



HAL
open science

Etude et modélisation compacte du transistor FinFET ultime

Nicolas Chevillon

► **To cite this version:**

Nicolas Chevillon. Etude et modélisation compacte du transistor FinFET ultime. Micro et nanotechnologies/Microélectronique. Université de Strasbourg, 2012. Français. NNT : 2012STRAD016 . tel-00750928

HAL Id: tel-00750928

<https://theses.hal.science/tel-00750928>

Submitted on 19 Nov 2012

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.

Ecole Doctorale Mathématiques, Sciences de l'Information
et de l'Ingénieur

UdS

THÈSE

présentée pour obtenir le grade de

Docteur de l'Université de Strasbourg

Discipline : Sciences de l'ingénieur

Spécialité : Microélectronique

par

Nicolas Chevillon

**Etude et modélisation compacte du transistor
FinFET ultime**

Soutenue publiquement le 13 juillet 2012

Membres du jury :

Directeur de thèse : M. Christophe LALLEMENT, Professeur, InESS, Strasbourg

Rapporteur : Mme Daniela MUNTEANU, Chargé de recherche, IM2NP, Marseille

Rapporteur : M. Jean-Pierre RASKIN, Professeur,
Université Catholique de Louvain (UCL), Belgique

Examineur : M. Daniel MATHIOT, Professeur, InESS, Strasbourg

Examineur : M. Jean-Michel SALLESE, Maître d'enseignement et de recherche,
École Polytechnique Fédérale de Lausanne (EPFL), Suisse

Examineur : M. Fabien PRÉGALDINY, Maître de conférences, InESS, Strasbourg

Remerciements

Mes premiers remerciements vont tout naturellement à mon directeur de thèse, Christophe Lallement pour m'avoir donné l'occasion de réaliser cette thèse. Je le remercie vivement pour sa patience, ses conseils et son soutien qui ont contribué au bon déroulement de cette thèse. Un très grand merci pour les corrections et l'aide apportées à la rédaction de ce manuscrit.

Je tiens à remercier mes encadrants de thèse. Merci à Fabien Prégaldiny pour son expertise en physique du solide et en modélisation des composants, ainsi que pour les corrections apportées à ce manuscrit. Merci à Morgan Madec pour son expertise dans la simulation de circuits et pour m'avoir fait découvrir les joies de l'enseignement.

Je remercie profondément Jean-Michel Sallese, de m'avoir fait profiter de sa connaissance de la physique solide. Merci pour toutes nos discussions enthousiasmantes et pour l'apport de ses idées.

Je veux remercier Ashkhen Yesayan pour sa contribution précieuse dans la modélisation du FinFET ultime, et notamment dans la modélisation des effets canaux courts par le modèle de potentiel minimum.

Je tiens également à remercier Yann Leroy pour nos discussions sur la physique des composants et son support dans la rédaction de cette thèse.

Je remercie également Nicolas Collin pour le support informatique indispensable qu'il a apporté tout au long de cette thèse.

Je remercie le directeur du laboratoire, Daniel Mathiot, de m'avoir accueilli au sein du laboratoire InESS et d'avoir accepté de présider le jury.

Je remercie aussi très fortement Daniela Munteanu et Jean-Pierre Raskin pour avoir accepté d'être les rapporteurs de ma thèse.

Je tiens à exprimer ma reconnaissance à mes collègues du laboratoire InESS pour les discussions très utiles tout au long de ce travail. Merci à Anne-Sophie Cordan, Jean-Baptiste Kammerer, Jacques Michel, Norbert Dumas. Merci à Dumitru Armeanu, Qing Sun, Mingchun Tang, Yves Gendrault, Adam Raba, Jérôme Heitz, Cyril Kern et Ahmed Quadri.

Enfin, un grand merci à ma famille et à mes amis pour leur soutien, et sans qui je ne serais pas arrivé jusque-là.

Table des matières

Notations	ix
Introduction générale	1
1 La technologie multigrille avancée	5
1.1 L'évolution de la technologie des transistors MOSFET	5
1.1.1 Limitations de la réduction d'échelle des MOSFET	6
1.1.2 Vers de nouvelles architectures de transistors MOSFET	7
1.1.3 Les dispositifs multigrilles	8
1.2 Le premier processeur à base de transistors multigrilles : le « Ivy Bridge » d'Intel	10
Conclusion	11
Références	12
2 Modélisation compacte du transistor MOSFET double grille canal long	13
2.1 État de l'art des modèles compacts du DG MOSFET	14
2.1.1 Modèle de Yuan Taur	14
2.1.2 Modèle PSP FinFET	16
2.1.3 Modèle de Benjamin Iñiguez	18
2.2 Le modèle « InESS-EPFL » du DG MOSFET non dopé	20
2.2.1 Structure du dispositif modélisé	20
2.2.2 Corps du modèle	21
2.2.3 Aspects en température	24
2.3 Le modèle du DG MOSFET dopé	25
2.3.1 État de l'art	25
2.3.2 Développement du modèle basé sur un concept de largeur équivalente de silicium	26
2.3.3 Validation du modèle et discussion	32
2.3.4 Conclusion de la modélisation du DG MOSFET dopé	39
Conclusion	39
Références	39
3 Modélisation compacte du transistor FinFET ultime	43
3.1 Introduction	44
3.2 Le modèle FinFET ultime	46
3.2.1 Effets canaux courts	46
3.2.2 Dégradation de la mobilité	57
3.2.3 Région de saturation	58
3.2.4 Effets de mécanique quantique	59
3.2.5 Modèle de transcapacités	62
3.2.6 Discussion et résultats	66
Conclusion	70

Références	71
4 Vers le modèle « universel » de MOSFET multigrille	73
4.1 Etat de l'art	74
4.2 Modélisation de transistors multigrilles selon un concept de paramètres géométriques équivalents	75
4.2.1 Définition de la largeur équivalente pour les transistors à effets de champs de section de grille rectangulaire	75
4.2.2 Généralisation de la largeur équivalente à toutes géométries à grille enrobante	80
4.2.3 Un cas d'étude : le GAA MOSFET à grille cylindrique	81
4.2.4 Extension de la modélisation à des transistors MOSFET à grille partiellement enrobante	84
4.2.5 Le cas du MOSFET bulk triple grille	90
4.3 Modélisation de l'effet de la température	93
4.4 Paramètres du modèle multigrille	94
4.5 Validation	94
4.5.1 Validation TCAD	95
4.5.2 Validation par rapport aux mesures expérimentales	98
Conclusion	100
Références	100
5 Environnement de travail logiciel	103
5.1 Simulation numérique de dispositifs MOSFET multigrilles	104
5.1.1 Méthode des éléments finis	104
5.1.2 Modèles et méthodes utilisés	105
5.1.3 Constitution du maillage	107
5.2 Extraction de paramètres du modèle FinFET	109
5.2.1 Caractéristiques électriques pour l'extraction de paramètres	109
5.2.2 Méthodologie de la procédure d'extraction de paramètres de modèle	111
5.2.3 Procédure complète automatisée d'extraction de paramètres du modèle FinFET	112
5.3 Environnement de travail	113
5.3.1 Plate-forme de simulation numérique et d'extraction des paramètres du modèle	113
5.3.2 Outils pour l'étude du comportement physique des transistors.	116
5.3.3 Méthodes d'extraction de paramètres électriques	116
Conclusion	117
Références	117
6 Simulation de circuits numériques	119
6.1 Simulations niveau-transistor de circuits numériques	120
6.1.1 Généralités	120
6.1.2 Modèle compact du transistor FinFET pour un simulateur de circuits	121
6.1.3 Simulation de portes logiques	124
6.2 Modélisation niveau-porte	126
6.2.1 Principe	126
6.2.2 Procédure de paramétrage du modèle niveau-porte	128
6.2.3 Méthode d'extraction des temps de propagation	130
6.2.4 Résultats d'extraction	131
6.3 Simulation de l'influence de variations de process	132

6.3.1	Etude de la porte NOT	132
6.3.2	Etude de circuits CMOS complexes	133
	Conclusion	135
	Références	136
7	Modélisation du transistor MOSFET double grille sans jonction	137
7.1	Etat de l'art	138
7.2	Étude électrostatique dans les DG MOSFET sans jonction	139
7.2.1	Limitations du modèle de DG MOSFET à base de jonctions pour simuler les dispositifs sans jonction	139
7.2.2	Solution approchée de l'équation de Poisson	141
7.2.3	Approche générale du calcul de la charge mobile	143
7.2.4	Expressions approchées pour les charges	144
7.2.5	Tension de seuil dans les transistors sans jonction	149
7.3	Développement de l'expression de courant	151
7.3.1	Courant en accumulation	152
7.3.2	Courant en déplétion	152
7.3.3	Traitement général du courant dans le DG MOSFET sans jonction	152
7.3.4	Comparaison avec des simulations numériques	154
	Conclusion	157
	Références	158
	Conclusion et perspectives	159
	Annexes	161
A	Simulation de circuits numériques	161
A.1	Structure du code Verilog-A du modèle compact du transistor FinFET	161
A.2	Code spectre de la simulation d'une porte NOT	164
	Bibliographie personnelle	167

Notations

Constantes fondamentales

Désignation	Nom et valeur
Charge électrique élémentaire	$e = 1.602 \cdot 10^{-19} \text{ C}$
Permittivité diélectrique du vide	$\varepsilon_0 = 8.854 \cdot 10^{-12} \text{ F.m}^{-1}$
Permittivité diélectrique du silicium	$\varepsilon_{si} = 11.8 \cdot \varepsilon_0$
Permittivité diélectrique du dioxyde de silicium	$\varepsilon_{ox} = 3.9 \cdot \varepsilon_0$
Constante de Boltzmann	$k_b = 1.380658 \cdot 10^{-23} \text{ J.K}^{-1}$
Gap du silicium	$E_{G_{si}} = 1.08 \text{ eV}$
Constante de Planck	$h = 6.63 \cdot 10^{-34} \text{ J.s}$

Notations

Désignation	Nom	Unité
Concentration intrinsèque du silicium	n_i	m^{-3}
Température du semi-conducteur	T	K
Tension thermodynamique	$U_T = (k_b T)/e$	V
Longueur de grille	L	m
Longueur équivalente	L_{eq}	m
Largeur du film de silicium	W_{Si}	m
Hauteur du film de silicium	H_{Si}	m
Largeur équivalente de film de silicium non dopé	W_{eq}	m
Hauteur équivalente de film de silicium non dopé	H_{eq}	m
Épaisseur nominale d'oxyde grille	t_{ox}	m
Épaisseur équivalente d'oxyde grille	$t_{ox,eq}$	m
Épaisseur d'oxyde de la grille supérieure	$t_{ox,top}$	m
Rayon du film cylindrique de silicium	R_{Si}	m
Capacité nominale d'oxyde grille	C_{ox}	F.m^{-2}
Capacité équivalente d'oxyde grille	$C_{ox,eq}$	F.m^{-2}
Capacité du film de silicium	C_{Si}	F.m^{-2}
Mobilité totale	μ	$\text{m}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$
Mobilité à champ faible	μ_0	$\text{m}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$
Mobilité effective transverse	μ_{\perp}	$\text{m}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$

Tension de seuil canal long	V_{to}	V
Longueur d'échelle générale	λ_G	m
Facteur de forme	α	-
Niveau de dopage du canal dopé P	N_a	m^{-3}
Niveau de dopage du canal dopé N	N_D	m^{-3}
Dopage des régions source-drain	$N_{S/D}$	m^{-3}
Potentiel des jonctions source-canal et drain-canal	V_{bi}	V
Potentiel de Fermi	Φ_f	V
Potentiel de quasi-Fermi	V_{ch}	V
Potentiel électrostatique	Ψ	V
Potentiel électrostatique au centre du canal	Ψ_0	V
Différence de travaux de sortie	$\Delta\Phi$	V
Polarisation de grille	V_g	V
Polarisation de drain	V_d	V
Polarisation de source	V_s	V
Tension source-grille	V_{GS}	V
Tension source-drain	V_{DS}	V
Tension de grille corrigée des SCE	V_{gN}	V
Tension de grille corrigée des SCE et des QME	V_{gN}^{QM}	V
Courant de drain	I_D	A
Courant spécifique	I_{SP}	A
Courant de drain en état « on » / « off »	$I_{on/off}$	A
Densité de charges d'inversion	Q_i	C.m^{-2}
Densité intrinsèque de charges mobiles	q_{int}	C.m^{-2}
Densité de charges du canal	Q_m	C.m^{-2}
Densité de charges de grille	Q_g	C.m^{-2}
Densité de charges à la source	Q_s	C.m^{-2}
Densité de charges au drain	Q_d	C.m^{-2}
Charge totale de grille	Q_G	C
Charge totale d'inversion à la source	Q_S	C
Charge totale d'inversion au drain	Q_D	C
Facteur de pente	n	-
Transconductance	g_m	Ω^{-1}
Conductance de sortie	g_{ds}	Ω^{-1}
Longueur de diffusion des porteurs majoritaires	L_{MCD}	m
Résistance intrinsèque du canal	R_{ch}	Ω
Champ électrique transverse en forte inversion	$E_{\perp SI}$	V.m^{-1}
Champ électrique transverse induit par les effets canaux courts	$E_{\perp WI}$	V.m^{-1}
Vitesse de saturation des porteurs	v_{sat}	m.s^{-1}

Tension de saturation de drain	V_{Dsat}	V
Tension de drain effective	V_{Deff}	V
Variation du potentiel minimum	$\Delta\Psi_{Smin}$	V
Longueur équivalente du canal	L_{eq}	m
Longueur de la région de saturation	ΔL	m
Temps de propagation bas-vers-haut	tp_{LH}	s
Temps de propagation haut-vers-bas	tp_{HL}	s
Paramètre de mobilité du modèle FinFET	e_1	V/m
Paramètre de mobilité du modèle FinFET	e_2	$(V/m)^2$
Paramètre de la tension de saturation du modèle FinFET/- MuGFET	θ	-
Longueur de recouvrement (paramètre du modèle FinFET)	L_{ov}	m
Paramètre de mobilité du modèle MuGFET	E_0	V/m
Paramètre de mobilité du modèle MuGFET	E_1	$(V/m)^2$

Grandeurs électriques normalisées

Désignation	Nom
Densité de charges de grille normalisée	q_g
Densité de charges du canal normalisée	q_m
Densité de charges du canal à la source normalisée	q_{ms}
Densité de charges du canal au drain normalisée	q_{md}
Charge totale de grille normalisée	q_G
Charge totale de drain normalisée	q_D
Charge totale de source normalisée	q_S
Courant de drain normalisé	i
Tension de seuil canal long normalisée	v_{to}
Potentiel aux jonctions source/drain-canal normalisé	v_{bi}
Quasi-potentiel de Fermi normalisé	v_{ch}
Polarisation de drain normalisée	v_d
Polarisation de grille normalisée	v_g
Polarisation de source normalisée	v_s
Tension de grille normalisée corrigée des SCE	v_{gN}
Tension de grille normalisée corrigée des SCE et des QME	v_{gN}^{QM}
Tension de saturation de drain normalisée	v_{dsat}
Tension de drain effective normalisée	v_{def}
Potentiel de surface normalisé au niveau du drain	ψ_d
Potentiel de surface normalisé au niveau du point de saturation	ψ_{sat}
Variation normalisée du potentiel minimum	$\Delta\psi_{smin}$

Abréviations et acronymes

Nom	Signification	Traduction française
BOX	Buried OXide	Oxyde enterré
CLM	Channel length modulation	Modulation de longueur de canal
CMOS	Complementary MOS	Technologie MOS complémentaire
DG	Double Gate	Double grille
DIBL	Drain Induced Barrier Lowering	Abaissement de la barrière induit par le drain
FET	Field Effect Transistor	Transistor à effet de champ
FinFET	Fin Field Effect Transistor	Transistor à effet de champ en forme d'“aileron”
GAA	Gate-All-Around	Grille totalement enrobante
GIDL	Gate Induced Barrier Lowering	Abaissement de la barrière induit par la grille
ITRS	International Technology Roadmap for Semiconductors	“Feuille de route” internationale technologique pour les semi-conducteurs
MOS	Metal-Oxide-Semiconductor	Métal-Oxyde-Semi-conducteur
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor	Transistor à effet de champ Métal-Oxyde-Semi-conducteur
P(F)D-SOI	Partially (Fully) Depleted Silicon On Insulator	Silicium sur isolant partiellement (complètement) déplété
QG	Quadruple Gate	Quadruple grille
QME	Quantum Mechanical Effects	Effets de mécanique quantique
SCE	Short Channel Effect	Effet canal court (partage de charge)
SCEs	Short Channel Effects	Ensemble des effets canaux courts
SOI	Silicon On Insulator	Silicium sur isolant
SS	Subthreshold Slope	Pente sous le seuil
TAT	Trap-Assisted Tunneling	Effet tunnel assisté par les pièges
TCAD	Technology Computer Aided Design	Conception technologique assistée par ordinateur
TG	Triple Gate	Triple grille
ULSI	Ultra Large Scale Integration	Intégration à ultra grande échelle
VLSI	Very Large Scale Integration	Intégration à très grande échelle

Introduction générale

Depuis plus de 40 ans l'industrie de la microélectronique produit des puces de silicium qui réalisent des tâches de calculs et plus généralement des fonctions logiques. En effet, c'est en 1967 que Frank Wanlass invente le concept de la technologie CMOS qui exploite le transistor à effet de champ MOS pour concevoir des circuits de traitements de signaux numériques tels que des processeurs. Une course à la performance de ces puces a mené à une miniaturisation toujours plus importante des transistors MOS. Cela a permis la conception de circuits intégrés à très large et ultra-large échelle (VLSI-ULSI) comportant toujours plus de transistors. Ainsi, au fil des années, la complexité des tâches de traitements implémentées par des circuits intégrés sur une même surface de silicium a augmenté. De plus, la réduction des dimensions du transistor MOS augmente ses performances électriques.

En 1975, Gordon E. Moore quantifiait l'évolution de la densité des circuits intégrés. Il constata que le nombre de transistors par puce d'une même taille doublait tous les 18 mois. C'est ce que nous appelons la loi de Moore. Il ne s'agit en aucun cas d'une loi physique, mais d'un état de fait qui a ensuite poussé les industriels de la microélectronique à garder un rythme constant de l'amélioration des puces d'une génération à l'autre. Ce rythme n'est pas toujours resté constant, ainsi, depuis 1995 le constat du temps nécessaire au doublement du nombre de transistors est de deux ans. Au début des années 90, une organisation internationale, l'ITRS [1], est créée par des industriels et des universitaires pour planifier le futur de l'industrie de la microélectronique. L'évolution technologique de chaque élément du transistor est précisée année par année. Un plan à suivre est ainsi constitué pour coordonner les efforts d'améliorations techniques permettant la réduction d'échelle du transistor MOS.

Le transistor MOS bulk planaire est le transistor originel et conventionnel de la microélectronique. L'ITRS prévoit la fin de la miniaturisation du transistor MOS conventionnel en 2017. Des courants parasites apparaissent et augmentent avec la réduction de la longueur de grille. Cette dégradation des performances électriques provient principalement de la perte de contrôle électrostatique dans le transistor. Des solutions technologiques sont mises en oeuvre afin de réduire ces courants parasites. Malgré les diverses solutions proposées, il est de plus en plus difficile d'optimiser les procédés de fabrication des architectures classiques des transistors MOS. Ainsi, des solutions alternatives au MOSFET conventionnel sont développées pour les futurs noeuds technologiques. Il s'agit du transistor MOS SOI planaire *fully depleted* (FD) dont

la réduction peut s'étendre jusqu'en 2020 et des transistors MOS à multiple grilles (MuG) pour lesquels les prévisions de l'ITRS sont déterminé jusqu'en 2026.

L'émergence de nouvelles architectures telles que les transistors MOS à multiple grilles, comprenant les transistors double grille, FinFET ou à grille enrobante dits GAA (Gate-All-Around), permettent d'améliorer le contrôle électrostatique du dispositif et limitent ainsi les effets parasites généralement observés. Le transistor FinFET a une bonne compatibilité avec les chaînes de fabrication des transistors conventionnels et semble leur plus prometteur successeur. Il se pose alors naturellement le problème de la conception de circuits intégrant ces nouveaux dispositifs. Ainsi, le développement de modèles compacts performants, notamment pour le FinFET, devient une demande pressante. A la différence d'un modèle purement physique, un modèle compact utilise -idéalement- des expressions explicites (i.e. sans itération) pour le calcul des différentes grandeurs physiques telles que le courant de drain, les densités de charges, les transcapacités, etc. L'intérêt d'un modèle compact est de donner au concepteur un résultat de simulation le plus rapidement possible tout en conservant une précision suffisante à la bonne fonctionnalité du circuit simulé. En outre, un modèle compact se doit d'être lisible pour le concepteur, et le fait pour ce dernier de pouvoir effectuer un prédimensionnement à la main (avec le modèle) de son design est un réel atout. Depuis quelques années, plusieurs modèles compacts de MOSFET double grille et de FinFET ont été publiés. Cependant, actuellement aucun d'entre eux ne répond complètement à toutes les attentes des concepteurs, ainsi la réalisation d'un modèle complet de transistors à multiple grilles reste donc encore à faire.

Dans le Chapitre 1, nous allons présenter en détails l'évolution technologique du transistor MOS et introduire ensuite les avantages de la technologie MOSFET à multiple grilles dans la poursuite de la miniaturisation des dispositifs MOS.

Ensuite, dans le Chapitre 2, nous dresserons un état de l'art des modèles compacts de MOSFET double grille à canal faiblement dopé, suivi de la présentation de notre propre modèle compact de ce dispositif. Nous poursuivrons par le développement d'un nouveau modèle compact du MOSFET double grille à canal dopé élaborée comme une extension du modèle faiblement dopé.

Le modèle compact explicite du FinFET ultime sera détaillé dans le Chapitre 3. Nous rappellerons tout d'abord les travaux précédents effectués dans notre groupe à l'InESS sur le modèle compact du FinFET. Puis, nous développerons les avancées de notre modélisation statique et dynamique de transistors FinFET de petites géométries. Nous allons présenter l'expression du courant de drain en prenant en compte la dégradation de la pente sous le seuil, l'effet SCE, l'effet DIBL, la saturation de la vitesse des porteurs, la modulation de la longueur du canal et les effets quantiques. Nous détaillerons aussi notre modèle dynamique en présentant les expressions des transcapacités.

Le chapitre 4 est dédié à la modélisation compacte de transistors à multiple grilles de grandes géométries. En se basant sur le modèle canal long du transistor MOSFET double grille faiblement dopé, nous avons élaboré un modèle compact générique valide pour une grande variété d'architectures de dispositifs à multiple grilles. Une validation sur des mesures expérimentales sera présentée.

Dans le Chapitre 5, nous présenterons l'environnement logiciel complet établi au cours de la thèse pour permettre et faciliter notre travail de recherche. Nous décrirons ainsi l'outil de simulation numérique indispensable à nos travaux. Nous montrerons l'environnement de mesures et d'extractions de paramètres bâti autour de l'outil logiciel de caractérisation IC-CAP. Celui-ci permet notamment la caractérisation du FinFET, et l'extraction puis l'optimisation des paramètres du modèle. Enfin nous exposerons le développement d'un ensemble de programmes, que nous avons développés en langage Python, nous permettant le traitement des différentes données que nous manipulons et facilitant notre travail de recherche.

