of Surface Impurity Concentration", *Electron Devices, IEEE Transactions on*, vol. 41, no. 12, pp. 2357-2362, 1994.
- [Takagi'94-2] S.-I. Takagi, A. Toriumi, M. Iwase, H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFETs: Part 2-Effects of Surface Orientation", *Electron Devices, IEEE Transactions on*, vol. 41, no. 12, pp. 2363-2368, 1994.

- [Wang'94] Z. Wang, "Modélisation de la Piézorésistivité du Silicium : Application à la Simulation de Dispositifs M.O.S.", *thèse de doctorat*, l'Université des Sciences et Techniques de Lille Flandre-Artois, p. 16, 1994.
- [Widiez'05] J. Widiez, T. Poiroux, M. Vinet, B. Previtali, M. Mouis and S. Deleonibus, "Experimental low field transport investigation in sub-0.1 μ m Ultra-thin SOI Single and Double Gate MOSFETs" in proc. of the *Silicon Nanoelectronics Workshop 2005*, June 2005.
- [Widiez'05-2] J. Widiez, "Etude, fabrication et caractérisation de transistors CMOS double grille planaires déca-nanométriques", *thèse de doctorat*, INPG, 2005.
- [Wortman'65] J. J. Wortman, R. A. Evans, "Young's Modulus, Shear Modulus, and Poisson's ratio in Silicon and Germanium", *J. Appl. Phys. Dig.*, vol.36, no. 1, pp. 153-156, 1965.
- [Yang'03] M. Yang, M. leong, L. Shi, K. Chan, V. Chant, A. Chout, E. Gusev, K. Jenkins, D. Boyd, Y. Ninomiya, D. Pendleton, Y. Surpris, D. Heenan, J. Ott, K. Guarini, C. D'Emic, M. Cobb, P. Mooney, B. To, N. Rovedo, J. Benedict, R. MO and H. Ng, "High Performance CMOS Fabricated on Hybrid Substrate With Different Crystal Orientations", *IEDM'03 Technical Digest*, pp. 453-456, 2003.
- [Yang'06] M. Yang, K. Chan, A. Kumar, S.-H. Lo, J. Sleight, L. Chang, R. Rao, S. Bedell, A. Ray, J. Ott, J. Patel, C. D'Emic, J. Rubino, Y. Zhang, L. Shi, S. Steen, E. Sikorski, J. Newbury, R. Meyer, B. To, P. Kozlowski, W. Graham, S. Maurer, S. Medd, D. Canaperi, L. Deligianni, J. Tornello, G. Gibson, T. Dalton, M. leong, G. Shahidi, "Silicon-on-Insulator MOSFETs with Hybrid Crystal Orientations", *VLSI Tech. Dig.*, 2006. (à paraître)

Conclusions et perspectives du manuscrit

Conclusions du manuscrit :

Supportée par les investissements massifs de l'industrie du semiconducteur, la recherche en microélectronique évolue très vite. Ce travail de thèse s'inscrit dans ce contexte. Les transistors FinFETs sont actuellement des dispositifs très attractifs, comme le montre le nombre important de publications récentes sur ce sujet.

L'objectif de cette thèse était la fabrication et la caractérisation de transistors multigrilles non planaires (familles des FinFETs). Des transistors Ω FETs fonctionnels jusqu'à une longueur de grille de 10 nm intégrés avec *high-k* et grille métallique ont été réalisés. Ils montrent de bonnes performances, en particulier en ce qui concerne le contrôle des courants de fuite I_{OFF} .

Les effets particuliers aux architectures verticales ont également été étudiés. Il a été montré que la dispersion de la tension de seuil due aux coins dans les transistors Triple-grille est faible si les canaux sont faiblement dopés. Le couplage latéral permet de plus de supprimer certains effets observés dans les structures planaires complètement désertées. L'influence de la face arrière est ainsi complètement écrantée, conduisant à des dispositifs naturellement très résistants aux radiations. L'effet de couplage du drain (DIVSB) à travers l'oxyde enterré est également très bien contrôlé pour les structures à grilles verticales. Le modèle analytique proposé est à même de décrire l'évolution du potentiel dans l'oxyde enterré pour les structures tridimensionnelles de ce type.