Le transistor FinFET est principalement dédié à la conception de circuits numériques de technologie CMOS. L'objectif d'établir le modèle compact de ce dispositif est ainsi de permettre la simulation de circuits numériques basés sur le FinFET. Dans le chapitre 6, nous traiterons de la simulation de circuits numériques simples en présentant tout d'abord l'implémentation du modèle compact dans un langage de description matériel, le langage Verilog-A. La simulation de circuits numériques d'un grand nombre de transistors requiert un long temps de calcul malgré l'emploi de modèles compacts. Il est alors nécessaire de décomposer de tels circuits en fonctions logiques plus élémentaires sur différents niveaux d'abstraction. Chaque sous-circuit ainsi défini est associé à un modèle niveau-porte qui tient uniquement compte de la fonction logique du sous-circuit et de ses temps de propagation caractérisant ses performances temporelles. Nous avons développé le modèle niveau-porte de portes logiques à partir de la simulation de ces portes en utilisant le modèle compact, puis comparé les résultats entre ces deux niveaux de simulations pour des circuits plus complexes. Une étude statistique des paramètres du modèle niveau-porte est conduite en faisant varier des paramètres géométriques du modèle compact du transistor FinFET.

Enfin, le chapitre 7 présentera la modélisation compacte d'un transistor d'une architecture originale comparé au transistor MOS conventionnel. Il s'agit du transistor MOSFET double grille sans jonction. Il a été démontré expérimentalement qu'une structure de MOSFET dont le niveau de dopage est uniforme et fort de la source au drain, peut fonctionner en commutation. Le régime de fonctionnement du canal est en déplétion et/ou accumulation. Dans le cas des transistors traditionnels, la difficulté de réaliser des profils de dopant des jonctions toujours plus abrupts est un frein à la miniaturisation. Pour les transistors sans jonction, le problème disparaît puisque le niveau de dopage est uniforme.

Le manuscrit se terminera par une conclusion générale des travaux effectués au cours de cette thèse, et par un passage en revue des perspectives de nos travaux de recherche sur la modélisation compacte des transistors à multiple grilles.

Références

- [1] International Technology Roadmap for Semiconductors (ITRS), rap. tech., 2011. adresse : <http://public.itrs.net/>.

Chapitre 1

La technologie multigrille avancée

Sommaire

1.1	L'évolution de la technologie des transistors MOSFET	5
1.1.1	Limitations de la réduction d'échelle des MOSFET	6
1.1.2	Vers de nouvelles architectures de transistors MOSFET	7
1.1.3	Les dispositifs multigrilles	8
1.2	Le premier processeur à base de transistors multigrilles : le « Ivy Bridge » d'Intel	10
	Conclusion	11
	Références	12

1.1 L'évolution de la technologie des transistors MOS-FET

Quand en 1962, les ingénieurs Steven Hofstein et Frederic Heiman inventent le transistor Metal-Oxide-Semi-conducteur à effet de champ (MOSFET), l'industrie de la microélectronique connaît une avancée technologique importante qui lui permet un fort développement qui ne s'est jamais démenti depuis. Le transistor MOSFET est l'élément de base de toutes les puces électroniques constituant nos ordinateurs, téléphone portable, etc . Ces puces sont des circuits intégrés de technologie CMOS. Cette technologie a permis de concevoir sur une seule puce plusieurs fonctions électroniques réalisées à partir de milliers de transistors : c'est le début de la conception de circuits intégrés à très grande échelle et de haute densité appelée communément VLSI (*Very-Large-Scale Integration*).

Ces composants semi-conducteurs, en référence aux propriétés du silicium utilisé pour leur fabrication, voient ainsi leurs dimensions régulièrement réduites. En plus de l'augmentation de la densité de transistors par puce, cela permet d'améliorer leur performances et donc celles des puces. En 1965, Gordon E. Moore, un des co-fondateurs d'Intel, constate pour la première fois une tendance sur le nombre de transistors des premiers circuits intégrés développés. Selon ces observations, le nombre de transistors intégrés dans une puce doublerait tous les ans. Cette

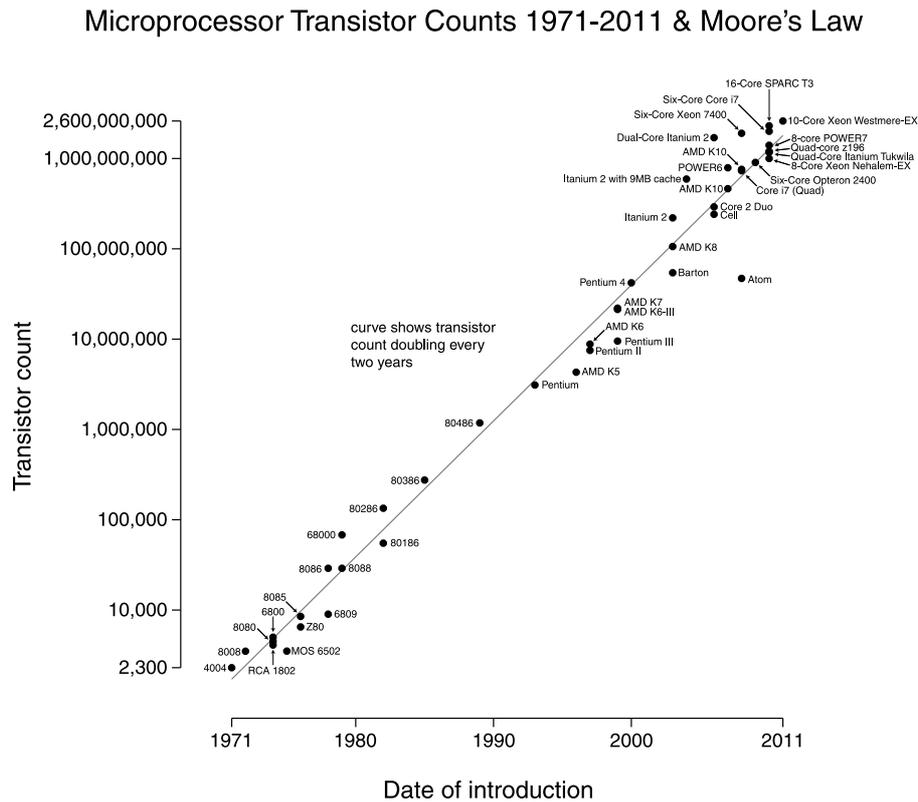


Figure 1.1 – *Loi de Moore. Tracé du nombre de transistor par CPU selon les dates de mises sur le marché.*

tendance qui est nommée alors la loi de Moore, se transforme en un but à atteindre par les industriels de la microélectronique. Elle est révisée à la baisse par Gordon E. Moore en 1975 et en 1995 avec désormais un doublement tous les deux ans. La loi de Moore est représentée à la Figure 1.1.

1.1.1 Limitations de la réduction d'échelle des MOSFET

Afin de suivre les spécifications de la roadmap ITRS et de la Loi de Moore, les composants intégrés sur puce d'abord micrométriques, sont considérablement réduits et atteignent des dimensions nanométriques de nos jours. L'impact lié à la réduction de la longueur de grille induit des effets parasites qui impliquent une dégradation des performances électriques du transistor [1]. Les effets de canaux courts, par exemple, impactent fortement le contrôle électrostatique de la grille sur le canal [2, 3]. La tension de seuil du transistor est ainsi dégradée. La polarisation de drain accentue ce phénomène par un effet connu sous le nom de DIBL (Drain Induced Barrier Lowering) et contribue de ce fait à une augmentation du courant I_{OFF} [4]. Afin d'améliorer au mieux ce contrôle du canal, une solution technologique consiste à diminuer l'épaisseur d'oxyde de grille. Mais malgré l'avantage indéniable que cela procure sur le plan électrostatique, des phénomènes parasites sont alors observés tels que l'apparition de courants de fuite de grille et une quantification des porteurs dans la couche d'inversion. Des solutions existent pour contre-carrer ces effets parasites mais l'adaptation de l'architecture du transistor conventionnel pour les noeuds technologiques à venir, reste un véritable défi.

1.1.2 Vers de nouvelles architectures de transistors MOSFET

Depuis le début des années 90, une organisation internationale, l'ITRS [5], composée d'industriels et d'universitaires, planifie le futur de l'industrie de la microélectronique et la réduction d'échelle du transistor MOSFET. Selon cette organisation, la miniaturisation jusqu'aux plus petites longueurs de grille des transistors MOSFET nécessite un changement d'architecture. En effet, la poursuite de l'augmentation du niveau de dopage déjà fort du MOSFET bulk conventionnel pour limiter les effets canaux courts n'est plus possible. Cette poursuite est difficile d'un point de vue technologique, le niveau de fluctuation du dopage du silicium devient difficilement maîtrisable. De plus la mobilité des charges du canal est fortement réduite, impliquant une baisse du courant de drain. Les transistors planaires réalisés sur substrat SOI représentent une très bonne alternative au transistor MOSFET conventionnel (Figure 1.2(a)). En effet, l'utilisation d'un substrat SOI permet de réaliser des transistors sur des films de conduction ultra minces appelés UTB (Ultra Thin Body) possédant une épaisseur de semi-conducteur inférieure à 10 nm. Ces transistors ont alors des films de conduction complètement désertés et sont communément appelés FDSOI (Fully Depleted SOI) (Figure 1.2(b)). Le transistor FDSOI est une avancée majeure dans le développement d'architectures de dispositifs car il optimise le couplage entre la grille et le canal. Cette amélioration du couplage électrostatique repose en fait sur l'intégration d'un isolant (oxyde) entre la zone active et le substrat. Cette couche d'isolant enterrée (typiquement d'une centaine de nanomètres d'épaisseur) appelée BOX (Burried Oxide), intercalée entre le substrat et la zone active de silicium, est obtenue grâce au procédé Smart-Cut®, produit par la société Soitec, et utilise le principe du collage moléculaire [6].

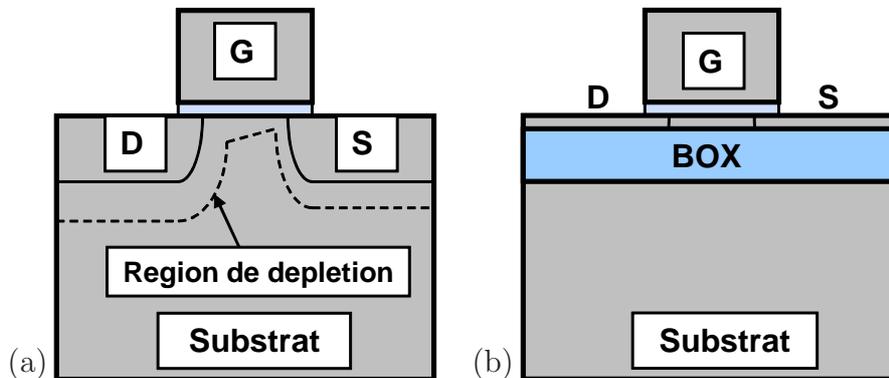


Figure 1.2 – Coupe d'un transistor MOSFET (a) bulk planaire et (b) SOI FD.

Que ce soit au niveau du dispositif ou au niveau du circuit, la technologie SOI est plus avantageuse que la technologie bulk [7] pour les raisons principales résumées ici :

- **Isolation du substrat** : comme le BOX isole les régions de source/drain du substrat, les capacités des jonctions et les pertes de courant sont réduites. Cela augmente la vitesse de commutation et réduit la consommation d'énergie comparé à la technologie bulk.
- **Fiabilité** : le SOI est insensible aux effets de radiations singuliers, puisque la paire électron-trou est générée dans le substrat de silicium épais au lieu de la fine couche active. Cependant, la technologie SOI est plus sensible que la technologie bulk aux effets

de dose totale étant donnée sa large couche d'oxyde enterrée.

- **Matériau en surface** : puisque le processus smart-cut peut être adapté au transfert de n'importe quel matériau semi-conducteur en surface de l'oxyde enterré, les wafers de SOI contraint (sSOI) et d'isolant sur Silicium-Germanium (SGOI) peuvent être facilement fabriqués.
- **Densité augmentée** : contrairement au bulk, ni les puits ni les caissons ne sont nécessaires à l'isolation des dispositifs entre les uns et les autres, permettant des conceptions plus compactes.
- **Architectures de nouveaux dispositifs** : l'utilisation du substrat SOI facilite la fabrication de dispositifs multigrilles d'échelle nanométrique qui sont mieux immunisés contre les effets canaux courts.

Il existe aussi des inconvénients à l'utilisation de la technologie SOI :

- **Auto-échauffement** : la présence de l'isolant ne permet pas la dissipation assez rapide de la chaleur du dispositif et produit un auto-échauffement. Cela mène à une dégradation du courant du dispositif lorsque le courant à dissiper est important.
- **Coût important du wafer** : le cout du wafer est environ deux à trois fois plus chère que le wafer bulk. Néanmoins le surcout de la technologie SOI n'est que de 10% si nous considérons le processus complet de fabrication.

Les avantages de la technologie SOI pour les circuits CMOS surpassent ainsi grandement ses inconvénients. Pour cette raison, de grandes compagnies comme IBM, AMD, Sony, Freescale intègrent déjà des transistors SOI dans leur produit.

Le transistor FDSOI est un très bon candidat pour la succession du transistor conventionnel, mais celui-ci est néanmoins limité par son architecture planaire pour les noeuds technologiques ultimes au niveau du contrôle électrostatique de la grille sur le canal de conduction [8]. En effet, les transistors à multiple grilles font appel à des architectures 3D et semblent ainsi être mieux immunisés face aux effets de canaux courts.

1.1.3 Les dispositifs multigrilles

Pour des raisons de commodité, nous nommerons les transistors à multiple grilles, transistors multigrilles. L'immunité des transistors contre les effets canaux courts peut être sensiblement améliorée par l'ajout de grilles dans la structure autour du canal de conduction [9]. Cette modification de l'architecture permet d'envisager des structures telles que les transistors double grille et FinFET où le film de silicium est contrôlé par deux grilles. L'émergence de ces nouveaux dispositifs est de plus en plus prononcée dans les secteurs R&D qui s'attachent à trouver des solutions performantes et, si possible, moins coûteuses permettant le remplacement éventuel du transistor conventionnel. Aussi, l'architecture planaire classique fait progressivement place aux architectures à trois dimensions. Les architectures 3D permettent l'ajout de canaux de conduction et laissent ainsi la possibilité d'augmenter la densité de courant du transistor ce qui

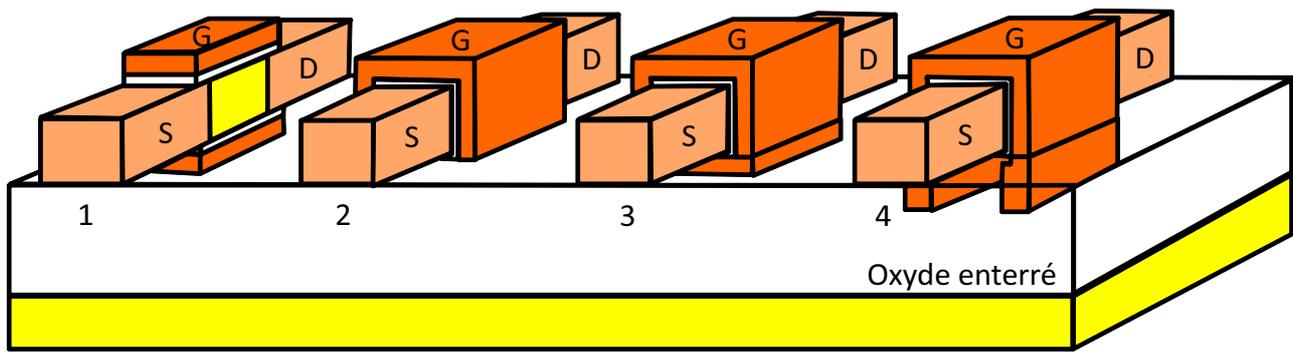


Figure 1.3 – Structures de transistors MOS multigrilles, 1 : double grille, 2 : triple grille, 3 : quadruple grille, 4 : Π -Gate.

est un réel avantage. Les propriétés isolantes du substrat SOI sont aussi utilisées pour réaliser ces dispositifs 3D afin de prévenir les risques de parasitage avec le substrat. Les dispositifs MOSFET multigrilles font donc partie des architectures les plus prometteuses pour la fabrication des transistors de générations futures.

Le transistor MOSFET double grille (DG MOSFET) (Figure 1.3(1)) fait partie de la catégorie des dispositifs multigrilles. Le film de silicium est inséré entre deux grilles. L'architecture de ce transistor s'apparente à un MOSFET conventionnel possédant un axe de symétrie qui serait situé au milieu du film de silicium. Le transistor double grille introduit la notion d'inversion volumique où le canal d'inversion ne se trouve plus seulement en surface mais dans tout le volume du film de silicium en régime d'inversion faible. Cet effet permet notamment d'améliorer la mobilité des porteurs et est typiquement observé dans les transistors multigrilles puisque le film de silicium mince est alors exposé à des effets de confinement [10]. Toutefois, les procédés technologiques sont difficilement maîtrisables pour ce dispositif et les problèmes de variabilité peuvent être conséquents, notamment le problème de non-alignement des deux grilles qui engendre de fortes capacités parasites.

Les transistors FinFET sont eux aussi des dispositifs émergents qui possèdent une très bonne immunité face aux effets de canaux courts [1]. Ceux-ci sont réalisés verticalement et sont constitués de telle sorte qu'un aileron ou « Fin » en anglais est gravé dans le silicium sur lequel est déposée une grille comme représenté Figure 3.1. La grille vient donc recouvrir deux côtés de l'aileron assurant de cette manière un bon contrôle électrostatique. Alors que le dimensionnement du transistor planaire double grille peut s'opérer aisément par modification de la longueur et de la largeur du transistor, le FinFET est limité par sa largeur qui correspond à la hauteur de l'aileron. En effet, les procédés de gravure du silicium ne permettent pas de réaliser des ailerons de grandes dimensions. Néanmoins, en déposant une grande épaisseur d'oxyde sous la grille supérieure comme cela est également visible Figure 1.4, le transistor FinFET devient un transistor double grille dont les grilles sont auto-alignées, et donc beaucoup plus facile à produire que le transistor double grille planaire. Cet avantage surpasse les limites de dimensionnement du transistor FinFET. Si cette couche d'oxyde négligeant l'influence de la grille supérieure est d'épaisseur égale aux couches des grilles latérales, nous nommerons le transistor, transistor triple grille. Il possède l'avantage par rapport au transistor FinFET de disposer d'un canal de

conduction supplémentaire (Figure 1.3(2)).

Enfin, les transistors à quadruple grille, à grille enrobante ou encore GAA (Gate-All-Around) représentent de nos jours, la solution la plus appropriée pour s'affranchir des effets de canaux courts des noeuds technologiques ultimes (Figure 1.3(3)). Le film de silicium est complètement enrobé par la grille ce qui améliore le contrôle électrostatique du canal de conduction par rapport aux architectures vues précédemment. Cependant ces types de dispositifs sont plus difficiles à fabriquer et plus coûteux en demandant des étapes supplémentaires dans le processus de production.

Des dispositifs intermédiaires entre les dispositifs à grille enrobante et les dispositifs à triple grilles peuvent également être employés. Par exemple, le transistor Π -Gate (Figure 1.3(4)) est un transistor triple grille, et possédant donc ses avantages de fabrication, dont les grilles latérales se prolongent dans l'oxyde enterré pour améliorer le contrôle électrostatique du canal.

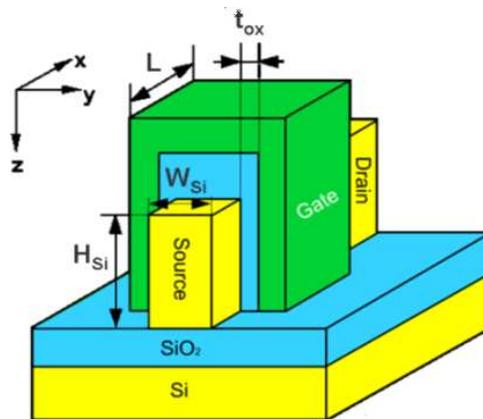


Figure 1.4 – Structure 3D du FinFET modélisé dans cette thèse

1.2 Le premier processeur à base de transistors multigrilles : le « Ivy Bridge » d'Intel

Le dernier né des processeurs d'Intel est le modèle Ivy Bridge avec 1,4 milliards de transistors sur une puce de 160 mm² [11]. Il intègre notamment quatre CPU et un processeur graphique plus puissant que la génération précédente (Figure 1.6). La grande particularité de ce nouveau modèle de processeurs qui retient notre attention dans le cadre de cette thèse, est le fait que pour la première fois dans l'histoire de la microélectronique il est constitué de transistors « 3D », soit des transistors MOSFET bulk tri-gate (dénomination d'Intel) d'une finesse de gravure de 22 nm. Tous les autres processeurs jusque là étaient réalisés à base de transistors MOSFET bulk planaire. La Figure 1.5 propose une vue au microscope électronique du transistor tri-gate. Nous voyons nettement sur cette image la structure à multiple barreaux de silicium en parallèle qui est inhérente à la conception de circuits à base de transistors multigrilles. En effet, toutes les dimensions du barreau de chaque transistor sont fixées pour imposer une faible perte de courant lorsque le transistor est bloqué, soit une plus faible consommation d'énergie. La multiplication

des barreaux en parallèle permet quant à elle d'augmenter le courant total du transistor et ainsi d'améliorer ses performances.

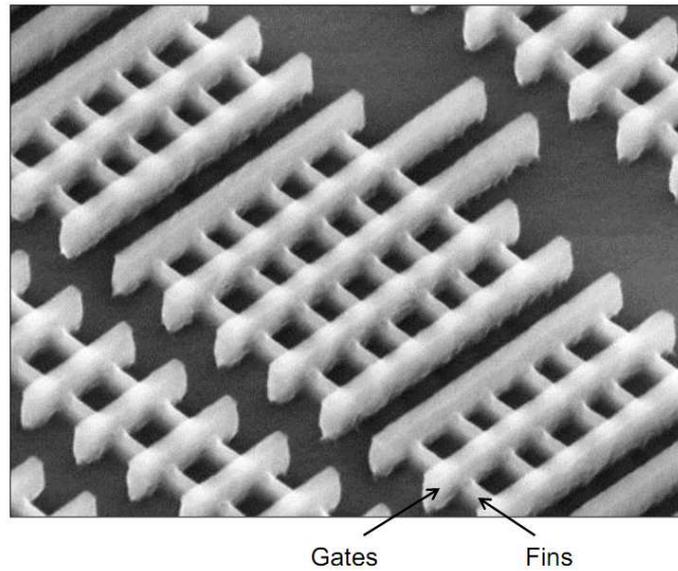


Figure 1.5 – Vue au microscope électronique du transistor tri-gate 22 nm d'Intel (Source : Intel).

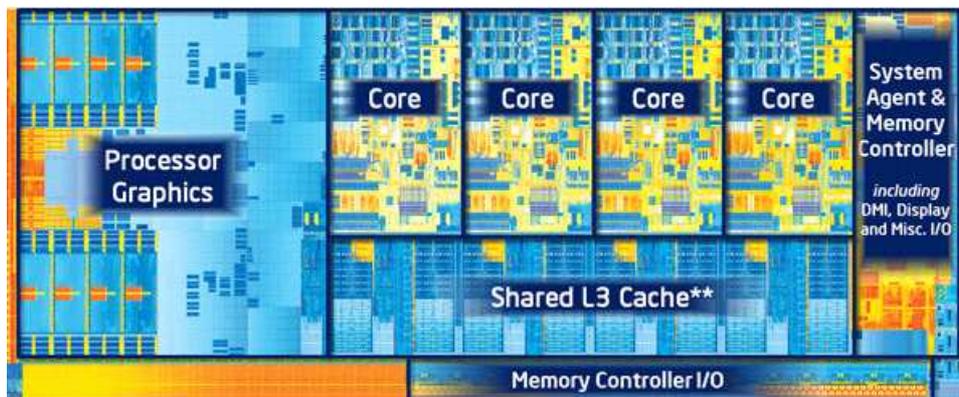


Figure 1.6 – Puce du processeur Ivy Bridge (Source : Intel).

Conclusion

Les objectifs en termes d'intégration et de performances imposés par le marché de la microélectronique amènent les industriels à concevoir des transistors de tailles toujours plus réduites. Cette réduction de dimensions n'est pas sans conséquences puisque les performances du transistor en sont alors fortement dégradées. Des phénomènes parasites importants conduisent alors par exemple à la conduction d'un courant élevé à l'état bloqué du transistor ou encore à l'altération du contrôle de la commutation (de l'état bloqué à celui de passant). La dégradation de performances est alors telle que les procédés de fabrication standards ne peuvent plus être conservés en l'état. Les architectures de transistors à grilles multiples représentent aujourd'hui les solutions les plus prometteuses pour remplacer le transistor conventionnel. L'application de transistors multigrilles est d'ors-et-déjà réalisée dans des produits commerciaux. Le

développement de circuits à base de tels dispositifs nécessite le développement de modèles compacts adaptés à leur fonctionnement 3D.

Nous proposons de développer dans cette thèse la modélisation compacte du transistor FinFET SOI pour les plus ultimes géométries de ce dispositif. L'exploitation de points communs de fonctionnement des transistors multigrilles nous a permis d'établir une modélisation compacte générique pour un large ensemble de transistors multigrilles à canaux long.

Références

- [1] M. TANG, F. PRÉGALDINY, C. LALLEMENT et J.-M. SALLESE, “Explicit compact model for ultranarrow body FinFETs”, *IEEE Trans. Electron Devices*, vol. 56, n° 7, pp. 1543–1547, 2009.
- [2] K. BJÖRKQVIST et T. ARNBORG, “Short channel effects in MOS-transistors”, *Physica Scripta*, vol. 24, pp. 418–421, 1981.
- [3] H. S. LEE, “An analysis of the threshold voltage for short-channel IGFET’s”, *Solid-State Electron.*, vol. 16, pp. 1407–1417, 1973.
- [4] R. R. TROUTMAN, “VLSI limitations from drain-induced barrier lowering”, *IEEE Trans. Electron Devices*, vol. 26, n° 4, pp. 461–469, 1979.
- [5] International Technology Roadmap for Semiconductors (ITRS), rap. tech., 2011. adresse : <http://public.itrs.net/>.
- [6] M. BRUEL, “Silicon on insulator material technology”, vol. 31, n° 14, pp. 1201–1202, 1995.
- [7] S CRISTOLOVEANU, “Silicon on insulator technologies and devices : from present to future”, *Solid-State Electron.*, vol. 45, pp. 1403–1411, 2001.
- [8] G. K. CELLER et S CRISTOLOVEANU, “Frontiers of silicon-on-insulator”, *Journal of Applied Physics*, vol. 93, n° 1, 2003.
- [9] J.-T. PARK et J.-P. COLINGE, “Multiple-gate SOI MOSFETs : device design guidelines”, *IEEE Trans. Electron Devices*, vol. 49, n° 12, pp. 2222–2229, 2002.
- [10] F. BALESTRA, S. CRISTOLOVEANU, M. BENACHIR, J. BRINI et T. ELEWA, “Double-gate silicon-on-insulator transistor with volume inversion : a new device with greatly enhanced performance”, *IEEE Electron Device Lett.*, vol. 8, n° 9, pp. 410–412, 1987.
- [11] S. DAMARAJU, V. GEORGE, S. JAHAGIRDAR, T. KHONDKER, S. MILSTREY R. and Sarkar, S. SIERS, I. STOLERO et A. SUBBIAH, “A 22nm IA multi-CPU and GPU system-on-chip”, dans *IEEE Solid-State Circuits Conference Digest of Technical Papers*, Intel, 2012, pp. 56–57.

Chapitre 2

Modélisation compacte du transistor MOSFET double grille canal long

Sommaire

2.1	État de l’art des modèles compacts du DG MOSFET	14
2.1.1	Modèle de Yuan Taur	14
2.1.2	Modèle PSP FinFET	16
2.1.3	Modèle de Benjamin Iñiguez	18
2.2	Le modèle « InESS-EPFL » du DG MOSFET non dopé	20
2.2.1	Structure du dispositif modélisé	20
2.2.2	Corps du modèle	21
2.2.3	Aspects en température	24
2.3	Le modèle du DG MOSFET dopé	25
2.3.1	État de l’art	25
2.3.2	Développement du modèle basé sur un concept de largeur équivalente de silicium	26
2.3.3	Validation du modèle et discussion	32
2.3.4	Conclusion de la modélisation du DG MOSFET dopé	39
	Conclusion	39
	Références	39

Le transistor MOSFET double grille symétrique est le transistor multigrille le plus simple à étudier et à modéliser parmi les transistors multigrilles. Il dispose d’un plan de symétrie et l’équation de Poisson à deux dimensions est suffisante pour représenter les phénomènes électrostatiques au sein du canal. Sa résolution exacte peut être obtenue selon différentes méthodes qui seront présentées dans l’état de l’art de ce chapitre. Nous présentons dans la deuxième partie de ce chapitre la modélisation compacte du transistor MOSFET double grille développée par notre équipe. Ce modèle est la base de nos travaux de modélisation du transistor FinFET comme nous le verrons au chapitre 3.

La première partie de ce chapitre traite de l'état de l'art de la modélisation compacte du transistor MOSFET double grille. Nous détaillerons ensuite notre propre modèle compact de ce dispositif, valide sur une structure symétrique à canal long non dopé, à partir duquel sont développés plusieurs des modèles présentés dans cette thèse. Nous porterons une attention à la dépendance à la température du dispositif qui est pleinement prédite par le modèle. Dans une dernière partie, nous présenterons une nouvelle modélisation du transistor double grille qui prend en compte le niveau de dopage du canal, quel que soit son niveau, de faiblement dopé à fortement dopé. Le modèle proposé est une extension du modèle non dopé en définissant une largeur équivalente du canal dépendante du niveau de dopage.

2.1 État de l'art des modèles compacts du DG MOSFET

Nous considérons ici un transistor MOSFET double grille complètement déplété (Figure 2.1). Depuis 2000, différents modèles compacts prédisent le comportement du transistor double grille sur tout ou partie de ses régimes de fonctionnement. Parmi ces modèles, nous rappelons ici les approches choisies par Yuan Taur, le groupe de recherche du modèle PSP et Benjamin Iñiguez. La définition de l'orientation (x-y) utilisée est montrée à la Figure 2.1. La largeur du film de silicium est notée W_{Si} , la hauteur est H_{Si} , la longueur de grille est L et l'épaisseur d'oxyde de grille est t_{ox} .

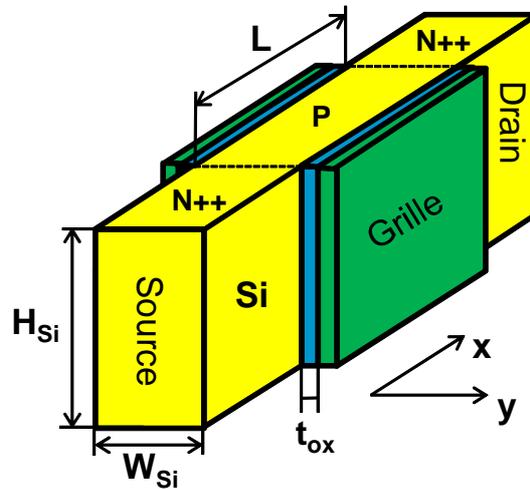


Figure 2.1 – Structure d'un MOSFET double-grille.

2.1.1 Modèle de Yuan Taur

Le premier modèle du DG MOSFET était proposé par Yuan Taur pour les transistors non dopés, en 2000 [1]. Ensuite, par une série de travaux, lui et son équipe ont développé un modèle complet [2-5].

Il résout l'équation de Poisson en 1-D dans le cas d'un transistor à canal long de type-n,

donc avec une densité de trous négligeable :

$$\frac{d^2\Psi(y)}{dy^2} = \frac{e \cdot n_i}{\varepsilon_{Si}} \cdot e^{\frac{\Psi(y)-V_{ch}}{k_b T/e}} \quad (2.1)$$

où V_{ch} est le potentiel de quasi-Fermi. Celui-ci est supposé constant dans la direction-y.

Après deux intégrations, et avec la condition aux limites $d\Psi/dy|_{y=W_{Si}/2} = 0$, il trouve :

$$\Psi(y) = V_{ch} - \frac{2 \cdot k_b \cdot T}{e} \cdot \ln \left[\frac{W_{Si}}{2 \cdot \beta} \cdot \sqrt{\frac{e^2 \cdot n_i}{2 \cdot \varepsilon_{Si} \cdot k_b \cdot T}} \cdot \cos \left(\frac{2 \cdot \beta \cdot y}{W_{Si}} \right) \right] \quad (2.2)$$

où $0 \leq y \leq W_{Si}/2$ et $\beta(x)$ est une constante (dans la direction-y) à déterminer par la condition aux limites :

$$\varepsilon_{ox} \cdot \frac{V_g - \Delta\Phi - \Psi(y = \pm \frac{W_{Si}}{2})}{t_{ox}} = \pm \varepsilon_{Si} \cdot \frac{d\Psi}{dy} \Big|_{y=\pm \frac{W_{Si}}{2}} \quad (2.3)$$

En remplaçant (2.2) dans (2.3), nous avons

$$\frac{e \cdot (V_g - \Delta\Phi - V_{ch})}{2 \cdot k_b \cdot T} - \ln \left(\frac{2}{W_{Si}} \cdot \sqrt{\frac{2 \cdot \varepsilon_{Si} \cdot k_b \cdot T}{e^2 \cdot n_i}} \right) = f_r(\beta) \quad (2.4)$$

$$\text{où } \begin{cases} f_r(\beta) = \ln \beta - \ln(\cos \beta) + \frac{2}{\alpha} \cdot \beta \cdot \tan \beta \\ \alpha = C_{ox}/C_{Si} \end{cases}$$

Le courant de drain est calculé par l'intégrale de Pao-Sah [6] :

$$I_{ds} = \mu \cdot \frac{H_{Si}}{L} \cdot \int_0^{V_{ds}} Q_i(V) \cdot dV_{ch} = \mu \cdot \frac{H_{Si}}{L} \cdot \int_{\beta_s}^{\beta_d} Q_i(V) \cdot \frac{dV_{ch}}{d\beta} \cdot d\beta \quad (2.5)$$

En utilisant la loi de Gauss et (2.2), nous pouvons dire :

$$Q_i = 2 \cdot \varepsilon_{Si} \cdot \frac{2 \cdot k_b \cdot T}{e} \cdot \frac{2 \cdot \beta}{W_{Si}} \quad (2.6)$$

En remplaçant $dV_{ch}/d\beta$ par la différentiation de (2.4) et Q_i par (2.6) dans (2.5), Taur obtient l'expression suivante du courant de drain :

$$I_{ds} = \mu \cdot \frac{H_{Si}}{L} \cdot \frac{4 \cdot \varepsilon_{Si}}{W_{Si}} \cdot \left(\frac{2 \cdot k_b \cdot T}{e} \right)^2 \cdot (g_r(\beta_s) - g_r(\beta_d)) \quad (2.7)$$

où $g_r(\beta) = \beta \cdot \tan \beta - \frac{\beta^2}{2} + \frac{\beta^2}{\alpha} \cdot \tan^2 \beta$. Afin de calculer le courant de drain, β_s et β_d sont indispensables. Ces deux variables intermédiaires sont calculées par (2.4) en remplaçant V_{ch} par V_s et V_d pour β_s et β_d respectivement par la méthode de Newton-Raphson.

A partir de l'expression du courant de drain, le comportement dynamique (g_m, g_{ds}, C_{ij} pour $i, j = g, d, s$) est pris en compte en 2006 [3].

Après une redéfinition du potentiel électrostatique dans le canal en incluant l'influence des source et drain, les effets canaux courts (SCEs) sont inclus physiquement [7] et cette version du modèle est alors validée pour une longueur et une largeur du canal minimum de 24 nm et 5 nm respectivement. Mais pour une largeur si petite ($W_{Si} \ll 20$ nm), les effets quantiques devraient être considérés mais ils ne sont pas encore pris en compte dans le modèle de Taur.

La résolution de β_s et β_d n'est pas explicite, et de plus les SCEs sont inclus d'une manière implicite. Cela freine l'utilisation de cette version du modèle de Taur dans les simulations de circuit où il est préférable de disposer d'un modèle explicite, ce qui veut dire rapide en temps de simulation, plus robuste numériquement et permettant aux concepteurs de faire des dimensionnements « à la main ».

Etant donné l'inconvénient de l'expression implicite, l'équipe de Taur a développé un algorithme explicite [5]. Les paramètres β_s et β_d sont alors résolus explicitement et le modèle de Taur devient un modèle explicite qui convient à la simulation de circuit. Néanmoins, l'algorithme explicite est compliqué, ce qui freine son utilisation dans le dimensionnement à la main. De plus, il n'y a pas encore de résultats montrés sur un dispositif court inférieur à 70nm [8].

2.1.2 Modèle PSP FinFET

Afin d'éviter le problème du calcul implicite, le groupe de recherche du modèle PSP (MOSFET Bulk) a développé sa version de modèle pour le FinFET [9].

Le point de départ du modèle PSP FinFET est le même que le modèle de Taur sauf que certaines grandeurs sont normalisées. Les tensions sont normalisées par rapport à la tension thermodynamique ($U_T = k_b \cdot T/e$) et les paramètres géométriques sont normalisés par rapport à la longueur de Debye ($L_D = \sqrt{(\varepsilon_{Si} U_T)/(2e N_A)}$).

L'équation de Poisson s'écrit, après la normalisation, comme :

$$\frac{d^2 \Psi}{dy^2} = \frac{\exp(\Psi - v_{ch})}{2} \quad (2.8)$$

La solution de (2.8) est :

$$\Psi(x, y) = v_{ch}(x) - \ln \left[\frac{\cos(\sqrt{\alpha(y)} \cdot y/2)^2}{\alpha(y)} \right] \quad (2.9)$$

où $\alpha(y)$ ne dépend pas de x comme β dans le modèle de Taur et il doit être résolu par la condition aux limites :

$$c_{ox}(v_g - \Psi_s(x)) = \Psi(x) \quad (2.10)$$

où $\Psi(x)$ est le potentiel de surface ($=\Psi(x, W_{Si}/2)$) et $(v_g - \Psi_s(x))$ est le champ à la surface ($=\partial\Psi/\partial y(x, W_{Si}/2)$).

L'expression du potentiel (2.9) diffère de (2.2) pour des raisons de normalisation comme dit précédemment.

La densité locale du courant est calculée en prenant en compte les phénomènes de conduction

et de diffusion :

$$-\mu \cdot H_{Si} \cdot q \cdot \frac{dv_{ch}}{dx} = -\mu \cdot H_{Si} \cdot \left(q \cdot \frac{d\Psi}{dx} - \frac{dq}{dx} \right) \quad (2.11)$$

Après intégration de (2.11) suivant la direction-y, l'expression suivante est obtenue :

$$I_{DS} = -\mu \cdot H_{Si} \cdot \int_{-W_{Si}/2}^{W_{Si}/2} q \cdot dy \cdot \frac{dv_{ch}}{dx} = -\mu \cdot H_{Si} \cdot Q \cdot \frac{dv_{ch}}{dx} = -\mu \cdot H_{Si} \cdot \left(\int_{-W_{Si}/2}^{W_{Si}/2} q \cdot \frac{d\Psi}{dx} \cdot dy - \frac{dQ}{dx} \right) \quad (2.12)$$

Puis, après un arrangement de (2.12), nous avons :

$$\int_{-W_{Si}/2}^{W_{Si}/2} q \cdot \frac{d\Psi}{dx} \cdot dy = Q \cdot \frac{dv_{ch}}{dx} + \frac{dQ}{dx} = \tilde{Q} \cdot \frac{d\Psi_s}{dx} \quad (2.13)$$

$$\text{avec } \begin{cases} \tilde{Q} = Q \cdot \left[1 + \frac{\gamma}{4} \cdot g(\theta) \right] \\ g(\theta) = \frac{\sin(2\theta) - 2\theta \cdot \cos(2\theta)}{\theta \cdot \tan(\theta) \cdot [2\theta + \sin(2\theta)]} \\ \theta = \sqrt{\alpha(y)} \cdot W_{Si}/4 \end{cases}$$

Une approximation est faite : $g(\theta) \simeq 0$, donc $\tilde{Q} \simeq Q$. L'expression du courant de drain devient alors :

$$I_{DS} = -\mu \cdot H_{Si} \cdot \left(\tilde{Q} \cdot \frac{d\Psi_s}{dx} - \frac{dQ}{dx} \right) \simeq -\mu \cdot H_{Si} \cdot \left(Q \cdot \frac{d\Psi_s}{dx} - \frac{dQ}{dx} \right) \quad (2.14)$$

Avec $2 \cdot C_{ox}(\Psi_s - v_g) = Q$, (2.14) devient :

$$I_{DS} = -\mu \cdot \frac{H_{Si}}{2 \cdot C_{ox}} \cdot \left(Q \cdot \frac{dQ}{dx} - 2 \cdot C_{ox} \cdot \frac{dQ}{dx} \right) \quad (2.15)$$

Après l'intégration de (2.15) de la source au drain, le courant de drain est finalement donné par :

$$I_{DS} = -\mu \cdot \frac{H_{Si}}{2 \cdot L \cdot C_{ox}} \cdot \bar{Q}^* \cdot \Delta Q^* \quad (2.16)$$

$$\text{où } \begin{cases} \bar{Q}^* = (Q_0^* + Q_L^*) / 2 \\ \Delta Q^* = Q_L^* - Q_0^* \\ Q_{0,L}^* = Q_{0,L} - 2 \cdot C_{ox} \end{cases} .$$

En comparant (2.15) et (2.16), nous avons la relation suivante :

$$\frac{dx}{L} = \frac{Q^* \cdot dQ^*}{\bar{Q}^* \cdot \Delta Q^*} \quad (2.17)$$

A l'aide de (2.17), les expressions de Q_I , Q_S et Q_D peuvent être déduites :

$$-Q_G = Q_I = \bar{Q} - \eta \cdot \Delta Q^* / 6 \quad (2.18)$$

$$Q_D = [\bar{Q} - (\eta + \eta^2/5 - 1) \cdot \Delta Q^* / 6] / 2 \quad (2.19)$$

$$Q_S = [\bar{Q} - (\eta - \eta^2/5 + 1) \cdot \Delta Q^* / 6] / 2 \quad (2.20)$$

avec $\eta = -\Delta Q^* / (2 \cdot \bar{Q}^*)$.

L'avantage de ce modèle canal long est l'expression explicite qui est souhaitée pour un modèle compact. L'intégration des effets de petites géométries dans le modèle du FinFET est similaire au modèle MOSFET-bulk (PSP). Cela exige donc beaucoup de paramètres ce qui rend le modèle complexe et donc très difficile pour une utilisation « calcul à la main ».

2.1.3 Modèle de Benjamin Iñiguez

Un autre modèle explicite est proposé par l'équipe de B. Iniguez [10]. Le point de départ est le même que celui de Sallese et al [11] (voir section 2.2), c'est-à-dire, l'équation de Poisson en négligeant la densité de trous et en 1-D :

$$\frac{d^2 \Psi(y)}{dy^2} = \frac{d^2(\Psi(y) - V_{ch})}{dy^2} = \frac{e \cdot n_i}{\varepsilon_{Si}} \cdot e^{\frac{e \cdot (\Psi(y) - V_{ch})}{k_b \cdot T}} \quad (2.21)$$

(2.21) est résolue comme [11] mais sans normalisation :

$$\begin{aligned} (V_{GS} - \Delta\phi - V_{ch}) + \frac{k_b \cdot T}{e} \cdot \ln \left(\frac{e \cdot n_i \cdot W_{Si}}{8 \cdot C_{ox} \cdot \frac{k_b \cdot T}{e}} \right) - \frac{k_b \cdot T}{e} \cdot \ln \left(\frac{C_{ox}}{C_{Si}} \right) \\ = \frac{Q}{2 \cdot C_{ox}} + \frac{k_b \cdot T}{e} \cdot \left(\ln \left(\frac{Q}{8 \cdot C_{ox} \cdot \frac{k_b \cdot T}{e}} \right) + \ln \left(\frac{C_{Si}}{C_{ox}} + \frac{Q}{8 \cdot C_{ox} \cdot \frac{k_b \cdot T}{e}} \right) \right) \end{aligned} \quad (2.22)$$

Le courant de drain est calculé comme [11] :

$$I_{DS} = \frac{\mu \cdot H_{Si}}{L} \cdot \int_0^{V_{DS}} Q(V_{ch}) \cdot dV_{ch} \quad (2.23)$$

Après la dérivation de (2.22), ils obtiennent une relation (2.24) qui facilite l'intégration de (2.23) :

$$dV_{ch} = -\frac{dQ}{2 \cdot C_{ox}} - \frac{k_b \cdot T}{e} \cdot \left(\frac{dQ}{Q} + \frac{dQ}{Q + 2 \cdot Q_0} \right) \quad (2.24)$$

où $Q_0 = 4 \cdot (k_b \cdot T/e) \cdot C_{Si}$. Puis, après l'intégration de (2.23) en utilisant (2.24) de la source (où $Q = Q_S$) au drain ($Q = Q_D$), ils aboutissent à l'expression du courant de drain :

$$I_{DS} = \frac{\mu \cdot H_{Si}}{L} \cdot \left[2 \cdot \frac{k_b \cdot T}{e} \cdot (Q_S - Q_D) + \frac{Q_S^2 - Q_D^2}{4 \cdot C_{ox}} + 8 \cdot \left(\frac{k_b \cdot T}{e} \right)^2 \cdot C_{Si} \cdot \ln \left(\frac{Q_D + 2 \cdot Q_0}{Q_S + 2 \cdot Q_0} \right) \right] \quad (2.25)$$

La résolution de (2.22) est similaire à [12], et est donc :

$$Q = 2 \cdot C_{ox} \cdot \left(-\frac{2 \cdot C_{ox} \cdot U_T^2}{Q_0} + \sqrt{\left(\frac{2 \cdot C_{ox} \cdot U_T^2}{Q_0} \right)^2 + 4 \cdot U_T^2 \cdot \ln^2 \left[1 + \exp \left(\frac{V_{GS} - V_{th} + \Delta V_{th} - V_{ch}}{2 \cdot U_T} \right) \right]} \right) \quad (2.26)$$

$$\text{où } \begin{cases} V_{th} = V_0 + 2 \cdot U_T \cdot \ln \left(1 + \frac{Q'}{2 \cdot Q_0} \right) \\ V_0 = \Delta \phi - U_T \cdot \ln \left(\frac{e \cdot n_i \cdot W_{Si}}{2 \cdot Q_0} \right) \\ \Delta V_{th} = \frac{\left(\frac{C_{ox} \cdot U_T^2}{Q_0} \right) \cdot Q'}{Q_0 + \frac{Q'}{2}} \\ Q' = C_{ox} \cdot \left(-\frac{2 \cdot C_{ox} \cdot U_T^2}{Q_0} + \sqrt{\left(\frac{2 \cdot C_{ox} \cdot U_T^2}{Q_0} \right)^2 + 4 \cdot U_T^2 \cdot \ln^2 \left[1 + \exp \left(\frac{V_{GS} - V_0 - V_{ch}}{2 \cdot U_T} \right) \right]} \right) \end{cases}$$

Les densités de charge d'inversion aux source (Q_s) et drain (Q_d) sont calculées par (2.26) en prenant $V = 0$ et $V = V_{DS}$ respectivement.