De par leurs structures, les transistors multigrilles non planaires ont des flancs verticaux dont l'orientation cristalline varie suivant la direction cristalline du transport. Outre une chute de la mobilité avec la longueur de grille, il a été montré qu'on ne retrouvait pas les effets prévus en changeant la largeur de l'aile ou la direction cristalline du dispositif. On observe également une amélioration de la mobilité pour les transistors étroits pour les électrons et les trous dans la direction $\langle 110 \rangle$. Par des simulations numériques, il a été montré que la contrainte induite par le procédé pouvait être responsable de cette amélioration de mobilité.

Perspectives :

En ce qui concerne ce travail de thèse, on peut dégager certaines pistes pour des études ultérieures. Il faudrait essayer de déterminer la limite technologique de gravure de l'aileron de silicium (la limite des technologies FinFETs en terme d'*aspect ratio*), et donc également la limite en terme de capacité d'intégration sur la plaque. Ce manuscrit a également montré que le comportement des transistors étroits était sans doute sensible aux contraintes s'appliquant sur le canal. En plus d'une compréhension en profondeur du phénomène, on pourrait comme sur les technologies planaires en contrôlant la contrainte sur le canal disposer d'un moyen élégant d'améliorer les performances. Pour ces applications où le courant circule dans de petites sections de silicium, l'autoéchauffement du canal pourrait aussi conduire à des dégradations notables des performances. Il reste à déterminer si cet autoéchauffement est significatif, en particulier en régime dynamique, et le cas échéant à voir si l'oxyde de grille et l'oxyde enterré constituent des barrières thermiques préoccupantes.

De manière plus générale, on constate que les difficultés pour améliorer les performances des transistors MOS sont de plus en plus importantes, mais que des limites auparavant jugées insurmontables ont été dépassées [Iwai'04]. Il est difficile de prédire actuellement quelles seront les prochaines architectures de transistors (ce qui est plutôt intéressant du point de vue du chercheur).

Vers l'année 2010, on espère l'introduction de dispositifs avec des longueurs de grille inférieures à 20 nm (noeud technologique 'hautes performances' 45 nm) [ITRS'05]. Pour atteindre ce noeud technologique, il faut pour suivre la *roadmap* continuer à améliorer le contrôle électrostatique et les propriétés de transport.

Les technologies simple-grille SOI complètement désertées sont intégrables jusqu'à des longueurs de grille de l'ordre de 20 nm, mais difficilement en dessous. Il faudrait en effet contrôler de très faibles épaisseurs de film (en dessous de 5 nm) avec une très bonne précision. Les problèmes technologiques comme la dispersion de l'épaisseur de silicium, l'épitaxie ou la consommation de silicium seront alors préoccupants.

Les transistors multigrilles ont montré qu'ils pouvaient être les prochaines étapes de la miniaturisation, et qu'ils pouvaient descendre jusqu'à des longueurs de grilles inférieures à 10 nm. Les transistors double grille sont séduisants, mais l'absence de technologies auto-alignées les rends pour l'instant peu pragmatiques. Les transistors de type FinFET sont par contre intrinsèquement auto-alignés, mais leur structure verticale soulève d'autres problèmes. Il a été démontré (y compris dans ce manuscrit) que ces transistors pouvaient être intégrées avec des oxydes haute-permittivité et des grilles métalliques (nécessaire puisque l'utilisation de canaux dopés est très difficile). Il a également été montré que leurs performances en terme de contrôle de la pente sous le seuil et des courants de fuite étaient excellentes grâce à leur structure. Cependant, pour disposer d'une technologie complète le courant de sortie est à améliorer. La surface sur la plaque d'un réseau FinFET doit être réduite, ce qui implique une réduction de la côte lithographique ou l'augmentation de la hauteur des transistors. La

mobilité pour les transistors courts, systématiquement observée plus faible que la mobilité des transistors longs, est également à améliorer. Il faudra ainsi élaborer des parades technologiques pour contrecarrer les dégradations de mobilité observées sur les canaux courts, par exemple en utilisant des canaux contraints. A l'échelle décanométriques, les résistances séries jouent un rôle déterminant dans la dégradation des performances et doivent être optimisées. A ces longueurs de grille, on atteint également des limites physiques. La réduction 'ultime' des dimensions conduit par exemple à un fort effet de courant tunnel source/drain, qui augmente par là même les courants de fuite. De manière générale, en dessous de 5/10 nm de dimension, les effets quantiques et la balisticité deviendront importants. L'évolution de la nanoélectronique à moyen terme nécessitera la prise en compte de ces phénomènes.