Comme dans le modèle de Taur, la transconductance (g_m) et la conductance de sortie (g_{ds}) sont calculées directement par les densités de charge d'inversion aux source et drain :

$$g_m = \frac{\partial I_{DS}}{\partial V_{GS}} = \frac{\mu \cdot H_{Si}}{L} \cdot (Q_s - Q_d) \quad (2.27)$$

$$g_{ds} = \frac{\partial I_{DS}}{\partial V_{DS}} = \frac{\mu \cdot H_{Si}}{L} \cdot Q_d \quad (2.28)$$

La charge totale dans le canal est obtenue en intégrant la densité de charge d'inversion tout au long du canal :

$$\begin{aligned} Q_I &= -H_{Si} \cdot \int_0^L Q \cdot dx = -H_{Si}^2 \cdot \frac{\mu}{I_{DS}} \cdot \int_0^{V_{DS}} Q^2 \cdot dV_{ch} \\ &= -H_{Si}^2 \cdot \frac{\mu}{I_{DS}} \cdot \int_{Q_s}^{Q_d} \left(\frac{Q^2}{2 \cdot C_{ox}} + U_T \cdot Q + U_T \cdot \frac{Q^2}{Q + 2 \cdot Q_0} \right) \cdot dQ \quad (2.29) \end{aligned}$$

Les charges totales de drain et de source sont calculées par les expressions suivantes :

$$Q_D = -H_{Si} \cdot \int_0^L \frac{x}{L} \cdot Q \cdot dx = -\frac{W_{Si}^3 \cdot \mu^2}{L \cdot I_{DS}^2} \cdot \int_{Q_s}^{Q_d} Q^2 \cdot \left(\left(\frac{Q^2 - Q_s^2}{4 \cdot C_{ox}} \right) + U_T \cdot \left(2 \cdot (Q - Q_s) - 2 \cdot Q_0 \cdot \ln \left(\frac{Q + 2 \cdot Q_0}{Q_s + 2 \cdot Q_0} \right) \right) \cdot \left(\frac{1}{2 \cdot C_{ox}} + U_T \cdot \left(\frac{1}{Q} + \frac{1}{Q + 2 \cdot Q_0} \right) \right) \right) \cdot dQ \quad (2.30)$$

$$Q_S = Q_I - Q_D \quad (2.31)$$

Une fois que les expressions des charges totales du canal, de source et de drain sont obtenues, les capacités peuvent être calculées par les relations suivantes :

$$C_{ij} = \begin{cases} -\frac{\partial Q_i}{\partial V_j} & \text{si } i \neq j \\ \frac{\partial Q_i}{\partial V_j} & \text{sinon} \end{cases} \quad \text{avec } i, j = G, D, S \quad (2.32)$$

où Q_i est la charge correspondant au i -ième terminal.

Le modèle est complété ensuite par la prise en compte des effets de petites géométries comme la tension de saturation, la résistance série, la modulation de la longueur de canal, l'effet DIBL [13] et les effets quantiques [14].

2.2 Le modèle « InESS-EPFL » du DG MOSFET non dopé

La base de nos travaux de modélisation est le modèle compact du transistor DG MOSFET symétrique à canal long faiblement dopé (Figure 2.1) développé par J.-M. Sallese *et al.* [11]. Ce modèle détermine le courant de drain du transistor sur la base du calcul explicite des charges.

2.2.1 Structure du dispositif modélisé

Les deux grilles du transistor (Figure 2.1) sont en métal et ont les mêmes travaux de sortie mid-gap. Les différences de travaux de sortie entre les grilles et le silicium sont donc nulles. Les grilles sont également connectées entre elles pour disposer d'un fonctionnement symétrique. Les épaisseurs des oxydes de grille sont identiques et valent t_{ox} . Le film de silicium, de largeur W_{Si} et de hauteur nommée H_{Si} , se décompose en trois grandes régions : la région de la source et la région du drain en silicium fortement dopé, et une région centrale d'un dopage inverse à celui des régions terminales. Les jonctions P-N ainsi créées sont distantes de la longueur de grille L .

2.2.2 Corps du modèle

Le corps du modèle est basé sur certaines hypothèses physiques. La mobilité des porteurs est considérée comme constante le long du canal d'inversion. Les effets de mécanique quantique ne sont pas pris en compte. L'effet de polydépétion n'est pas considéré étant donné l'utilisation de grilles métalliques de type mid-gap.

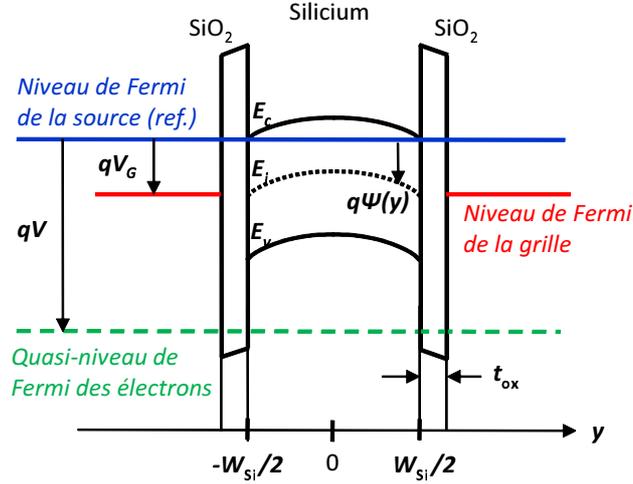


Figure 2.2 – Diagramme des bandes d'énergie suivant l'axe normal aux grilles.

La Figure 2.2 décrit le diagramme des bandes d'énergie à travers le film de silicium. Les axes x et y sont toujours considérés suivant la direction longitudinale et la direction transversale au canal respectivement. Le potentiel $\Psi(y)$ correspond à l'écart entre le niveau intrinsèque du silicium et le niveau de Fermi de la source. La source est utilisée comme référence pour les potentiels ($V_s = 0V$). V_{ch} correspond au potentiel du quasi-niveau de Fermi des électrons et est constant sur toute la largeur du film de silicium en raison de la répartition symétrique des porteurs dans le film.

Modèle de charges

Etant donné le faible niveau de dopage de la partie centrale du film de silicium, nous supposons que cette région est complètement déplétée. Ainsi, nous pouvons négliger les trous dans l'équation de Poisson, écrite en 1-D selon l'axe y et utilisant les statistiques de Boltzmann non dégénérés :

$$\frac{d^2\Psi(y)}{dy^2} = \frac{d^2(\Psi(y) - V_{ch})}{dy^2} = \frac{e \cdot n_i}{\epsilon_{Si}} \cdot e^{\frac{\Psi(y) - V_{ch}}{U_T}} \quad (2.33)$$

A partir de l'équation de Poisson, nous souhaitons obtenir une expression analytique du champ électrique à travers le film de silicium. Tout d'abord, nous établissons simplement une relation entre la dérivée seconde du potentiel et sa dérivée première :

$$\frac{d}{dy} \left[\left(\frac{d\Psi(y)}{dy} \right)^2 \right] = 2 \cdot \frac{d\Psi(y)}{dy} \cdot \frac{d^2\Psi(y)}{dy^2} \quad (2.34)$$

En combinant (2.33) et (2.34), nous obtenons :

$$\frac{d}{dy} \left[\left(\frac{d\Psi(y)}{dy} \right)^2 \right] = 2 \cdot \frac{d\Psi(y)}{dy} \cdot \frac{e \cdot n_i}{\varepsilon_{Si}} \cdot e^{\frac{\Psi(y)-V_{ch}}{U_T}} \quad (2.35)$$

Puis, en intégrant (2.35) du centre à l'interface canal/oxyde, nous pouvons finalement exprimer le champ électrique de surface :

$$\begin{aligned} E_s^2 &= \int_{\Psi_0}^{\Psi_s} d \left(\frac{d\Psi(y)}{dy} \right)^2 = \int_{\Psi_0}^{\Psi_s} \frac{2 \cdot e \cdot n_i}{\varepsilon_{Si}} \cdot e^{\frac{\Psi-V_{ch}}{U_T}} \cdot d\Psi \\ E_s^2 &= \frac{2 \cdot e \cdot n_i \cdot U_T}{\varepsilon_{Si}} \cdot \left(e^{\frac{\Psi_s-V_{ch}}{U_T}} + C_1 \right) \end{aligned} \quad (2.36)$$

où $\Psi_s = \Psi(\pm W_{Si}/2)$ est le potentiel de surface, $\Psi_0 = \Psi(0)$ est le potentiel au centre du canal et $C_1 (= -e^{\frac{\Psi_0-V_{ch}}{U_T}})$ est la constante d'intégration volontairement non exprimée dans l'expression du champ électrique à ce stade du développement du modèle.

A partir du champ électrique de surface, nous exprimons la densité de charges de grille $Q_g = E_s \cdot \varepsilon_{Si}$. La densité de charges considérée ici est celle contrôlée par une seule grille. Les potentiels peuvent donc s'exprimer en fonction de la densité de charges de grille et de la constante d'intégration :

$$\Psi_s - V_{ch} = U_T \cdot \ln \left(\frac{Q_g^2}{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} - C_1 \right) \quad (2.37)$$

Une autre relation de potentiels est obtenue en appliquant le théorème de Gauss à travers l'oxyde de grille :

$$V_g^* - \Psi_s = \frac{Q_g}{C_{ox}} \quad (2.38)$$

où $V_g^* = V_g - \Delta\Phi_{ms}$ et $\Delta\Phi_{ms}$ est la différence de travaux de sortie entre la grille et le silicium. En combinant (2.37) et (2.38), nous introduisons la tension de grille dans la relation entre les potentiels et la densité de charge de grille :

$$V_g^* - V_{ch} = \frac{Q_g}{C_{ox}} + U_T \cdot \ln \left(\frac{Q_g^2}{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} - C_1 \right) \quad (2.39)$$

L'intégration de la relation (2.36) aboutit à la solution analytique de la densité de charge de grille exprimée en fonction de C_1 :

$$Q_g = \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{-C_1} \cdot \tan \left(\sqrt{\frac{2 \cdot e \cdot n_i \cdot U_T}{\varepsilon_{Si}}} \cdot \frac{W_{Si}}{4 \cdot U_T} \cdot \sqrt{-C_1} \right) \quad (2.40)$$

La relation (2.39) est constituée de deux termes, l'un est proportionnel à Q_g et l'autre est de forme logarithmique. En forte inversion, l'impact de C_1 est très faible car le terme logarithmique est négligeable, c'est-à-dire que le couplage entre les grilles est négligeable en forte inversion. Par conséquent, C_1 représente le couplage électrostatique entre les deux électrodes de grille.

Cependant, en faible inversion, à cause de la répartition des porteurs dans toute la largeur du film de silicium, le couplage entre les grilles n'est plus négligeable et l'impact de C_1 doit être pris en compte. Un développement limité de (2.40) à l'ordre 1 autour de la valeur de $C_1 = 0$ en faible inversion permet ensuite de déterminer C_1 en fonction de Q_g :

$$C_1 \simeq -\frac{2 \cdot Q_g}{e \cdot n_i \cdot W_{Si}} \quad (2.41)$$

Le remplacement de C_1 dans (2.39) avec (2.41) permet de trouver la relation entre les charges et les potentiels :

$$V_g^* - V_{ch} - U_T \cdot \ln\left(\frac{2}{e \cdot n_i \cdot W_{Si}}\right) = \frac{Q_g}{C_{ox}} + U_T \cdot \ln(Q_g) + U_T \cdot \ln\left(1 + \frac{Q_g}{4 \cdot C_{Si} \cdot U_T}\right) \quad (2.42)$$

Normalisations des grandeurs électriques

Afin de simplifier l'utilisation des relations de notre modèle pour les concepteurs de circuits, nous adoptons une normalisation semblable à celle existante dans le modèle EKV [15] (modèle du MOSFET bulk). Le tableau 2.1 regroupe tous les facteurs de normalisation utilisés dans notre modèle.

Tableau 2.1 – Facteurs de normalisation des grandeurs électriques.

Grandeurs électriques	Facteurs de normalisation
Tensions	U_T
Courants	$I_{SP} = 4 \cdot \mu \cdot C_{ox} \cdot U_T^2 \cdot H_{Si}/L$
Charges	$Q_0 = 4 \cdot C_{ox} \cdot U_T$
Capacités	$C_{ox_total} = 2 \cdot C_{ox}$

Par convention, nous noterons les grandeurs non normalisées par des lettres majuscules, et les grandeurs normalisées par des lettres minuscules. Par exemple, la tension V est transformée en la tension normalisée $v = V/U_T$.

Ainsi, la relation charges-potentiels (2.42) normalisée s'écrit simplement :

$$v_g^* - v_{ch} - v_{to} = 4 \cdot q_g + \ln q_g + \ln\left(1 + q_g \cdot \frac{C_{ox}}{C_{Si}}\right) \quad (2.43)$$

où $C_{Si}(= \varepsilon_{Si}/W_{Si})$ est la capacité du film de silicium et v_{to} est la tension de seuil canal long :

$$v_{to} = \ln\left(\frac{8 \cdot C_{ox} \cdot U_T}{e \cdot n_i \cdot W_{Si}}\right) = -\ln\left(\frac{q_{int}}{2}\right) \quad (2.44)$$

q_{int} est la densité intrinsèque de charges mobiles.

Modèle de courant

Nous considérons le modèle de transport drift-diffusion pour le développement de notre

modèle de courant. Le courant de drain s'exprime donc ainsi :

$$I_D = -\mu \cdot \frac{H_{Si}}{L} \cdot \int_{V_s}^{V_d} Q_m dV_{ch} \quad (2.45)$$

où $Q_m (= -2 \cdot Q_g)$ est la densité de charge dans le canal.

Puis, en faisant apparaître la dérivée $\frac{dv_{ch}}{dq_m}$ de la relation (2.43) dans la relation (2.45) :

$$I_D = -\mu \cdot \frac{H_{Si}}{L} \cdot \int_{q_{ms}}^{q_{md}} (4 \cdot C_{ox} \cdot U_T \cdot q_m) \cdot \left(\frac{d(v_{ch} \cdot U_T)}{dq_m} \right) \cdot dq_m \quad (2.46)$$

le courant de drain normalisé devient :

$$i = \frac{I_D}{I_{SP}} = -q_m^2 + 2 \cdot q_m + 2 \cdot \frac{C_{Si}}{C_{ox}} \cdot \ln \left(1 - q_m \cdot \frac{C_{ox}}{2 \cdot C_{Si}} \right) \Big|_{q_{ms}}^{q_{md}} \quad (2.47)$$

Grâce à (2.47), le courant de drain peut être calculé directement par la densité de charges mobiles qui peut être résolue par (2.43). Cependant, (2.43) ne peut pas être résolue explicitement, ce qui limite l'utilisation de ce modèle dans les simulations de circuit. Afin de résoudre explicitement (2.43), un algorithme d'inversion a été développé [16]. Il permet d'obtenir $q_m = f(v_g, v_d)$. L'algorithme est détaillé en Annexe de [17].

2.2.3 Aspects en température

Nous proposons maintenant de mettre en évidence que le modèle du DG canal long prend déjà naturellement en compte l'effet de la température. La température se retrouve dans la concentration intrinsèque n_i et la tension thermodynamique U_T , qui sont incluses dans les facteurs de normalisation et, dans une moindre mesure, dans la tension de seuil normalisée (voir 2.44). Afin d'illustrer la validité de la dépendance à la température du modèle DG canal long, nous avons effectué des simulations numériques TCAD 2-D d'un DG MOSFET pour différentes températures, en imposant une mobilité constante de $\mu_0 = 1000 \text{ cm}^2/\text{V.s}$ à 300 K. La mise en œuvre de ces simulations numériques est décrite au Chapitre 5. La mobilité μ vaut ici la mobilité à champ faible μ_0 laquelle dépend uniquement de la température. Nous utilisons dans le modèle compact la loi de variation de la mobilité avec la température suivante :

$$\mu_0 = \mu_{0300K} \cdot \left(\frac{T}{300K} \right)^{-3/2} \quad (2.48)$$

La figure 2.3 montre le bon accord entre les résultats de ces simulations et ceux du modèle compact, pour une gamme de température de 275 à 400 K.

Nous venons de rappeler comment a été développé le modèle compact du MOSFET double

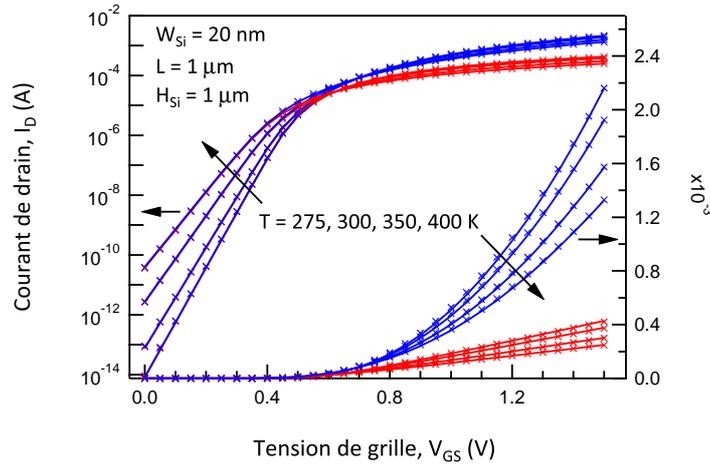


Figure 2.3 – Courant de drain d'un DG MOSFET en fonction de la tension de grille en régimes linéaire et saturé. Symboles : simulations 2-D, lignes : modèle DG [11].

grille à canal faiblement dopé [11], modèle à la base du travail de modélisation compacte de notre équipe de recherche. Nous allons maintenant présenter une extension de ce modèle, développé lors de cette thèse, pour des DG MOSFET symétriques à canal long et dopé, valide pour une large gamme de dopage.

2.3 Le modèle du DG MOSFET dopé

Dans cette section, nous proposons de développer un modèle analytique pour le DG MOSFET symétrique dopé à canal long, valide dans toutes les régions d'opération. Nous montrons qu'un DG MOSFET dopé peut être équivalent à un DG MOSFET non dopé dont la largeur du film de silicium est changée. La définition d'une largeur équivalente de silicium non dopé nous permet de modéliser les dispositifs dopés. Le décalage de tension de seuil induit par le dopage vis-à-vis d'un dispositif non dopé est implicitement pris en compte. Le formalisme développé pour le dispositif non dopé est ensuite appliqué en utilisant cette géométrie de canal équivalente. En considérant les mêmes paramètres physiques, nous démontrons que cette approche est en accord avec des simulations numériques TCAD. Ce concept est ainsi une base intéressante pour l'unification des DG MOSFET symétriques dopés et non dopés.

2.3.1 État de l'art

Aujourd'hui, le DG MOSFET est probablement le dispositif nanoélectronique utilisé en technologie CMOS le plus modélisé depuis le premier travail publié par Taur [1] il y a plus de dix ans. En particulier, la modélisation compacte analytique du DG MOSFET symétrique non dopé a été investiguée en détails [2, 11, 18, 19], incluant des caractéristiques avancées telles que les effets canaux courts (SCEs) et l'effet de quantification (QME) [7, 13, 20, 21]. D'autre part, moins de travail a été fait sur des dispositifs DG dont le canal est dopé. Les modèles dans [18] et [22] supposent un DG MOSFET fortement dopé. Le modèle présenté dans [23–26] est valide dans un large intervalle de dopage et aussi comme un cas limite pour les dispositifs non dopés.

Cependant ce modèle a besoin de fonctions d'interpolation, et utilise une expression empirique pour la différence de potentiels entre la surface et le centre de la couche de silicium. Dans [27], les auteurs proposent une solution basée sur le modèle développé dans [11] mais introduisent des approximations arbitraires pour assurer la « continuité » entre les dispositifs dopés et non dopés. Bien que, pour des questions d'appariement et de mobilité, la couche de silicium devrait idéalement être non dopée, de nombreux dispositifs ont leur canal dopé afin d'affiner leur tension de seuil. Ainsi un modèle compact pour DG MOSFET, valide dans un intervalle de concentration de dopage utilisé dans les processus de fabrication, incluant les dispositifs non dopés, est nécessaire.

Nous proposons de démontrer par la suite comment un film de silicium dopé peut être converti en un film de silicium non dopé de largeur équivalente de film de silicium, en raisonnant du point de vue des densités de charges du canal, et donc du point de vue du courant total. Des validations avec des simulations TCAD confirment que le concept de “largeur équivalente” est valide jusqu'à ce que le semi-conducteur atteigne la limite de la déplétion totale. Au delà, le couplage entre les grilles disparaît, signifiant que le concept du DG MOSFET n'est plus. La transition entre la déplétion totale et la déplétion partielle dans des DG MOSFET dopés a été traitée par Liu et al. [28]. Les auteurs calculent les potentiels, les charges et le courant selon une méthode itérative approximée combinée à l'intégrale double de Pao-Sah, méthode qui requière donc un calcul numérique par itération. Le formalisme simple présentée ici pour des DG MOSFET symétriques totalement déplétés dopés et non dopés, s'avère précis même lorsque le dopage dépasse le critère de déplétion partielle rapporté dans [28]. De plus, cette approche est particulièrement adaptée aux questions de conceptions analogiques des MOSFET multigrilles, et peut facilement être implémentée dans des simulateurs électriques.

2.3.2 Développement du modèle basé sur un concept de largeur équivalente de silicium

Nous supposons un DG MOSFET de type N : sa couche de semi-conducteur de largeur W_{Si} et de hauteur H_{Si} est dopé P avec un niveau de dopage N_a , et ses électrodes de source et drain sont fortement dopé N. Une représentation énergétique de la structure considérée est montrée à la Figure 2.4.

2.3.2.1 Electrostatique dans le DG MOSFET dopé

Supposant que le DG MOSFET dopé opère dans le mode “déplétion-inversion”, nous proposons de négliger la contribution des trous dans la densité de charges locale. Cette hypothèse a été analysée par Zhou et al. [29], qui exprime une solution rigoureuse de l'équation de Poisson pour les DG MOSFET symétriques non dopés. Il est reporté que cette simplification génère des erreurs mineures lorsque le transistor fonctionne loin de la condition de bandes plates (particulièrement pour les transistors dopés [29]). Cependant, la prise en compte des trous empêche

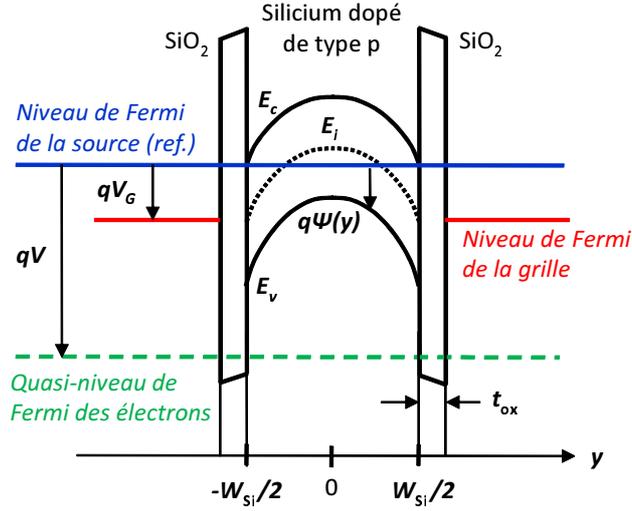


Figure 2.4 – Diagramme de bandes du DG MOSFET symétrique dopé.

l'obtention de toute solution analytique.

Nous proposons ici de résoudre l'équation de Poisson appliquée au film de silicium dopé N_a , en ne considérant que les électrons comme charges mobiles :

$$\frac{d^2\psi}{dy^2} = \frac{e}{\varepsilon_{Si}} \cdot \left(n_i \cdot e^{\frac{\psi - V_{ch}}{U_T}} + N_a \right) \quad (2.49)$$

où N_a est la concentration du dopage. Tous les potentiels sont référencés par rapport au potentiel de Fermi de la source, comme montré à la Figure 2.4. A la différence du DG MOSFET non dopé [2, 30], où $N_a = 0$, la relation (2.49) n'a pas de solution analytique, et ainsi, nous devons chercher une solution approchée.

Une solution peut être obtenue en suivant la méthode utilisée pour développer le corps du modèle non dopé à la section 2.2.2. Ainsi, en multipliant des deux côtés de l'équation de Poisson par $2 \cdot d\psi/dy$, nous obtenons :

$$2 \cdot \frac{d\psi}{dy} \cdot \frac{d^2\psi}{dy^2} = \frac{d}{dy} E^2 = \frac{2 \cdot e}{\varepsilon_{Si}} \cdot \left(n_i \cdot e^{\frac{\psi - V_{ch}}{U_T}} + N_a \right) \cdot \frac{d\psi}{dy} \quad (2.50)$$

En intégrant (2.50) de $y=0$ à $y=W_{Si}/2$, étant donné le champ électrique nul au centre (imposé par la symétrie du transistor), nous pouvons écrire :

$$E_s^2 = \frac{2 \cdot e \cdot n_i \cdot U_T}{\varepsilon_{Si}} \cdot \left[\left(e^{\frac{\psi_s - V_{ch}}{U_T}} - e^{\frac{\psi_0 - V_{ch}}{U_T}} \right) + \frac{N_a}{n_i} \cdot \frac{\psi_s - \psi_0}{U_T} \right] \quad (2.51)$$

où $\psi_s = \psi(W_{Si}/2)$ et E_s sont respectivement le potentiel de surface et le champ électrique à l'interface semi-conducteur/isolant, et où $\psi_0 = \psi(0)$ est le potentiel au centre du film de silicium. Tout comme dans le développement du modèle non dopé, nous proposons d'introduire la variable C_1 dans l'équation du champ électrique :

$$C_1 = -e \frac{\psi_0 - V_{ch}}{U_T} \quad (2.52)$$

Cette variable est une constante d'intégration du potentiel, fonction du potentiel du centre du film.

Puisque le théorème de Gauss lie le champ électrique de surface à la charge de grille, $Q_g = E_s \cdot \varepsilon_{Si}$, la relation (2.51) peut aussi être écrite :

$$\psi_s - V_{ch} = U_T \cdot \ln \left(\frac{Q_g^2}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} - C_1 - \frac{N_a}{n_i} \cdot \frac{\psi_s - \psi_0}{U_T} \right) \quad (2.53)$$

D'autre part, le potentiel de grille est relié à la densité de charge de grille par l'intermédiaire de la chute de tension à travers la capacité de grille C_{ox} , soit :

$$V_G - \Delta\phi - \psi_s = \frac{Q_g}{C_{ox}} \quad (2.54)$$

En combinant (2.53) avec (2.54), la charge de grille peut être exprimée en termes de potentiels de canal et de grille :

$$V_G - \Delta\phi - V_{ch} = \frac{Q_g}{C_{ox}} + U_T \cdot \ln \left(\frac{Q_g^2}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} - C_1 - \frac{N_a}{n_i} \cdot \frac{\psi_s - \psi_0}{U_T} \right) \quad (2.55)$$

La relation (2.55) est le pendant de la relation (2.39) obtenue pour le cas non dopé.

En supposant que la densité de charge mobile peut être négligée en régime de faible inversion, la solution de l'équation de Poisson (2.49) dans ce régime est :

$$\psi(y) = \psi_0 + \frac{e}{2 \cdot \varepsilon_{Si}} \cdot N_a \cdot y^2 \quad (2.56)$$

Nous pouvons ainsi exprimer la différence de potentiel entre la surface et le centre du film de silicium :

$$\psi_s - \psi_0 = \frac{e \cdot N_a \cdot W_{Si}^2}{8 \cdot \varepsilon_{Si}} \quad (2.57)$$

En introduisant l'expression (2.57) dans (2.55), nous obtenons alors une nouvelle relation entre charges et potentiels dans laquelle seule la variable C_1 est encore inconnue :

$$V_G - \Delta\phi - V_{ch} = \frac{Q_g}{C_{ox}} + U_T \cdot \ln \left(\frac{Q_g^2}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} - C_1 - \left(\frac{e \cdot N_a \cdot W_{Si}}{2} \right)^2 \cdot \frac{1}{2 \cdot e \cdot \varepsilon_{Si} \cdot U_T \cdot n_i} \right) \quad (2.58)$$

2.3.2.2 Détermination de C_1

De même que pour le DG non dopé, nous proposons de calculer le terme inconnu C_1 de la fonction logarithmique à partir de la valeur qu'il prend en faible inversion.

Dans le modèle que nous avons développé pour le cas non dopé, C_1 a été déterminé à partir du champ électrique de surface $E(W_{Si}/2)$ en faible inversion, obtenue de la solution exacte du potentiel. Dans le cas du silicium dopé, la situation est différente puisque aucune solution analytique existe, rendant impossible la détermination de C_1 par le champ électrique de surface.

Cependant, C_1 est aussi relié à la chute de potentiel au centre du semi-conducteur ψ_0 par (2.52). Et en faible inversion, ce potentiel est lié au potentiel $\psi(y)$ du film de silicium dans la relation (2.56), et donc à la densité de charges à travers les statistiques du semi-conducteur. Nous proposons donc d'établir une expression de la variable C_1 en fonction de la densité de charges mobiles. En supposant un semi-conducteur non dégénéré, la densité de charges mobiles sur la largeur W_{Si} du film de silicium en faible inversion, est défini par :

$$-Q_m = 2 \cdot e \cdot n_i \cdot \int_{-W_{Si}/2}^0 e^{\frac{\psi(y)-V_{ch}}{U_T}} dy \quad (2.59)$$

En insérant la relation (2.56) dans (2.59), nous introduisons le potentiel ψ_0 dans la définition de la densité de charges :

$$-\frac{Q_m}{2} = e \cdot n_i \cdot \int_{-W_{Si}/2}^0 e^{\frac{\psi_0 - V_{ch} + \frac{e}{2 \cdot \epsilon_{Si}} \cdot N_a \cdot y^2}{U_T}} dy \quad (2.60)$$

La relation (2.60) peut être manipulée pour mettre en évidence C_1 :

$$\begin{aligned} -\frac{Q_m}{2} &= e \cdot n_i \cdot e^{\frac{\psi_0 - V_{ch}}{U_T}} \cdot \int_{-W_{Si}/2}^0 e^{\frac{e}{2 \cdot \epsilon_{Si} \cdot U_T} \cdot N_a \cdot y^2} dy \\ &= -C_1 \cdot e \cdot n_i \cdot \int_{-W_{Si}/2}^0 e^{\frac{e}{2 \cdot \epsilon_{Si} \cdot U_T} \cdot N_a \cdot y^2} dy \end{aligned} \quad (2.61)$$

Finalement, nous obtenons une expression où la variable C_1 est proportionnelle à la charge mobile globale du semi-conducteur selon une quantité qui dépend à la fois de la concentration du dopage et de la largeur du semi-conducteur :

$$C_1 = \frac{Q_m}{2} \cdot \frac{1}{e \cdot n_i \cdot \int_{-W_{Si}/2}^0 e^{\frac{e}{2 \cdot \epsilon_{Si} \cdot U_T} \cdot N_a \cdot y^2} dy} \quad (2.62)$$

Nous pouvons remarquer que la relation (2.62) revient à l'expression de la constante d'intégration (2.41) obtenue pour le cas non dopé avec $N_a = 0$:

$$\begin{aligned} C_1 &= \frac{Q_m}{2} \cdot \frac{1}{e \cdot n_i \cdot \int_{-W_{Si}/2}^0 e^{\frac{e}{2 \cdot \epsilon_{Si} \cdot U_T} \cdot N_a \cdot y^2} dy} \\ &\stackrel{\text{non dope}}{=} \frac{-Q_g}{e \cdot n_i \cdot \int_{-W_{Si}/2}^0 dy} = \frac{-Q_g}{e \cdot n_i \cdot \frac{W_{Si}}{2}} \end{aligned} \quad (2.63)$$

En plus de définir la variable C_1 en fonction de la charge mobile totale Q_m , nous devons faire de même pour la densité de charges de grille Q_g , qui parallèlement, inclut également la

contribution des accepteurs ionisés Q_{dep} du semi-conducteur dopé :

$$-2 \cdot Q_g = Q_m + Q_{dep} = Q_m - e \cdot N_a \cdot W_{Si} \quad (2.64)$$

En insérant (2.62) et (2.64) dans la relation (2.58), nous obtenons une relation entre la densité de charges mobiles et les potentiels :

$$V_G - \Delta\phi - V_{ch} = \frac{\left(\frac{-Q_m}{2} + X\right)}{C_{ox}} + U_T \cdot \ln \left(\frac{\left(\frac{-Q_m}{2} + X\right)^2}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} + \frac{\frac{-Q_m}{2}}{e \cdot n_i \cdot \int_{-W_{Si}/2}^0 e^{\frac{X}{W_{Si} \cdot \varepsilon_{Si} \cdot U_T} \cdot y^2} dy} - \frac{X^2}{2 \cdot e \cdot \varepsilon_{Si} \cdot U_T \cdot n_i} \right) \quad (2.65)$$

où $X = e \cdot N_a \cdot W_{Si}/2$ représente la moitié des charges fixes introduites par le dopage du silicium. En réarrangeant les termes, nous avons :

$$V_G - \left(\Delta\phi + \frac{X}{C_{ox}} \right) - V_{ch} = \frac{\left(\frac{-Q_m}{2}\right)}{C_{ox}} + U_T \cdot \ln \left(- \left(\frac{Q_m}{2} \right) \right) + U_T \cdot \ln \left(\frac{\frac{-Q_m}{2}}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} + \frac{1}{e \cdot n_i} \left(\frac{1}{\int_{-W_{Si}/2}^0 e^{\frac{X}{W_{Si} \cdot \varepsilon_{Si} \cdot U_T} \cdot y^2} dy} - \frac{X}{\varepsilon_{Si} \cdot U_T} \right) \right) \quad (2.66)$$

Sous cette forme, la relation entre charges et potentiels peut être mise en regard avec cette même relation du cas non dopé (2.42) réécrite ici :

$$V_G - \Delta\phi - V_{ch} = \frac{-\left(\frac{Q_m}{2}\right)}{C_{ox}} + U_T \cdot \ln \left(- \left(\frac{Q_m}{2} \right) \right) + U_T \cdot \ln \left(\frac{-\frac{Q_m}{2}}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} + \frac{1}{e \cdot n_i \cdot \frac{W_{Si}}{2}} \right) \quad (2.67)$$

Nous pouvons remarquer la similitude de ces deux relations. Par identification, nous établissons que la part de différence de travaux de sortie due au niveau de dopage du silicium est prise en compte par le terme $\frac{X}{C_{ox}}$. Nous pouvons également identifier la largeur physique W_{Si} du film de silicium non dopé à une expression dépendant du niveau de dopage du film dopé et de sa largeur physique.

Nous proposons ainsi de définir une largeur de semi-conducteur équivalente W_{eq} :

$$\frac{1}{\frac{W_{eq}}{2}} = \frac{1}{\int_{-W_{Si}/2}^0 e^{\frac{X}{W_{Si} \cdot \varepsilon_{Si} \cdot U_T} \cdot x^2} dx} - \frac{X}{\varepsilon_{Si} \cdot U_T} \quad (2.68)$$

Cette identification géométrique est l'élément central du développement de notre modélisation du transistor DG dopé. Elle permet d'établir une formulation unique de la relation entre la densité de charges mobiles et les potentiels, quel que soit le niveau de dopage du film de silicium.

Finalement, la relation (2.66) est réécrite :

$$\begin{aligned} V_G - \left(\Delta\phi + \frac{e \cdot N_a \cdot \frac{W_{Si}}{2}}{C_{ox}} \right) - V_{ch} \\ = \frac{\left(\frac{-Q_m}{2} \right)}{C_{ox}} + U_T \cdot \ln \left(- \left(\frac{Q_m}{2} \right) \right) + U_T \cdot \ln \left(\frac{\frac{-Q_m}{2}}{2 \cdot \varepsilon_{Si} \cdot e \cdot n_i \cdot U_T} + \frac{1}{e \cdot n_i \cdot \frac{W_{eq}}{2}} \right) \end{aligned} \quad (2.69)$$

Nous pouvons remarquer que lorsque la concentration du dopage est faible, la relation (2.68) donne :

$$\frac{W_{eq}}{2} \stackrel{\text{faiblement dope}}{\simeq} \int_{-W_{Si}/2}^0 dx = \frac{W_{Si}}{2} \quad (2.70)$$

impliquant que la largeur équivalente revient à la largeur physique du cas faiblement dopé. La signification de ce paramètre sera éclaircie plus tard.

2.3.2.3 Normalisation et développement de l'expression du courant

Puisque la relation (2.69) a la même forme que dans le cas non dopé, nous définissons les charges et potentiels normalisés de la même manière, soit $v = V/U_T$ et $q = Q/(4 \cdot C_{ox} \cdot U_T)$.

La normalisation de la relation entre potentiels et charges mobiles pour le DG MOSFET dopé aboutit à :

$$v_g - v_{ch} - \left(\Delta\phi + 2 \cdot q_d - \ln \left(\frac{q_{\text{int.eq}}}{2} \right) \right) = 4 \cdot \left(\frac{-q_m}{2} \right) + \ln \left(\frac{-q_m}{2} \right) + \ln \left(1 + \frac{-q_m}{2} \cdot \frac{C_{ox}}{C_{eq}} \right) \quad (2.71)$$

où $q_d = (e \cdot N_a \cdot W_{Si})/(4 \cdot C_{ox} \cdot U_T)$ et $q_{\text{int.eq}} = (e \cdot n_i \cdot W_{eq})/(4 \cdot C_{ox} \cdot U_T)$ sont respectivement la densité normalisée de charges de déplétion et la densité normalisée de charges intrinsèques équivalentes (en valeurs absolues). C_{eq} est la capacité équivalente du semi-conducteur donnée par $C_{eq} = \varepsilon_{Si}/W_{eq}$, qui dépend donc du niveau de dopage. La relation (2.71) révèle que la densité de charges mobiles est affectée par le dopage à travers deux termes impliquant chacun un décalage dans la tension de seuil. L'influence de ces termes est étudiée à la section 2.3.3.3. La formulation commune avec le cas non dopé permet l'utilisation de l'algorithme d'inversion numérique, proposée dans [16], pour rendre explicite la solution de la relation (2.71) et ainsi obtenir la densité de charge q_m en fonction du niveau de dopage.

Le courant est normalisé par le courant spécifique $I_{SP} = 4 \cdot \mu \cdot C_{ox} \cdot U_T^2 \cdot H_{Si}/L$ selon $i = I_D/I_{SP}$ [11]. La formulation commune de la relation charges-potentiels avec le cas non dopé aboutit directement à une expression du courant normalisé similaire à celle du transistor non

dopé, en considérant que la capacité du film de silicium est la capacité équivalente C_{eq} :

$$i = -q_m^2 + 2 \cdot q_m + 2 \cdot \frac{C_{eq}}{C_{ox}} \cdot \ln \left(1 - q_m \cdot \frac{C_{ox}}{2 \cdot C_{eq}} \right) \Big|_{q_{mS}}^{q_{mD}} \quad (2.72)$$

Il est intéressant de remarquer que la largeur équivalente influence le courant normalisé à travers C_{eq} et à travers la charge normalisée q_m obtenue par la relation (2.71). Ensuite, tous les résultats obtenus pour le DG MOSFET non dopé [11, 18, 20] reste valide pour le DG MOSFET dopé, étant donné que le travail de sortie est corrigé des effets du dopage [voir relation (2.66)] et que la largeur du semi-conducteur est remplacée par sa définition équivalente pour le dopé. Ceci est une conclusion vraiment importante puisqu'elle fournit une formulation cohérente des charges et du courant des DG MOSFET dopés et non dopés en fonctionnement symétrique.

2.3.3 Validation du modèle et discussion

Nous proposons maintenant d'évaluer la justesse du modèle en le mettant en regard avec des simulations TCAD. Dans le but de clairement identifier l'influence du dopage, nous choisissons de simuler des géométries pour lesquelles les SCEs sont négligeables. Nous considérons donc un DG MOSFET à canal long (1 μm) avec une épaisseur d'oxyde de silicium de 1.5 nm, et nous travaillons à une mobilité constante de 0.1 $\text{m}^2/\text{V}\cdot\text{s}$. La hauteur H_{Si} a également été fixée à 1 μm . La concentration du dopage N_a et la largeur de silicium W_{Si} sont les paramètres qui varieront tout au long des simulations.

2.3.3.1 Largeur équivalente

La Figure 2.5 montre la dépendance de la largeur équivalente de semi-conducteur en fonction de la concentration du dopage pour différentes largeurs de silicium.

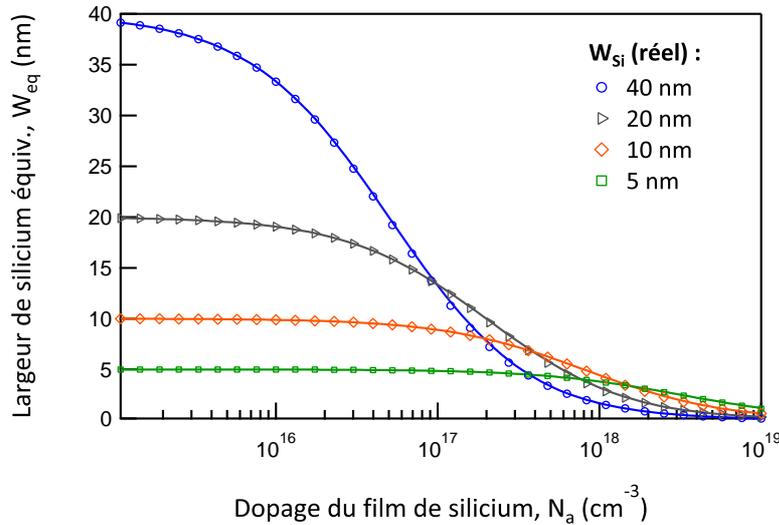


Figure 2.5 – Dépendance de la largeur équivalente selon la concentration du dopage pour les largeurs de silicium 10, 20 et 40 nm. Lignes : solution exacte [relation (2.68)]. Cercles : solution approchée [relation (2.73)].

Pour des couches de silicium relativement larges, la dépendance de la largeur équivalente est vraiment prononcée. Par exemple, pour une largeur de silicium de 40 nm, la largeur équivalente est d'environ 33 nm avec une concentration du dopage aussi basse que 10^{16} cm^{-3} . Elle atteint 13 nm pour un dopage de 10^{17} cm^{-3} . Nous pouvons faire confiance à ces calculs jusqu'à une concentration de 10^{18} cm^{-3} . En effet, le modèle peut être défaillant à donner un résultat correct pour de hauts niveaux de dopage, puisque dans ce cas les statistiques de Fermi-Dirac devraient être utilisées, et possiblement, le film de silicium peut aussi devenir partiellement déplété. Même si des travaux mentionnent que des effets de mécanique quantique (QME) sont présents dans des couches de silicium de moins de 15 nm de largeur [21, 31], nous ignorons ces effets dans notre développement du modèle. D'après la Figure 2.5, la largeur équivalente est plus petite pour des dispositifs larges que pour des étroits lorsque le dopage dépasse quelques 10^{17} cm^{-3} .

La relation (2.68) peut être simplifiée. L'exponentiel dans la relation (2.68) étant importante seulement pour de faibles valeurs de X , nous pouvons approximer (2.68) par son développement d'ordre 1 en séries de Taylor selon le terme X , aboutissant à :

$$W_{eq} \simeq \frac{2 \cdot W_{Si} \cdot \varepsilon_{Si} \cdot U_T \cdot (12 \cdot \varepsilon_{Si} \cdot U_T + X \cdot W_{Si})}{24 \cdot \varepsilon_{Si}^2 \cdot U_T^2 + 12 \cdot W_{Si} \cdot \varepsilon_{Si} \cdot U_T + X^2 \cdot W_{Si}^2} \quad (2.73)$$

Une comparaison entre l'expression exacte (2.68) et la simplifiée (2.73) de W_{eq} est montrée à la Figure 2.5. Nous observons que la relation (2.73) est une très bonne approximation de (2.68).

Enfin, il est important de garder à l'esprit que la largeur équivalente ne peut pas remplacer la largeur physique lorsque qu'on veut traiter de l'électrostatique. En effet, si la largeur équivalente était utilisée pour calculer le potentiel dans le silicium déplété, nous obtiendrions un résultat faux. Pour un potentiel de grille donné, Ψ_s est correctement évalué puisque la charge totale est bien prédite, mais ensuite, en utilisant (2.57) avec W_{eq} au lieu de W_{Si} , nous trouverions une valeur fautive pour Ψ_0 (le potentiel au centre du canal). Malgré tout, en faible inversion, Ψ_0 peut toujours être calculé à partir de la définition de C_1 dans (2.52) en utilisant (2.62) avec (2.69).

La largeur équivalente doit être uniquement utilisée lors du calcul des densités de charges mobiles globales et de paramètres associés.

2.3.3.2 Densité de charges mobiles et courant

La densité de charges mobiles et le courant obtenus par le modèle ont été comparés avec des simulations TCAD. La densité de charges mobiles a été évaluée dans les simulations à travers le courant qui parcourt le canal sous une tension $V_{DS} = 1 \text{ mV}$, condition pour considérer que la charge mobile est presque uniforme de la source au drain. Les mêmes paramètres physiques ont été utilisés pour les simulations TCAD et les simulations analytiques du modèle, c'est-à-dire qu'aucun paramètre additionnel d'ajustement n'a été introduit dans le modèle.

La Figure 2.6 représente la densité de charges mobiles, en valeur absolue, en fonction de la tension de grille pour différentes largeurs de silicium. La concentration du dopage est fixée à

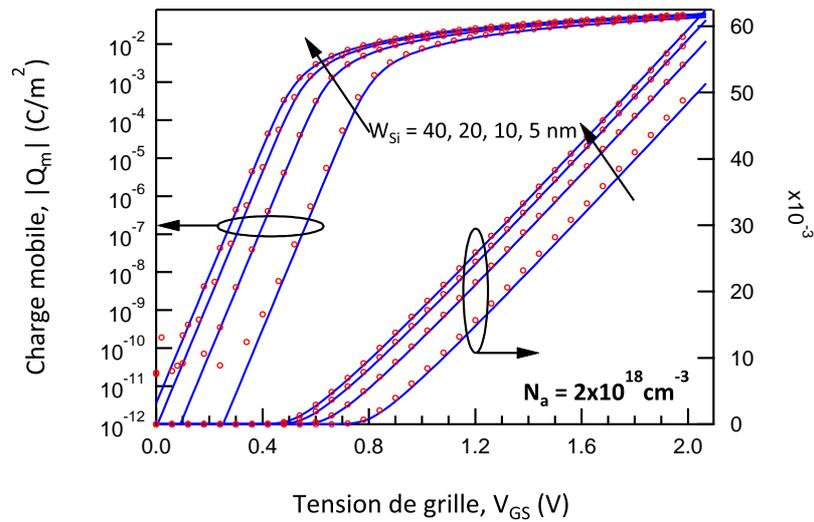


Figure 2.6 – Densité de charge mobile en fonction de la tension de grille pour différentes largeurs de silicium. La concentration du dopage a été fixée à $2 \times 10^{18} \text{ cm}^{-3}$. Ligne : modèle. Symboles : simulations TCAD.