On peut penser que l'avenir sera peut être à des structures encore plus agressives, comme les matrices de nanofils. On combinerait alors le contrôle électrostatique avec un fort courant de sortie. Il reste à voir si ces structures ne se heurteront pas à des problèmes de procédé insolubles. Des améliorations importantes peuvent également être faites sur le *design*, quelque soit le type de technologie.

Ce manuscrit ne s'attachant qu'à l'étude d'un dispositif en particulier et sous un certain angle de vue, il n'a sûrement pas la prétention de détenir la clé de l'avenir de l'industrie de la microélectronique. Si l'Histoire nous apprend quelque chose, c'est bien l'humilité : il ne faut pas oublier que le XIX^{ème} siècle considérait que toute la physique avait globalement été comprise, avant que les deux grandes théories du XX^{ème} siècle que sont la mécanique quantique et la relativité bouleversent nos certitudes. La 'loi' de Moore est une loi empirique, il peut être intéressant de regarder s'il n'existe pas d'évolutions analogues pour des industries de haute technologie plus vieilles.

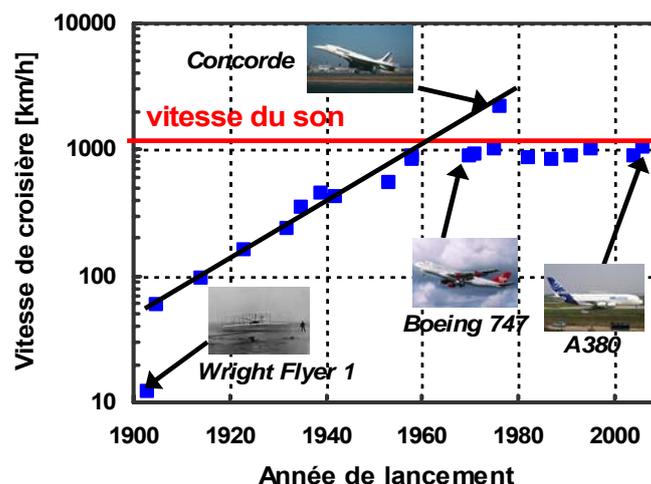


Fig. 144: Courbe de la vitesse de croisière de modèles emblématiques d'avions commerciaux en fonction de leur date de lancement (d'après [Airliners]).

A ce titre, une comparaison de l'évolution de l'avionique civile est frappante. Si on trace la courbe d'évolution de la vitesse de croisière des avions commerciaux en fonction de leur

année de lancement (Fig. 144), on constate une loi de type exponentielle (comme la 'loi' de Moore) très régulière suivie d'un palier marqué à partir des années 1970. Il existe à l'heure actuelle des produits très performants comme le supersonique *Concorde*, mais les coûts de fabrication et d'exploitation de celui-ci sont trop élevés pour qu'il devienne prédominant sur le marché. C'est peut-être là que se trouve l'avenir de la microélectronique : des produits grand public où la vitesse de commutation des transistors n'évolue quasiment plus. Les améliorations se feront alors sur le *design* et plus généralement sur l'architecture des puces. Des applications à très grande vitesse pourront être développées, mais leur coût prohibitif les cantonnera dans des applications spécifiques où ils seront indispensables.

[Airliners] <http://www.airliners.net>, site généraliste sur l'aéronautique.

[ITRS'05] *International Technology Roadmap for Semiconductors (ITRS)*, édition 2005, "Process Integration Devices, and Structures", "High Performance Logic Technology Requirements", p. 11, <http://www.itrs.net/Links/2005ITRS/PIDS2005.pdf>, 2005.

[Iwai'04] H. Iwai, "CMOS Technology Future", proceedings of the 5th *International Caracas Conference on Devices, Circuits, and Systems*, pp. 179-182, 2004.