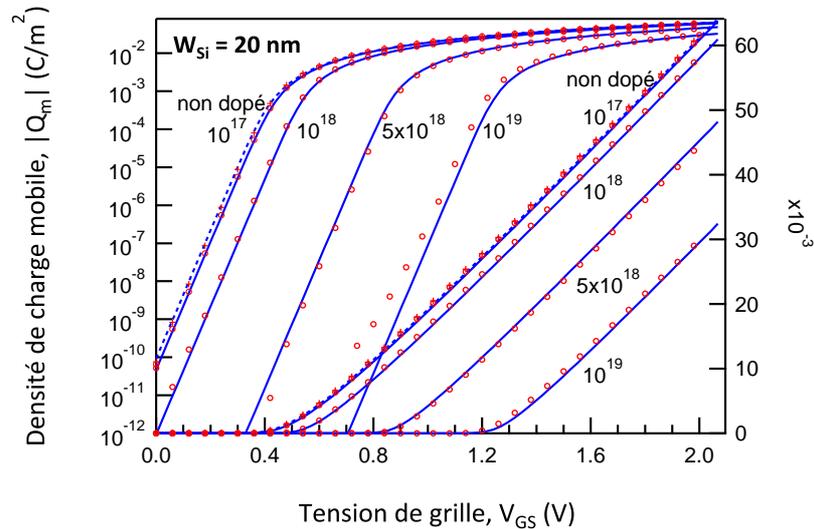


Figure 2.7 – Densité de charge mobile en fonction de la tension de grille pour différentes concentrations du dopage. La largeur de silicium a été fixée à 20 nm. Ligne : modèle. Symboles : simulations TCAD.

$2 \times 10^{18} \text{ cm}^{-3}$. Une bonne concordance entre le modèle et les simulations TCAD est obtenue de la faible à la forte inversion. Comme attendu de la dépendance de la largeur équivalente montrée à la Figure 2.5, plus le canal de silicium est étroit, et moins le dopage ne réduira la tension de seuil.

L'effet de la concentration du dopage est mise en évidence à la Figure 2.7. Pour une largeur donnée de silicium de 20 nm, le dopage varie d'un niveau non dopé à un niveau de 10^{19} cm^{-3} , ce qui est une valeur assez grande. De nouveau, l'accord entre le modèle et les simulations numériques est très bon. Nous pouvons remarquer la faible différence qu'il y a entre être non dopé et être dopé à 10^{17} cm^{-3} pour une largeur de silicium de 20 nm.

Lorsque le dopage excède 10^{19} cm^{-3} , le modèle est défaillant à donner une dépendance correcte des charges en fonction du potentiel, particulièrement en inversion faible et modérée.

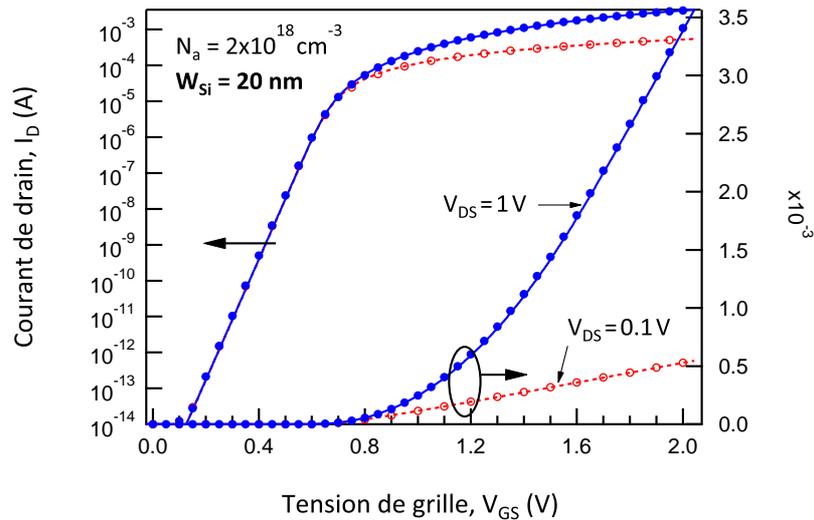


Figure 2.8 – Courant de drain en fonction de la tension de grille en régimes linéaire et saturé. La largeur de silicium a été fixée à 20 nm, et la concentration du dopage à $2 \times 10^{18} \text{ cm}^{-3}$. Ligne : modèle. Symboles : simulations TCAD.

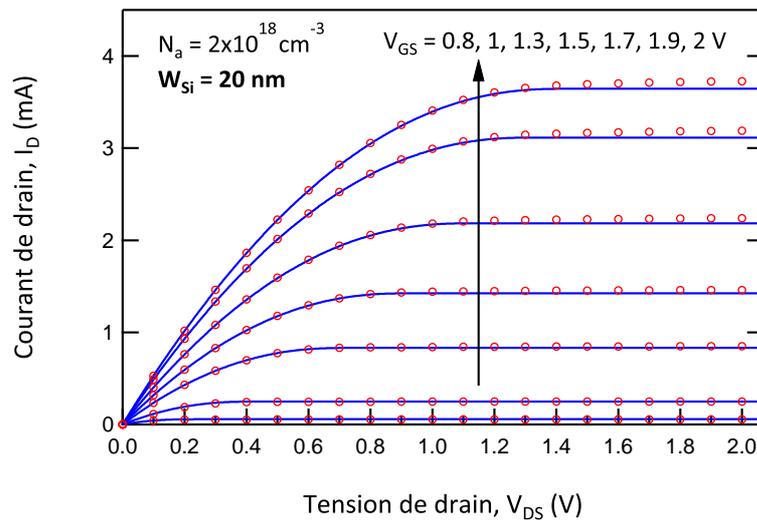


Figure 2.9 – Courant de drain en fonction de la tension de drain pour différentes tensions de grille. La largeur de silicium a été fixée à 20 nm, et la concentration du dopage à $2 \times 10^{18} \text{ cm}^{-3}$. Ligne : modèle. Symboles : simulations TCAD.

Des simulations numériques révèlent que cette situation correspond à une couche de silicium partiellement déplétée. Cela signifie que le DG MOSFET revient à deux MOSFET simple grille partiellement déplétés connectés en parallèle. Le modèle n'est plus adéquat lorsque le concept de DG disparaît.

La limite de validité de ce modèle a aussi été évaluée pour une largeur de silicium de 40 nm. Des écarts commencent à se produire pour des dopages supérieurs à $5 \times 10^{18} \text{ cm}^{-3}$. Clairement cette limite s'abaisse lorsque le silicium devient plus épais. A ce stade, il convient de mentionner que même si le modèle ignore la déplétion partielle [28], la densité de charges mobiles est bien prise en compte pour des densités de dopage excédant malgré tout celles rapportées dans [28].

Le courant en fonction de la tension de grille est montré à la Figure 2.8 pour une couche

de silicium de 20 nm ayant une concentration du dopage de $2 \times 10^{18} \text{ cm}^{-3}$. L'accord entre les simulations TCAD et le modèle est toujours bon, à la fois en régime linéaire ($V_{DS} = 100 \text{ mV}$) et en régime de saturation ($V_{DS} = 1 \text{ V}$), confirmant la justesse du modèle.

L'influence de la tension de drain est davantage visible à la Figure 2.9. Pour le courant en fonction de la tension de drain, l'accord est bon du régime linéaire au régime saturé. Cependant, des effets canaux courts ont lieu à fort V_{DS} à travers une modulation de longueur du canal, malgré notre considération initiale d'avoir choisi des dispositifs assez longs. Cela devient plus clair lorsque la tension de grille excède 1,5 V. Le courant de la simulation TCAD ne sature plus, et des discordances apparaissent avec la valeur constante prédite par le modèle. Cependant, même si les SCEs avaient été introduits dans le modèle analogique pour dispositifs non dopés [32], ces développements ne sont pas le sujet de cette partie. Finalement, même pour la plus grande concentration de dopage (10^{18} cm^{-3}) et avec une largeur de silicium (40 nm) utilisées dans [28], notre modèle est capable de prédire les densités de charges mobiles avec une grande précision.

2.3.3.3 Considérations de la tension de seuil dans le DG dopé

Intuitivement, doper le semi-conducteur modifiera aussi la tension de seuil, mais comme nous allons voir, cela requiert une attention particulière.

Différentes définitions de la tension de seuil existent. Ici, nous supposons que la tension de seuil peut être définie comme la tension de grille qui annule la densité de charges lorsque cette dernière est extrapolée depuis la forte inversion [11]. Cette définition est "grossière" mais elle est encore très utile. Cela mène à une simple relation, permettant de comprendre comment cette quantité est affectée par le dopage. Rappelons que, dans le DG MOSFET non dopé, la tension de seuil est exprimée par (2.44) :

$$V_{T_Non_dope} = \Delta\Phi - U_T \cdot \ln\left(\frac{q_{int}}{2}\right) \quad (2.74)$$

et en utilisant la même définition pour le cas dopé, nous avons :

$$V_{T_Dope} = \Delta\Phi + \frac{e \cdot N_a \cdot W_{Si}}{2 \cdot C_{ox}} - U_T \cdot \ln\left(\frac{q_{int,eq}}{2}\right) \quad (2.75)$$

Ensuite, la différence entre les tensions de seuil pour les DG dopé et non dopé ayant la même largeur physique, est simplement donnée par :

$$V_{T_Dope} - V_{T_Non_dope} = \frac{e \cdot N_a \cdot W_{Si}}{2 \cdot C_{ox}} + U_T \cdot \ln\left(\frac{W_{Si}}{W_{eq}}\right) \quad (2.76)$$

La relation (2.76) révèle que l'effet du dopage est double.

Le premier terme du côté droit de l'équation représente l'effet du dopage qui modifie V_T à travers la capacité d'oxyde de grille totale. Il s'agit d'un effet purement électrostatique induit par la densité de charges fixes dans le semi-conducteur à travers l'isolant de grille. La nature du

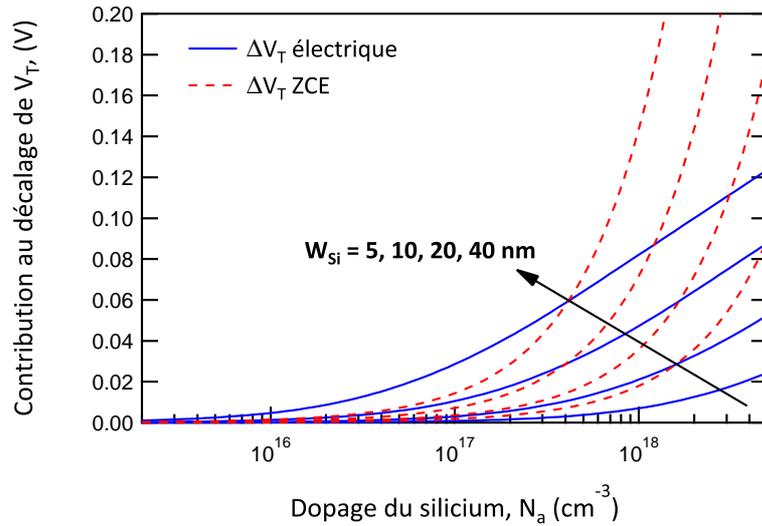


Figure 2.10 – Effet du dopage sur le décalage de tension de seuil entre des canaux de silicium non dopé et dopé pour différentes largeurs de silicium. Ligne pointillée : contribution de l’effet purement électrostatique, et ligne continue : contribution de l’effet lié au semi-conducteur.

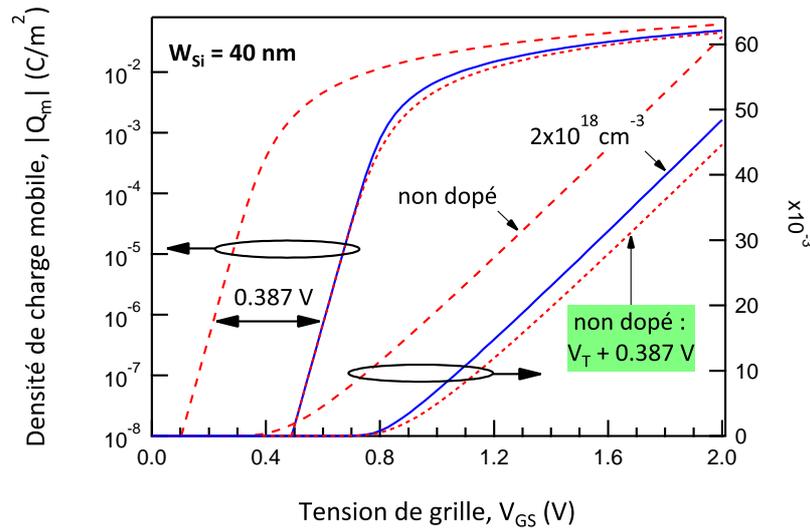


Figure 2.11 – Comparaison de la densité de charges mobiles entre des DG MOSFET dopé et non dopé. L’effet du dopage n’est pas assimilable à un décalage de tension de seuil du dispositif non dopé.

semi-conducteur n’a pas d’influence ici, et il n’y pas de dépendance explicite de la température (excepté la constante diélectrique de l’isolant qui change avec la température). Dans le second terme du côté droit de l’équation, l’autre effet du dopage est de changer la distribution de la charge à l’intérieur du semi-conducteur à travers le concept de largeur équivalente. Ici, les propriétés du semi-conducteur sont incluses dans la définition de la largeur équivalente ainsi que dans celle de la concentration équivalente de porteurs intrinsèques. La concentration intrinsèque physique n_i est donc impliquée elle aussi. Cette correction de tension de seuil dépendra fortement de la température (elle diminuera avec l’augmentation de la température), comme attendu par les termes liés aux statistiques du semi-conducteur. Ces deux différentes contributions sont tracées à la Figure 2.10 pour différentes largeurs de silicium. On peut voir que la contribution du terme “semi-conducteur” (lignes continues de la Figure 2.10) domine pour des couches rela-

tivement larges (au-dessus de 20 nm) et pour des densités de dopage inférieures à $4 \times 10^{17} \text{ cm}^{-3}$. Pour de plus grands dopages et de plus fines couches, cette contribution reste une grande part du décalage de tension de seuil. Ainsi, l'idée que le dopage modifie la tension de seuil selon un effet purement capacitif est trompeur.

Cependant, il est intéressant de voir si l'effet du dopage pourrait être équivalent à un simple décalage dans la tension de seuil, même si ce n'est pas strictement correct. La Figure 2.11 montre la charge mobile en fonction de la tension de grille pour un DG MOSFET non dopé et un dopé à $2 \times 10^{18} \text{ cm}^{-3}$, ayant tous les deux une largeur de leur couche de silicium de 40 nm. D'après la relation (2.75), la tension de seuil est en effet affectée par le dopage. Dans notre cas, sa valeur est d'environ 387 mV. Décaler les caractéristiques de courant du dispositif non dopé de cette quantité pour retrouver celles du dopé, semble être assez bon pour le régime de faible inversion. Mais manifestement, cela ne permet pas de donner le bon comportement en forte inversion. L'erreur serait pire si le décalage de tension de seuil permettait un bon accord en forte inversion. Dans ce cas, l'erreur en inversion faible et modérée serait très importante. Ainsi, l'effet du dopage ne peut pas être pris en compte comme une simple redéfinition de la tension de seuil en comparaison avec le dispositif non dopé. Le dopage affecte également la densité de charge dans un sens plus complexe, comme cela a été mis en évidence dans la section 2.3.2.3. Dans la relation charges-potentiels (2.71), la largeur équivalente influence la densité de charges en faible inversion à travers le terme logarithmique.

Cependant, pour de fines couches de silicium dopées modérément, une simple correction de la tension de seuil peut être suffisante. Néanmoins, Il faut faire attention à ne pas croire que cela est seulement dû à la correction "électrostatique", qui dans ce cas est celle qui contribue le moins (voir Figure 2.10).

2.3.3.4 Interprétation de la largeur équivalente

Le concept de largeur équivalente peut être compris par la forme du potentiel parabolique au sein du silicium déplété, une fois que la densité de charges mobiles devient négligeable (voir la relation (2.56)). Les corrections QME étant négligées, lorsque la tension de grille est augmentée, les charges mobiles ne peupleront pas toute la largeur du silicium. Contrairement à ce qui se passe dans les couches non dopées, l'inversion dite volumique n'a pas lieu. Ici, même en faible inversion, les porteurs seront toujours localisés aux interfaces entre le silicium et l'oxyde de silicium, là où le potentiel sera le plus bas.

Nous pouvons donner une explication qualitative de la largeur équivalente. Augmenter le niveau de dopage augmentera la densité de charges fixes, et par la même augmentera le champ électrique de surface (nous supposons ici la faible inversion, i.e. la densité de charges mobiles est négligeable). Cela signifie que la variation du potentiel de surface sera plus forte à l'interface Si/SiO₂. Par conséquent, la part de silicium correspondant à la région de plus fort potentiel électrique sera plus petite, et est de largeur W_{eq} . Ainsi la région de silicium peuplée de porteurs mobiles sera plus fine que le canal de silicium, ce qui illustre la signification de la largeur équivalente.

2.3.4 Conclusion de la modélisation du DG MOSFET dopé

Nous venons d'exposer comment un modèle pour le DG MOSFET dopé peut se ramener à celui du non dopé, à travers la définition d'une largeur équivalente de silicium, dépendant de la concentration du dopage. Ce travail a fait l'objet d'une publication dans une revue [33]. De multiples comparaisons avec des simulations numériques TCAD ont confirmé que cette approche est valide jusqu'à de relativement hauts niveaux de dopage, c'est-à-dire, jusqu'à ce que le silicium deviennent partiellement déplété. La base d'un formalisme cohérent pour des DG MOSFET symétriques dopés et non dopés est ainsi créée. De ces résultats, il ressort que la définition de la largeur équivalente, en conjonction avec des relations simples, peut être utilisée afin de simuler les charges et le courant dans des DG MOSFET dopés. Ainsi, du point de vue de la simulation de circuit, l'utilisation de ce modèle compact du transistor DG dopé ne demande pas plus de ressources de calcul qu'avec des transistors non dopés.

En perspective de ce travail, il conviendrait d'étudier la prise en compte des corrections des effets de mécanique quantique et des effets canaux courts dans le transistor DG. Cette approche de modélisation basée sur la définition d'une largeur équivalente peut être une base intéressante pour la modélisation compacte de dispositifs multigrilles dopés, tel que le transistor à effet de champ à grille enrobante GAA.

Conclusion

Nous avons présenté le formalisme de notre modélisation compacte du transistor MOSFET double grille faiblement dopé. Nous pouvons constater sa simplicité comparée aux modèles existant dans la littérature, notamment apportée par la normalisation des grandeurs électriques. La force de notre approche de modélisation est montrée par sa capacité à permettre de modéliser des dispositifs dopés de manière simple. Le concept de largeur équivalente développé dans la dernière partie de ce chapitre permet une simple extension du modèle non dopé pour modéliser des transistors double grille d'un faible à un fort niveau de dopage. Cette approche est basée sur la répartition dans la largeur du film de silicium des charges mobiles en régime sous-seuil. Cette technique de modélisation tout à fait intéressante sera par ailleurs réemployée dans la modélisation de transistors multigrilles au chapitre 4. Le modèle canal long du transistor MOSFET double grille est au coeur du travail de cette thèse et sera rappelé à plusieurs reprises dans les chapitres suivants de modélisation.

Références

- [1] Y. TAUR, "An analytical solution to a double-gate MOSFET with undoped body", *IEEE Electron Device Lett.*, vol. 21, n° 5, pp. 245–247, 2000.
- [2] Y. TAUR, X. LIANG, W. WANG et H. LU, "A continuous, analytic drain-current model for DG-MOSFETs", *IEEE Electron Device Lett.*, vol. 25, n° 2, pp. 107–109, 2004.
- [3] H. LU et Y. TAUR, "An analytic potential model for symmetric and asymmetric dg mosfets", *IEEE Trans. Electron Devices*, vol. 53, n° 5, 1161–1168, 2006.

- [4] H. LU, B. YU et Y. TAUR, “A unified charge model for symmetric double-gate and surrounding-gate MOSFETs”, *Solid-State Electron.*, vol. 52, n° 1, pp. 67–72, 2008.
- [5] B. YU, M. LU Huaxin Liu et Y. TAUR, “Explicit continuous models for double-gate and surrounding-gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 54, n° 10, pp. 2715–2722, 2007.
- [6] H. C. PAO et C. T. SAH, “Effects of diffusion current on characteristics of metal-oxide (insulator)-semiconductor transistors”, *Solid-State Electron.*, vol. 9, n° 10, pp. 927–937, 1966.
- [7] X. LIANG et Y. TAUR, “A 2-D analytical solution for SCEs in DG MOSFETs”, *IEEE Trans. Electron Devices*, vol. 51, n° 9, pp. 1385–1391, 2004.
- [8] J. SONG, B. YU et Y. YUAN Y. Taur, “A review on compact modeling of multiple-gate MOSFETs”, *IEEE Trans. Circuits and Systems*, vol. 56, n° 8, pp. 1858–1869, 2009.
- [9] G. D. J. SMIT, A. J. SCHOLTEN, G. CURATOLA, R. van LANGEVELDE, G. GILDENBLAT et D. B. M. KLAASSEN, “PSP-based scalable compact FinFET model”, dans *Proc. NTSI-Nanotech, WCM’07*, vol. 3, 2007, pp. 520–525.
- [10] O. MOLDOVAN, D. JIMENEZ, J. GUITART, F. A. CHAVES et B. INIGUEZ, “Explicit analytical charge and capacitance models of undoped double-gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 54, pp. 1718–1724, 2007.
- [11] J.-M. SALLESE, F. KRUMMENACHER, F. PRÉGALDINY, C. LALLEMENT, A. ROY et C. ENZ, “A design oriented charge-based current model for symmetric DG MOSFET and its correlation with the EKV formalism”, *Solid-State Electron.*, vol. 49, n° 3, pp. 485–489, 2005.
- [12] B. IÑÍGUEZ, D. JIMÉNEZ, J. ROIG, H. A. HAMID, L. F. MARSAL et J. PALLARÈS, “Explicit continuous model for long-channel undoped surrounding gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 52, n° 8, pp. 1868–1873, 2005.
- [13] F. LIME, B. IÑÍGUEZ et O. MOLDOVAN, “A quasi-two-dimensional compact drain-current model for undoped symmetric double-gate MOSFETs including short-channel effects”, *IEEE Trans. Electron Devices*, vol. 55, n° 6, pp. 1441–1448, 2008.
- [14] A. LAZARO, B. NAE, B. INIGUEZ, F. GARCIA, I. TIENDA-LUMA et A. GODOY, “A compact quantum model for fin-shaped field effect transistors valid from dc to high frequency and noise simulations”, *Journal of Applied Physics*, vol. 103, pp. 084 507–084 5079–, 2008.
- [15] J.-M. SALLESE, M. BUCHER, F. KRUMMENACHER et P. FAZAN, “Inversion charge linearization in MOSFET modeling and rigorous derivation of the EKV compact model”, *Solid-State Electron.*, vol. 47, n° 4, pp. 677–683, 2003.
- [16] F. PRÉGALDINY, F. KRUMMENACHER, B. DIAGNE, F. PÊCHEUX, J.-M. SALLESE et C. LALLEMENT, “Explicit modelling of the double-gate MOSFET with VHDL-AMS”, *Int. J. Numer. Model.*, vol. 19, n° 3, pp. 239–256, 2006.
- [17] M. TANG, “Etude et modélisation compacte du transistor FinFET”, thèse de doct., Université de Strasbourg, 2009.
- [18] O. MOLDOVAN, A. CERDEIRA, D. JIMENEZ, J.-P. RASKIN, V. KILCHYTSKA, D. FLANDRE, N. COLLAERT et B. INIGUEZ, “Compact model for highly-doped double-gate SOI MOSFETs targeting baseband analog applications”, *Solid-State Electron.*, vol. 51, n° 5, pp. 655–661, 2007.
- [19] A. ORTIZ-CONDE, F. J. GARCIA-SANCHEZ, J. MUCI, S. MALOBABIC et J. J. LIOU, “A review of core compact models for undoped double-gate SOI MOSFETs”, *IEEE Trans. Electron Devices*, vol. 54, n° 1, pp. 131–140, 2007.

-
- [20] B. DIAGNE, F. PRÉGALDINY, C. LALLEMENT, J.-M. SALLESE et F. KRUMMENACHER, “Explicit compact model for symmetric double-gate MOSFETs including solutions for small-geometry effects”, *Solid-State Electron.*, vol. 52, n° 1, pp. 99–106, 2008.
- [21] D. MUNTEANU, J. L. AUTRAN, X. LOUSSIER, S. HARRISON, R. CERUTTI et T. SKOTNICKI, “Quantum short-channel compact modelling of drain-current in double-gate MOSFET”, *Solid-State Electron.*, vol. 50, n° 4, pp. 680–686, 2006.
- [22] P. FRANCIS, A. TERAQ, D. FLANDRE et F. van de WIELE, “Modeling of ultrathin double-gate nMOS/SOI transistors”, *IEEE Trans. Electron Devices*, vol. 41, n° 5, pp. 715–720, 1994.
- [23] A. CERDEIRA, O. MOLDOVAN, B. IÑIGUEZ et M. ESTRADA, “Compact model for long channel symmetric doped DG MOSFETs”, dans *MOS-AK/ESSDERC/ESSCIRC Workshop*, Munchen, Germany, 2007.
- [24] A. CERDEIRA, B. INIGUEZ et M. ESTRADA, “Compact model for short channel symmetric doped double-gate MOSFETs”, *Solid-State Electron.*, vol. 52, n° 7, pp. 1064–1070, 2008.
- [25] A. CERDEIRA, O. MOLDOVAN, B. INIGUEZ et M. ESTRADA, “Modeling of potentials and threshold voltage for symmetric doped double-gate MOSFETs”, *Solid-State Electron.*, vol. 52, n° 5, pp. 830–837, 2008.
- [26] J. ALVARADO, B. IÑIGUEZ, M. ESTRADA, D. FLANDRE et A. CERDEIRA, “Implementation of the symmetric doped double gate MOSFET in Verilog-A for circuit simulation”, *Int. J. Numer. Model.*, vol. 23, n° 2, pp. 88–106, 2010.
- [27] L. ZHANG, J. HE, F. LIU, J. ZHANG et Y. SONG, “A unified charge-based model for symmetric DG MOSFETs valid for both heavily doped body and undoped channel”, dans *MIXDES Conf.*, Pozna, Poland, 2008, pp. 367–372.
- [28] F. LIU, J. HE, J. ZHANG, Y. CHEN et M. CHAN, “A non-charge-sheet analytical model for symmetric double-gate MOSFETs with smooth transition between partially and fully depleted operation modes”, *IEEE Trans. Electron Devices*, vol. 55, n° 12, pp. 3494–3502, 2008.
- [29] X. ZHOU, Z. ZHU, S. C. RUSTAGI, G. H. SEE, G. ZHU, S. LIN, C. WEI et G. H. LIM, “Rigorous surface-potential solution for undoped symmetric double-gate MOSFETs considering both electrons and holes at quasi non-equilibrium”, *IEEE Trans. Electron Devices*, vol. 55, n° 2, pp. 616–623, 2008.
- [30] Y. TAUR, “Analytic solutions of charge and capacitance in symmetric and asymmetric double-gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 48, n° 12, pp. 2861–2869, 2001.
- [31] J.-P. COLINGE, “Multiple-gate SOI MOSFETs”, *Solid-State Electron.*, vol. 48, n° 6, pp. 897–905, 2004.
- [32] M. TANG, F. PRÉGALDINY, C. LALLEMENT et J.-M. SALLESE, “Explicit compact model for ultranarrow body FinFETs”, *IEEE Trans. Electron Devices*, vol. 56, n° 7, pp. 1543–1547, 2009.
- [33] J.-M. SALLESE, N. CHEVILLON, F. PRÉGALDINY, C. LALLEMENT et B. IÑIGUEZ, “The equivalent-thickness concept for doped symmetric DG MOSFETs”, *IEEE Trans. Electron Devices*, vol. 57, n° 11, pp. 2917–2924, 2010.

Chapitre 3

Modélisation compacte du transistor FinFET ultime

Sommaire

3.1	Introduction	44
3.2	Le modèle FinFET ultime	46
3.2.1	Effets canaux courts	46
3.2.2	Dégradation de la mobilité	57
3.2.3	Région de saturation	58
3.2.4	Effets de mécanique quantique	59
3.2.5	Modèle de transcapacités	62
3.2.6	Discussion et résultats	66
	Conclusion	70
	Références	71

Ce chapitre est dédié à la modélisation compacte du transistor FinFET SOI. Sa structure est présentée à la Figure 3.1. Elle s'apparente à celle d'un MOSFET double grille disposé à la verticale. Une troisième grille surmonte la structure double grille, mais la constitution d'une épaisse couche d'oxyde de silicium entre cette grille et le film de silicium, permet d'ignorer son influence électrostatique sur le dispositif. Le but de cette structure est de fabriquer des transistors dont le comportement électrique est celui du transistor double grille, et ainsi pouvoir utiliser les modèles compacts existants de ce dispositif. En effet, la question de l'alignement des grilles du transistor double grille planaire ne se pose plus dans le cadre de la fabrication du transistor FinFET. Ainsi, nous considérerons tout au long de ce chapitre le transistor FinFET comme un transistor MOSFET double grille de largeur de film de silicium W_{Si} , de hauteur H_{Si} et de longueur de grille toujours égale à L . De plus le ratio (hauteur H_{Si}) / (largeur W_{Si}) doit être supérieure à une valeur seuil pour que le transistor FinFET puisse toujours être considéré comme un transistor double grille, sous peine de ne plus rendre négligeable l'influence de la grille supérieure.

La problématique de la modélisation du transistor FinFET ultime est ainsi celle de la

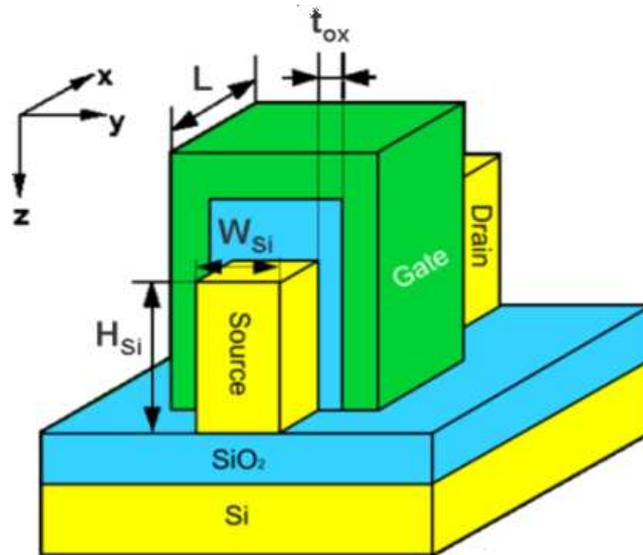


Figure 3.1 – Structure 3D du FinFET modélisé dans cette thèse

modélisation du transistor MOSFET double grille symétrique, telle qu'elle a été présentée au chapitre précédent. Le modèle compact « InESS-EPFL » du MOSFET double grille est le point de départ de notre travail de modélisation du transistor FinFET.

Nous introduisons ce chapitre par un rapide historique du développement des travaux de l'équipe jusqu'au commencement de cette thèse. Après le rappel des bases de la modélisation compacte du transistor FinFET de notre équipe, nous allons présenter les évolutions développées au cours de cette thèse sur le FinFET pour des dimensions ultimes. Nous disposons désormais d'une modélisation des effets canaux courts qui ne demande aucun paramètre empirique. Le modèle compact du FinFET tient également compte des effets de mécanique quantique selon l'axe normal aux deux grilles latérales sans paramètre empirique. Nous présenterons également une modélisation des transcapacités valide pour de petites géométries du transistor, indispensable aux simulations de circuits pour prendre en compte les temps de propagations des sous-circuits logiques. Davantage basée sur la physique, notre modélisation compacte du transistor FinFET ultime ne nécessite que peu de paramètres à extraire. Ce modèle prend en compte les aspects statique et dynamique du comportement du transistor FinFET.

3.1 Introduction

Dans notre équipe, un modèle compact du transistor FinFET pour de petites géométries existait suite aux travaux de deux précédents doctorants. Le modèle élaboré par Mingchung Tang [1], mon prédécesseur, prédisait le courant de drain de FinFETs de dimensions minimales $H_{Si} = 50$ nm, $W_{Si} = 3$ nm et $L = 25$ nm pour le film de silicium. Les effets canaux courts SCE et DIBL, et la dégradation de la pente sous-le-seuil était pris en compte dans le modèle compact, mais en se basant sur des expressions semi-empiriques qui demandaient 18 paramètres empiriques qu'il était nécessaire d'extraire. Pour les canaux étroits, de largeur W_{Si} inférieure à

15 nm, la quantification des charges mobiles était également prise en compte. Une expression semi-empirique intégrant 9 paramètres modélisait le décalage d'énergie de la première sous-bande du diagramme de bandes du canal induit par les effets de mécanique quantique. De plus, la dégradation de la mobilité n'était jusqu'alors pas encore considérée. Le modèle compact était validé à partir de résultats de simulations numériques TCAD 3D selon une mobilité constante. La modulation de longueur de canal induite par l'apparition d'une région de saturation au niveau du drain était modélisée, mais par des expressions qui n'étaient pas proche de la physique. D'autre part, un modèle dynamique existait, valable pour les canaux longs, mais de précision insuffisante pour les canaux courts. Ce modèle établissait toutes les transcaptivités du FinFET qui sont définies selon l'expression suivante :

$$C_{ij} = \begin{cases} -\frac{\partial Q_i}{\partial V_j} & \text{si } i \neq j \\ \frac{\partial Q_i}{\partial V_j} & \text{sinon} \end{cases} \quad \text{avec } i, j = G, D, S \quad (3.1)$$

où Q_i est la charge correspondant au i -ième terminal.

Suite à ces travaux qui ont permis une première modélisation du transistor FinFET jusqu'à des dimensions ultimes, mon objectif a été de trouver des solutions davantage basées sur la physique pour remplacer les modélisations semi-empiriques des effets canaux courts et de mécanique quantique, pour diminuer le nombre de paramètres à extraire du modèle. De plus, le modèle devait pouvoir tenir compte de la dégradation de la mobilité en forte inversion et nous devons étudier l'influence des canaux courts sur la mobilité quel que soit le régime de fonctionnement. Il était nécessaire de remplacer la modélisation de la modulation de longueur de canal. Enfin la modélisation des transcaptivités devait être valide pour les petites géométries. Le tableau 3.1 propose de comparer le nombre de paramètres du modèle compact entre sa version présentée dans cette thèse et celle élaboré en 2009 par le doctorant Mingchun Tang [1]. La modélisation des effets canaux courts et des effets de mécanique quantique (QME) ne nécessite désormais plus aucun paramètres empiriques.

Tableau 3.1 – *Comparaison du nombre de paramètres du modèle FinFET entre la version actuelle et la précédente.*

Effet modélisé	Modèle précédent [1]	Modèle actuel
Roll-off, DIBL, pente sous-le-seuil (SS)	18	0
Modulation de longueur de canal (CLM)	2	1
Mobilité	2	2
Quantification (QME)	9	0
Capacité de recouvrement	1	1
Total	32	4

3.2 Le modèle FinFET ultime

Le FinFET (Figure 3.1) est constitué d'une grande épaisseur d'oxyde de 50 nm sous la grille supérieure, et d'une épaisseur d'oxyde de 1,5 nm pour les deux grilles latérales. La section horizontale du FinFET est très similaire à la structure d'un MOSFET double grille conventionnel. Cela est particulièrement vrai pour les FinFETs très étroits qui ont une hauteur H_{Si} supérieure à 50 nm. La perte de contrôle électrostatique du canal par l'influence combinée des grilles latérales et de la grille supérieure aux angles supérieurs du canal, dit « effets de coin », peut être négligée si le film de silicium est peu dopé ou si H_{Si}/W_{Si} est très grand. Ainsi le coeur physique du modèle compact de FinFET peut être dérivé de celui d'un modèle de MOSFET double grille symétrique [2, 3].

Tout d'abord, la densité de charge de grille q_g est exprimée en terme de potentiels de canal et de grille (v_{ch} et v_g , respectivement) soit

$$v_g - v_{to} - v_{ch} = 4 \cdot q_g + \ln q_g + \ln(1 + \alpha \cdot q_g) \quad \text{avec } \alpha = \frac{C_{ox}}{C_{Si}} \quad (3.2)$$

où les potentiels sont normalisés selon $v = V/U_T$ (U_T est la tension thermodynamique) et les densités de charges selon $q = Q/(4C_{ox}U_T)$. C_{ox} est la capacité d'oxyde de grille ($= \varepsilon_{ox}/t_{ox}$), C_{Si} est la capacité de la couche de silicium ($= \varepsilon_{Si}/W_{Si}$) et v_{to} représente la tension de seuil canal long ($= \ln[(8C_{ox}U_T)/(en_iW_{Si})]$). L'équation (3.2) est résolue sans calcul itératif en utilisant l'algorithme d'inversion établi dans [3].

En définissant le courant de drain normalisé par $i = I_D/(4\mu C_{ox}U_T^2 H_{Si}/L)$, nous obtenons une relation entre le courant et les densités de charge source et drain, donnée par

$$i = -q_m^2 + 2 \cdot q_m + \frac{2}{\alpha} \cdot \ln \left(1 - \alpha \cdot \frac{q_m}{2} \right) \Big|_{q_{ms}}^{q_{md}} \quad (3.3)$$

où q_m est la densité de charge mobile ($= -2q_g$).

Les équations du corps du modèle seront référencées durant le développement du modèle compact statique et dynamique du FinFET. En particulier, afin d'inclure selon une méthode physique les effets canaux courts tout comme les effets de mécanique quantique dans le modèle complet, nous verrons plus tard que la connaissance et la compréhension de (3.3) est importante.

3.2.1 Effets canaux courts

Les effets canaux courts (SCEs) correspondent à une dégradation de la tension de seuil et de la pente sous-le-seuil lors de la réduction de la longueur des canaux du FinFET et l'augmentation de la tension source-drain. Nous avons étudié le potentiel électrique et la répartition des charges mobiles au sein de différentes géométries de canal afin de montrer l'origine physique des SCEs. Rappelons que le transistor FinFET est un transistor à canal faiblement dopé, le rendant facilement complètement déplété. La déplétion correspond au fait que les charges mobiles majoritaires du canal se recombinent avec les charges mobiles majoritaires des régions

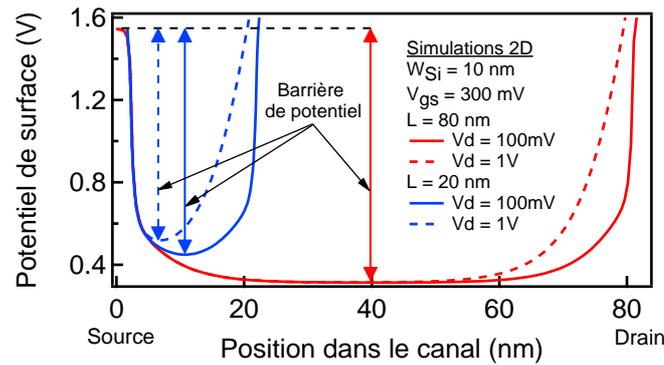


Figure 3.2 – Potentiel du canal de la source au drain pour les deux longueurs $L = 20$ nm et $L = 80$ nm, et pour $W_{Si} = 10$ nm. Simulations 2D.

source et drain qui sont elles en beaucoup plus grande concentration. Ainsi, la concentration des porteurs majoritaires et minoritaires du canal est intrinsèque, et les porteurs majoritaires des source et drain diffusent dans le canal (cf Figure 3.7). Dans la direction source-drain, le transistor FinFET est constitué de deux jonctions PN dont les régions les moins dopées sont communes, soit le film de silicium où le canal d'inversion apparaît. Cependant, étant donné les petites géométries de canal qui nous occupent et qui permettent d'atteindre un état de déplétion totale, les jonctions ne sont jamais à l'équilibre. Ainsi, le potentiel électrique du canal de source à drain est l'association des profils de potentiel des deux jonctions PN côte à côte, avec un potentiel qui atteint son minimum au centre du canal si le drain n'est pas polarisé. Néanmoins, pour les besoins de notre étude, nous nous proposons de décomposer le profil du canal en trois parties selon sa vitesse de variation selon l'axe entre source et drain, soit les deux profils de potentiel abrupts proches des jonctions à l'origine de la barrière de potentiel du canal, et le profil quasi-plat au centre du canal (cf Figure 3.2 pour $L = 80$ nm). Le terme « quasi-plat » dû à l'état de non-équilibre des jonctions sera par la suite remplacé par le terme « plat » pour des raisons de commodité. Pour des canaux de longueur inférieure à une certaine longueur, la région centrale à profil plat n'existe plus, et les transistors sont dit ultra-courts (cf Figure 3.2 pour $L = 20$ nm).

Nous allons maintenant nous attacher à étudier et modéliser les effets canaux courts tant sur le plan électrostatique que des densités de charges. Intéressons-nous tout d'abord à l'influence de la longueur du canal sur le profil de potentiel longitudinal.

3.2.1.1 Influence des canaux courts sur le potentiel électrostatique

Pour les plus courts canaux, la réduction de leur longueur produit, pour une largeur W_{Si} donnée, une modification de la tension de seuil. Cette dégradation est accentuée par l'effet de l'abaissement de la barrière induite par le drain (DIBL). En effet, la barrière de potentiel est la différence de potentiels entre celui de la source et le potentiel le plus faible du film de silicium qui est le potentiel minimum. Lorsque la longueur du film est suffisamment petite, l'augmentation du potentiel du drain entraîne avec lui une augmentation du potentiel minimum et donc une diminution de la barrière de potentiel. La simulation numérique du profil de potentiel du film

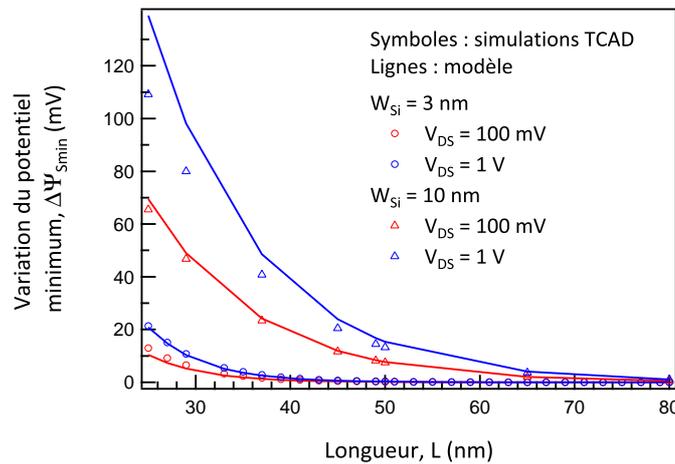


Figure 3.3 – *Potential minimum en fonction de la longueur du canal, pour une largeur $W_{Si} = 3$ nm et $W_{Si} = 10$ nm.*

de silicium en faible inversion, de la source au drain (Figure 3.2), montre l'influence possible à la fois de la longueur du film et de la tension source-drain sur la barrière de potentiel du transistor. Une bonne approximation du courant de drain en inversion volumique est de considérer uniquement la diffusion des porteurs libres, de la source vers le drain, qui est fonction de la barrière de potentiel. La réduction de la barrière implique l'augmentation du courant de drain. Ainsi, le roll-off de la tension de seuil (indépendamment de la réduction normale de la tension de seuil par la réduction de la longueur de grille) et l'effet DIBL ont une même origine physique, la variation du potentiel minimum pour une même polarisation de la grille. Nous avons extrait de simulations numériques la variation du potentiel minimum, noté $\Delta\Psi_{Smin}$, en fonction de la longueur de grille et pour deux tensions de drain, en prenant le potentiel minimum du canal long comme référence (Figure 3.3). Dans le cas des canaux ultra-courts, soit lorsque le potentiel minimum est modifié, le profil de potentiel le long du canal n'est plus « plat », même s'il n'y a pas de tension source-drain appliquée. Les profils de potentiel des jonctions source/canal et drain/canal constituent l'ensemble du profil de la source au drain. L'approximation du canal graduel ne peut donc plus s'appliquer. Le modèle dit canal long qui demande cette hypothèse n'est donc plus valable.