Bibliographie de l'auteur

Articles:

- ✓ **R. Ritzenthaler**, S. Cristoloveanu, O. Faynot, C. Jahan, A. Kuriyama, L. Brevard, and S. Deleonibus, "Lateral Coupling and Immunity to Substrate Effect in Ω FET Devices", *Solid-State Electronics*, vol. 50, no. 4, pp. 558-565, 2006.
- ✓ S. Cristoloveanu, **R. Ritzenthaler**, A. Ohata, O. Faynot, "3D Size Effects in Advanced SOI Devices", *International Journal of High Speed Electronics and Systems*, vol. 16, no. 1, pp. 9-30, 2006.
- ✓ T. Ernst, **R. Ritzenthaler**, O. Faynot, S. Cristoloveanu, "A Model of Fringing Fields in Short-Channel Planar and Triple-Gate SOI MOSFETs", accepté à *Transactions on electron devices*.

Conférences internationales:

- ✓ O. Faynot, G. Barna, **R. Ritzenthaler**, P. Gidon, "Scaling of Multiple-Gate Fully Depleted SOI, Transistors", proceedings of *Int. Conf. on Solid State Devices and Materials (SSDM)*, pp. 764-765, 2004.
- ✓ S. Cristoloveanu, **R. Ritzenthaler**, A. Ohata, O. Faynot, "3D Size Effects in Advanced SOI Devices", 4th *Advanced Workshop on 'Frontiers in Electronics' (WOFE)*, Aruba, December 17th - December 22nd, 2004. **(papier invité)**
- ✓ **R. Ritzenthaler**, O. Faynot, C. Jahan, S. Cristoloveanu, "Quantum Induced Suppression of Corner Effects in Double, Triple and Quadruple-Gate SOI Transistors", *EUROSOI Network 1st Workshop on Silicon-On-Insulator Technology Devices and Circuits*, pp. 75-76, 2005.
- ✓ **R. Ritzenthaler**, O. Faynot, C. Jahan, S. Cristoloveanu, "Corner and Coupling Effects in Multiple-Gate FETs", *207th meeting of the Electrochemical Society, Silicon-on-Insulator technology and devices session XII*, pp. 283-288, 2005.
- ✓ C. Jahan, O. Faynot, M. Cassé, **R. Ritzenthaler**, L. Brévard, L. Tosti, X. Garros, C. Vizioz, F. Allain, A.M. Papon, H. Dansas, F. Martin, M. Vinet, B. Guillaumot, A. Toffoli, B. Giffard, S. Deleonibus, "10nm Ω FETs transistors with TiN metal gate and HfO₂", *VLSI Tech. Dig.*, pp. 112-113, 2005.
- ✓ O. Weber, P. Scheiblin, **R. Ritzenthaler**, T. Ernst, F. Andrieu, F. Ducroquet, J.-F. Damlencourt, Y. Le Tiec, A.-M. Papon, H. Dansas, L. Brévard, A. Toffoli, B.

- Guillaumot, S. Deleonibus, "A Novel Locally Engineered (111) V-channel pMOSFET Architecture with Improved Drivability Characteristics for Low-Standby power (LSTP) CMOS Applications", *VLSI Tech. Dig.*, pp. 156-157, 2005.
- ✓ **R. Ritzenthaler**, O. Faynot, T. Poiroux, C. Jahan, S. Cristoloveanu, "Position-Dependent Threshold in FinFETs and Triple-gate FETs", *EUROSOI Network 2nd Workshop on Silicon-On-Insulator Technology Devices and Circuits*, 2006.
 - ✓ **R. Ritzenthaler**, O. Faynot, C. Jahan, A. Kuriyama, L. Brevard, S. Deleonibus, S. Cristoloveanu, "Immunity to Substrate Effect in Advanced Ω FET Devices", proceedings of the *35th European Solid-State Device Research conference (ESSDERC)*, pp. 81-84, 2005.
 - ✓ T. Ernst, C. Dupré, C. Isheden, E. Bernard, **R. Ritzenthaler**, V. Maffini-Alvaro, J. Cluzel, A. Toffoli, C. Vizioz, S. Borel, F. Andrieu, F. De Crecy, V. Delaye, D. Lafond, G. Rabillé, J.-M. Hartmann, M. Rivoire, B. Guillaumot, A. Suhm, P. Rivallin, O. Faynot, G. Ghibaudo, S. Deleonibus, "Novel 3D integration process for highly scalable Nano-Beam stacked-channels GAA (NBG) CMOSFETs with HfO₂/TiN gate stack", *IEDM'06 Technical Digest*, 2006. (accepté)
 - ✓ **R. Ritzenthaler**, C. Dupré, X. Mescot, O. Faynot, T. Ernst, J.-C. Barbé, C. Jahan, L. Brévard, F. Andrieu, S. Deleonibus, S. Cristoloveanu, "Mobility behavior in narrow Ω -gateFETs devices", *IEEE Int. SOI conf.*, 2006. (accepté)
 - ✓ A. Ohata, **R. Ritzenthaler**, O. Faynot, S. Cristoloveanu, "Special Size Effects in Advanced SOI Transistors", *7th Symposium Diagnostics & Yield in Advanced Silicon Devices and Technologies for ULSI Era*, Warsaw, June 26th - June 28th, 2006. (**papier invité**)
 - ✓ **R. Ritzenthaler**, M. Gaillardin, K. Akarvardar, O. Faynot, C. Jahan, S. Cristoloveanu, "Modeling the Back-Gate Coupling Effect in Triple-, Pi- and Omega-GateFETs", accepté au *211th meeting of the Electrochemical Society, Silicon-on-Insulator technology and devices session XIII*, 2007.

TITRE: Architectures avancées des transistors FinFETs – réalisation, modélisation et caractérisation électrique.

RESUME:

La réduction des dimensions des transistors pour les prochains nœuds technologiques soulève des problèmes technologiques de plus en plus difficiles à surmonter. De nouvelles architectures dites 'multigrilles' sont proposées afin de pouvoir poursuivre la miniaturisation. Parmi ces dispositifs, les transistors verticaux FinFETs sur SOI (Silicium-sur-isolant) sont de prometteurs candidats.

Dans cette étude de thèse, des transistors FinFETs ont été fabriqués jusqu'à une longueur de grille de 10 nm. Les transistors montrent d'excellentes performances, en particulier en ce qui concerne le contrôle des effets de canaux courts. Nous proposons également une étude des effets électrostatiques spécifiques intervenant dans ce type de dispositifs. Les effets de coins et les couplages de la face arrière et du drain à travers l'oxyde enterré sont ainsi mis en lumière. Il est montré que les effets de coins sont négligeables si le dopage est faible et que les couplages à travers l'oxyde enterré sont écartés pour des dispositifs étroits. Ceci rend les structures FinFETs attractives pour la miniaturisation des circuits intégrés MOS.

Les propriétés de transport dans les canaux verticaux et horizontaux sont également mises en lumière. Nous montrons et discutons l'effet du changement de plan et d'orientation cristalline ainsi que l'impact des contraintes mécaniques.

SPECIALITE:

Micro- et Nano-électronique

MOTS-CLES:

Microélectronique, transistor à effet de champ (MOSFET), silicium sur isolant (SOI), transistors multigrilles, FinFET, Ω FET, effets de canal court, couplage électrique, mobilité

TITLE: Advanced architectures for FinFETs transistors – processing, device modelling and characterization.

ABSTRACT:

The dimensions downscaling for the next nodes of the microelectronics industry is handicapped by technological problems more and more difficult to overcome. Multigate MOS architectures have been proposed to continue further the downscaling. Among them, vertical FinFET structures on SOI (Silicon-on-insulator) are promising candidates.

During this PhD, FinFET transistors with gate length down to 10 nm were processed. These transistors exhibit excellent performance, especially in term of short-channel effects control. A detailed investigation of the specific electrostatic effects in FinFET-like devices was performed. It is shown that the corner effect is small if the body remains undoped and that the coupling from the back-gate and drain through the buried oxide is screened for narrow devices. This makes the FinFETs attractive for applications in CMOS advanced integrated circuits.

Transport properties in the vertical channels were also investigated. The effect of crystallographic axis and surfaces is discussed, as well as the impact of the mechanical strain.

KEY-WORDS:

Microelectronics, field effect transistor (MOSFET), silicon on insulator (SOI), multiple-gate transistors, FinFET, Ω FET, short-channel effects, electrical coupling, mobility

- **LETI:** Laboratoire d'Electronique et des Technologies de l'Information, CEA Grenoble, 17 avenue des martyrs, 38054 Grenoble Cedex 9, FRANCE.
- **IMEP:** Institut de Microélectronique, Electromagnétisme et Photonique (IMEP), MINATEC, 3 Parvis Louis Néel, Avenue Félix Esclangon, 38016 Grenoble Cedex 1, FRANCE.