Modélisation

Nous proposons une solution pour nous permettre de continuer à utiliser le modèle canal long pour des dispositifs ultra-courts. Nous savons donc que le courant de drain en faible inversion est imposé par le potentiel minimum, quelle que soit la géométrie si nous considérons le transport de charge uniquement par dérive/diffusion. En faible inversion et pour un canal long non dopé, le potentiel minimum existe d'une interface Si/SiO₂ à l'autre du film de silicium et est imposé par la tension de grille à travers l'oxyde de grille. En effet, en étant placé en régime de faible inversion, la densité de charges du canal n'implique pas un champ électrique suffisant à travers la couche d'oxyde de grille pour avoir une différence de potentiel entre la grille et le canal. Le potentiel de surface peut être supposé égal à la tension de grille V_g [4]. La tension de

grille appliquée au transistor équivaut donc au potentiel minimum pour un canal long. Ainsi, une même variation du potentiel minimum dans le canal demande une même variation de la valeur de la tension de grille à utiliser dans le modèle compact. Cette relation est la clé de notre solution pour tenir compte des effets canaux courts dans notre modélisation du transistor FinFET. En effet, si nous pouvons prédire la valeur du potentiel minimum dans les dispositifs ultra-courts, en fonction notamment de la longueur du canal et de la tension de drain, nous pouvons directement l'utiliser dans le modèle compact canal long pour naturellement prédire le courant de drain en tenant compte des effets canaux courts. Ainsi en gardant la forme du modèle canal long (section 2.2), nous l'étendons pour le cas des canaux courts en remplaçant simplement V_g par V_{gN} défini comme suit :

$$V_{gN} = V_g + \Delta\Psi_{Smin} \quad (3.4)$$

Calcul de la variation du minimum de potentiel

Pour calculer $\Delta\Psi_{Smin}$ nous utilisons la solution exacte du profil de potentiel le long du canal définie dans [4]. Les auteurs de ce papier résolvent l'équation de Poisson en 2D avec l'hypothèse que la densité de charges mobiles dans le canal est négligeable. Il s'agit donc ici de la résolution du laplacien à deux dimensions en tenant compte des conditions aux limites que sont les potentiels des électrodes de grilles, source et drain. Cette solution du potentiel est ainsi uniquement valable lorsque le FinFET est en régime de faible inversion. Nous utilisons l'expression introduite dans [4] valide pour toutes les structures de telle sorte que $L/\lambda_G > 1,5$, où λ_G est la longueur d'échelle générale, donnée par :

$$\varepsilon_{Si} \cdot \tan(\pi \cdot t_{ox}/\lambda_G) = \varepsilon_{ox} \cdot \tan[\pi/2 - \pi \cdot W_{Si}/(2\lambda_G)] \quad (3.5)$$

Le ratio $r = L/\lambda_G$ reflète l'importance des effets canaux courts : plus ce ratio est petit, et plus les SCEs sont forts. Par exemple, pour $W_{Si} = 3$ nm et $L = 50$ nm, les SCEs sont négligeables étant donné le ratio de 5,5, alors que pour la même longueur mais avec $W_{Si} = 10$ nm, les effets DIBL et roll-off sont très importants avec dans ce cas un ratio $r = 2,8$. Sachant que $L/2$ est une localisation correcte du potentiel minimum le long du canal [5] et en utilisant l'expression du potentiel développée dans [4], nous obtenons la relation suivante pour $\Delta\Psi_{Smin}$:

$$\Delta\Psi_{Smin} = F_s \cdot [2 \cdot (V_{bi} - V_G) + V_D] \quad (3.6)$$

où V_{bi} est le potentiel des jonctions source/drain et F_s est une fonction issue de [4] que nous rappelons ici :

$$F_s = \cos\left(\frac{\pi \cdot W_{Si}}{2 \cdot \lambda_G}\right) \cdot \sinh(\pi \cdot r/2) \cdot \operatorname{csch}(\pi \cdot r) \cdot f_w \quad (3.7)$$

où

$$f_w = \frac{2 \cdot \lambda_G^2 \cdot \tan(\pi \cdot t_{ox}/\lambda_G) \cdot \sin[\pi \cdot W_{Si}/(2\lambda_G)]}{\pi^2 \cdot t_{ox} \cdot [\sin(\pi \cdot W_{Si}/\lambda_G) \cdot t_{ox} \cdot \operatorname{csc}(2 \cdot \pi \cdot t_{ox}/\lambda_G) + W_{Si}/2]} \quad (3.8)$$

Etant données les hypothèses de calcul de la solution analytique, nous rappelons que la

variation de potentiel minimum $\Delta\Psi_{Smin}$ calculée n'est valable qu'en faible inversion. Cependant, les valeurs de variation de potentiel issues de l'expression analytique pour des tensions de grilles en forte inversion, sont des valeurs négligeables comparées à la tension de grille (Figure 3.4). Ainsi, la relation 3.4 est considérée valide dans tous les régimes de fonctionnement. De plus, la modélisation analytique du potentiel minimum dépend de la tension V_{DS} entre source et drain. L'effet d'abaissement de barrière de potentiel induite par le drain (DIBL) est donc déjà prise en compte dans la relation 3.4.

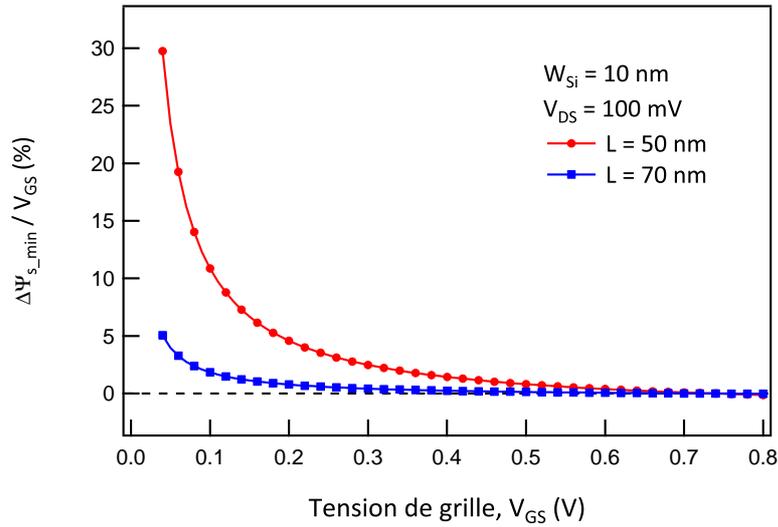


Figure 3.4 – Variation du potentiel minimum $\Delta\Psi_{Smin}$ par rapport à la tension de grille en pourcentage, en fonction de la tension de grille.

Maintenant que les formes normalisées des équations ci-dessus sont définies, nous sommes capables de calculer explicitement la densité de charge mobile par l'algorithme établi dans [3]. Les variables associées à la charge mobile q_m sont redéfinies comme suit :

$$q_m \equiv q_m(v_{gN}, v_{ch}, W_{Si}, L) \quad (3.9)$$

où

$$v_{gN} = v_g + \Delta\psi_{smin} \quad (3.10)$$

avec $\Delta\psi_{smin} = \Delta\Psi_{Smin}/U_T$.

Ainsi, en calculant la charge normalisée q_m en fonction de v_{gN} dans (3.3), nous conservons l'équation du courant de drain fonction de la charge mobile, qui est étendue aux dispositifs canaux courts en utilisant le modèle du potentiel minimum. Il est intéressant de remarquer qu'à travers cette nouvelle formulation plus physique des effets canaux courts, il est nullement nécessaire d'employer ni des paramètres empiriques, ni des fonctions mathématiques de lissage ou encore des calculs itératifs dépendant des potentiels appliqués au transistor. Ceci est une remarque très importante puisqu'il assure une formulation cohérente des charges et du courant dans des FinFET d'échelle très réduite et faiblement dopés. Nous apportons ici une amélioration majeure en comparaison du modèle de FinFET semi-empirique exposé dans [6]. Ce travail a été publié dans [7].

Dégradation de la pente sous-le-seuil

Un autre effet important relié à la réduction sévère de la longueur de canal du FinFET, est la dégradation de la pente sous-le-seuil. En définissant la pente sous-le-seuil comme la variation de tension de grille nécessaire à la réduction du courant sous-le-seuil d'une décade, et sachant que le courant sous-le-seuil dépend principalement de la hauteur de la barrière d'énergie du canal, le facteur de pente n est calculé au potentiel minimum au centre du canal (ψ_{cmin}) :

$$n = \frac{\partial v_g}{\partial \psi_{cmin}} = \frac{1}{1 - 2 \cdot \sinh(\pi \cdot r/2) \cdot \operatorname{csch}(\pi \cdot r) \cdot f_w} \quad (3.11)$$

La dégradation sous-le-seuil est prise en compte en terme de tension de pincement, soit :

$$v_p = \frac{v_g - v_{to}}{n} = \frac{v_{po}}{n} \quad (3.12)$$

où v_{po} correspond à la conventionnelle tension de pincement. Le facteur de pente donné par (3.11) est défini dans la région sous-le-seuil et ainsi (3.12) a une signification physique seulement en faible inversion. Par conséquent, en forte inversion le facteur de pente ne doit pas affecter la tension de pincement et la classique définition de la tension de pincement (v_{po}) peut être utilisée. De part le formalisme de notre modèle, et en particulier grâce à la forme de la relation du courant de drain (3.3), nous n'avons nullement besoin d'employer de fonctions de lissage pour décrire physiquement le comportement du dispositif. En effet, en faible inversion, seuls les termes linéaire et logarithmique du côté droit de (3.3) sont significatifs, et la substitution de v_{po} par v_p mène à une modélisation cohérente et précise de la charge et du courant normalisés.

Discussion

La Figure 3.5 compare le profil de potentiel de surface de source à drain de la solution analytique avec les résultats de simulations 3D d'un transistor FinFET, pour la géométrie $H_{Si}=50$ nm, $W_{Si} = 10$ nm, $L = 25$ nm. La comparaison est effectuée selon la coupe transversale du FinFET à mi-hauteur du canal. Les résultats du modèle analytique s'accordent bien à la simulation numérique au centre du canal, et comme nous pouvions nous s'y attendre, un écart existe près des jonctions, et ce d'autant plus que les régions source et drain sont fortement dopées. En effet, la densité de charges totales, fixes et mobiles, n'est pas du tout négligeable sur une courte distance dans le canal depuis les électrodes source et drain, comme cela est illustré par la courbe de charge à la Figure 3.5. Néanmoins, nous constatons que la grande différence de dopage n'a que très peu d'influence sur la détermination du potentiel minimum pour cette géométrie ultime prise ici comme cas limite. Comme le potentiel minimum est la grandeur que nous exploitons dans le modèle compact, la différence de dopage des électrodes a une influence négligeable sur les résultats de la modélisation des effets canaux courts du transistor FinFET. Cela est particulièrement visible pour un dopage des source et drain à une concentration de 5×10^{21} cm⁻³, la concentration utilisée pour toutes les simulations TCAD

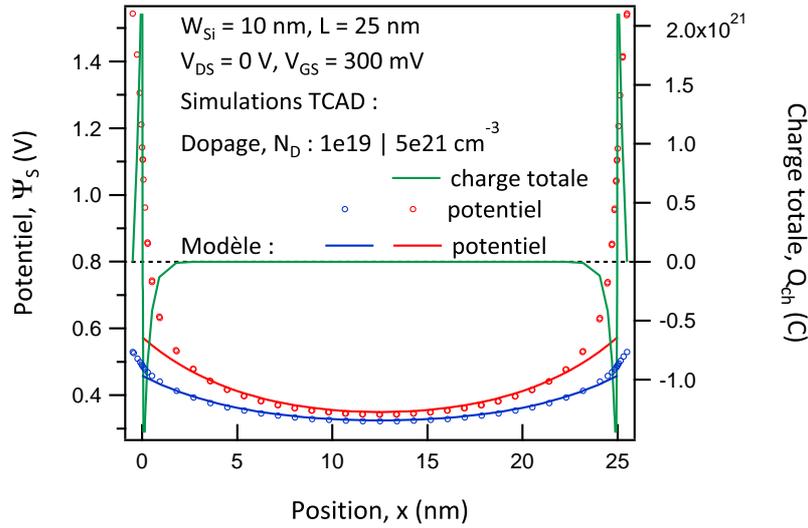


Figure 3.5 – Comparaison du profil de potentiel entre source et drain entre simulation et modèle analytique. Influence de la charge totale de la simulation sur le potentiel. Influence du niveau de dopage des régions source et drain.

2D et 3D des courants et des transcapacités de cette thèse. Cette différence flagrante pour cette extrême concentration s’explique simplement. La tension de diffusion des jonctions, ou tension de « built-in » V_{bi} , dans la simulation est en effet très différente de celle utilisée dans le calcul de la solution analytique. Le niveau de dopage est tel que le semi-conducteur est dégénéré, imposant d’autres lois de physique du solide que dans le cas non dégénéré, cas dans lequel est calculé la tension V_{bi} dans l’expression de la solution analytique. Ainsi, même avec des électrodes de source et drain constituées de semi-conducteur dégénérés, la modélisation de la variation du potentiel minimum par rapport au cas d’un transistor à canal long, par [4] est d’une qualité suffisante pour les besoins de la modélisation compacte du FinFET, comme le montre la Figure 3.3 comparant la variation du potentiel minimum en fonction de la longueur du canal entre la simulation TCAD et le modèle analytique.

Si nous observons la Figure 3.6 qui compare le roll-off de la tension de seuil issu des simula-

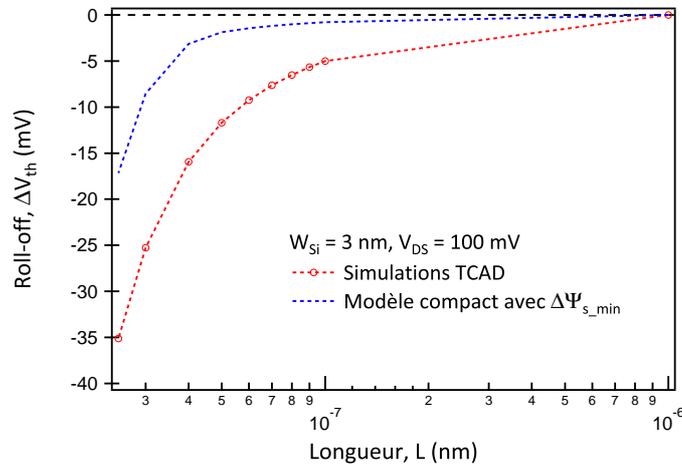


Figure 3.6 – Comparaison du roll-off des simulations TCAD 3D avec le roll-off du modèle compact ne prenant comme effet de petites géométries que la variation de potentiel minimum.

tions TCAD 3D avec celui du modèle compact qui prend en compte les effets canaux courts par la variation du potentiel minimum, nous comprenons que cette modélisation n'est pas suffisante pour correctement prendre en compte la dégradation de la tension de seuil. Ainsi, outre l'influence de la déformation du profil de potentiel dans les canaux ultra-courts maintenant prise en compte sans paramètres empiriques dans la modélisation compacte du transistor FinFET, il est intéressant de considérer le profil de concentration de la charge, de la source au drain, pour compléter la prise en compte des effets canaux courts.

3.2.1.2 Influence de la distribution de la densité de charges sur la longueur effective du canal

En observant la Figure 3.6, nous remarquons qu'une dégradation de tension de seuil existe à toutes les longueurs de canal. Cependant, le phénomène physique pris en compte par le modèle de variation de potentiel n'intervient que pour les canaux ultra-courts. Un autre effet physique que la variation du potentiel minimum induit donc une dégradation des performances du transistor en régime sous-le-seuil, et ce quelle que soit la longueur du canal.

Longueur effective du canal

L'observation du profil longitudinal de concentration de porteurs libres dans le canal à V_{DS} nulle (cf Figure 3.7) nous fait apparaître cet autre phénomène qui implique des effets canaux courts. La conduction électrique a pour support les porteurs de charge libres, et l'intensité de la conductivité dépend proportionnellement de la concentration de ces porteurs libres. Ainsi, le canal du point de vue électrique ne commence pas et ne se termine pas aux interfaces des régions P et N, mais nous le définissons plutôt par la région de conductivité minimale. En effet, les réservoirs de porteurs libres, que sont les régions de source et drain, s'étendent selon une diffusion sous la grille, rendant la longueur du canal effectif inférieure à celle de la grille. La

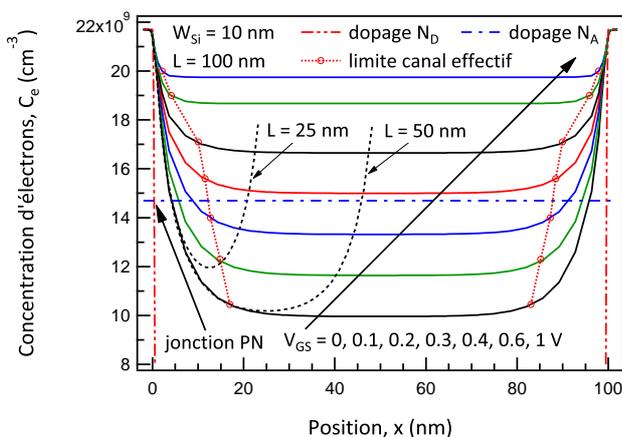


Figure 3.7 – Profil de source à drain de la concentration d'électrons libres de simulations TCAD 3D à $W_{Si} = 10$ nm, $L = 100$ nm et différentes tensions de grille (lignes continues). Profils de concentration du dopage. Délimitation du canal effectif en fonction de la tension de grille.

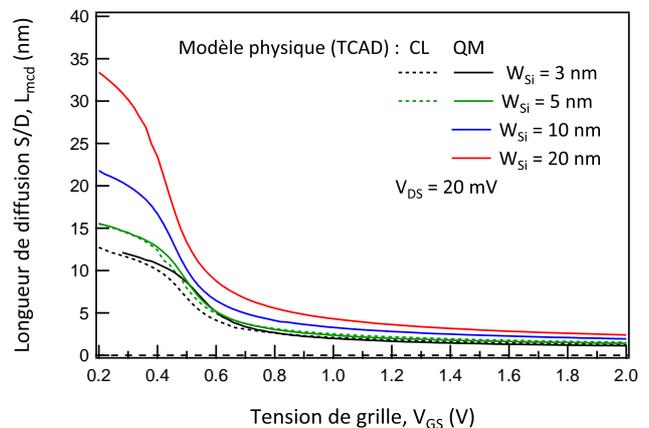


Figure 3.8 – Longueur de diffusion des porteurs majoritaires des source et drain dans le canal en fonction de la tension de grille, extraite de simulations TCAD 3D à différentes largeurs W_{Si} et selon un modèle physique classique et quantique.

Figure 3.7 nous montre par ailleurs que cette longueur effective du canal dépend de la tension de grille du transistor. L'augmentation de cette tension augmente la concentration de porteurs libres dans le canal, et la longueur de diffusion des porteurs libres de la source et du drain dans le canal est effectivement plus faible.

Selon cette constatation physique, la longueur du canal à utiliser dans le modèle compact est réduite par rapport à la longueur de grille du transistor, et ce quelle que soit cette longueur de grille. Ainsi l'erreur sur la longueur de canal demandée par le modèle compact est de plus en plus grande relativement à la longueur de grille lorsque celle-ci diminue. Et étant donné que le calcul du roll-off est basé sur la longueur de grille, la dégradation de tension de seuil doit augmenter progressivement avec la diminution de la longueur de grille. Il s'agit bien de ce qui est observé sur la courbe du roll-off des simulations TCAD.

Extraction de la longueur de diffusion des porteurs majoritaires dans le canal

Nous allons maintenant extraire la longueur de diffusion des porteurs majoritaires L_{MCD} dans le canal depuis les régions de source et de drain, afin de quantifier la longueur effective du canal. De nombreux papiers traitent de l'extraction de cette longueur effective [8]. Ils considèrent tous une longueur effective unique, extraite dans des conditions de très forte inversion, soit par exemple une tension de grille de 3 V dans [8]. Mais il n'est pas fait mention d'une longueur effective variant avec la tension de grille. Pour nous permettre l'étude et la quantification de la longueur de diffusion des porteurs de charge, et dans le même temps la longueur effective du canal, nous avons établi une méthode simple d'extraction valable quelle que soit la tension de grille.

Tout d'abord, nous considérons le canal complet du transistor comme trois résistances électriques en série (cf Figure 3.9). La résistance totale du canal s'écrit donc :

$$R_{totale} = 2 \cdot R_{extSD} + R_{ch} , \text{ avec } R_{extSD} = R_{extS} = R_{extD}$$

De plus, la concentration de charge mobile, et donc la conductivité électrique, dans le canal intrinsèque pouvant être considérée comme constante le long du canal, nous considérons que la résistance intrinsèque du canal R_{ch} est proportionnelle à la longueur effective du canal, soit

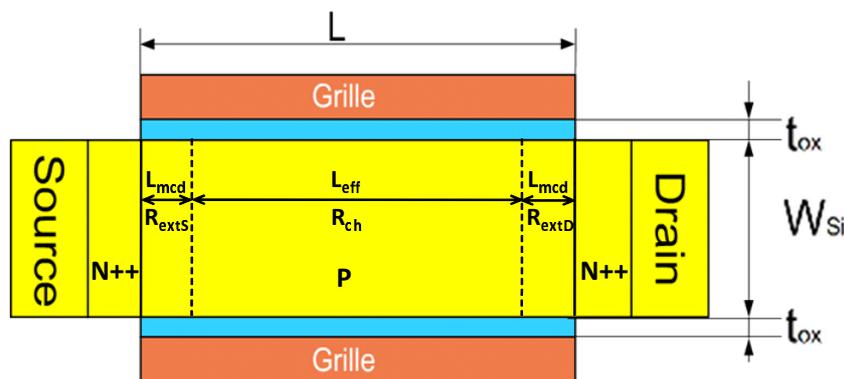


Figure 3.9 – Découpage du canal complet du FinFET selon une série de résistances.

$R_{ch} = a \cdot L_{eff} = a \cdot (L - 2 \cdot L_{MCD})$. Ainsi le courant de drain à une très faible tension source-drain peut simplement s'écrire :

$$I_D = \frac{V_{DS}}{2 \cdot R_{extSD} + a \cdot (L - 2 \cdot L_{MCD})}$$

Afin d'extraire la longueur L_{MCD} , nous considérons deux transistors FinFET de deux longueurs de grilles différentes L_1 et L_2 , et supérieures de quelques centaines de nanomètres des longueurs ultra-courtes, toute chose égale par ailleurs. De là nous exprimons le rapport des courants de drain des deux FinFET :

$$\frac{I_{D_1}}{I_{D_2}} = \frac{2 \cdot R_{extSD} + a \cdot (L_2 - 2 \cdot L_{MCD})}{2 \cdot R_{extSD} + a \cdot (L_1 - 2 \cdot L_{MCD})}$$

Les deux régions adjacentes de la source et du drain sont des régions de fortes concentrations de charges mobiles relativement à celle du canal. Ainsi, nous considérons que la résistance R_{extSD} est négligeable devant celle du canal étant donné les longueurs relativement longues des deux FinFET : avec $L_1 > L_2$, $R_{ch_1} > R_{ch_2} \gg R_{extSD}$. Ainsi, l'expression du rapport de courant de drain peut se simplifier en :

$$\frac{I_{D_1}}{I_{D_2}} = \frac{L_2 - 2 \cdot L_{MCD}}{L_1 - 2 \cdot L_{MCD}}$$

et la longueur de diffusion des porteurs majoritaires des régions source et drain L_{MCD} peut être exprimée en fonction de caractéristiques électriques :

$$L_{MCD}(V_{GS}, W_{Si}) = \frac{1}{2} \cdot \frac{I_{D_2}(V_{GS}, W_{Si}) \cdot L_2 - I_{D_1}(V_{GS}, W_{Si}) \cdot L_1}{I_{D_2}(V_{GS}, W_{Si}) - I_{D_1}(V_{GS}, W_{Si})} \quad (3.13)$$

Discussion

La figure 3.8 montre la longueur de diffusion des charges libres en fonction de la tension de grille pour différentes largeurs W_{Si} . Pour les largeurs de silicium inférieures à 10 nm, un effet de quantification des charges (présenté en section 3.2.4) apparaît en faible inversion et implique une augmentation de la tension de seuil du transistor. Ce décalage est également visible sur la Figure 3.8 entre le cas « classique (CL) » et le cas « quantique (QM) ». Les longueurs L_1 et L_2 utilisées pour l'extraction sont respectivement 400 nm et 200 nm, la tension V_{DS} est de 20 mV et les courants de drain sont issus de simulations TCAD 3D de FinFET. Nous remarquons que la longueur de diffusion est relativement constante en régime de forte inversion. Nous retrouvons dans ce cas la longueur effective habituellement extraite. Mais dès le passage en faible inversion, la longueur effective diminue rapidement et fortement. Les extrémités des longueurs de diffusion sont représentées sur la Figure 3.7 par les cercles rouges et ce pour différentes tensions de grilles. Nous y trouvons les longueurs de diffusion en adéquation avec leur réalité physique que nous pouvons voir sur les profils de concentration des charges mobiles. Les cercles encadrent la région du profil de concentration de charge relativement plate de manière similaire quelle que soit la tension de grille. De plus, nous comprenons l'origine de la diminution de la longueur de diffusion lorsque la tension de grille augmente. L'augmentation de la concentration de charge

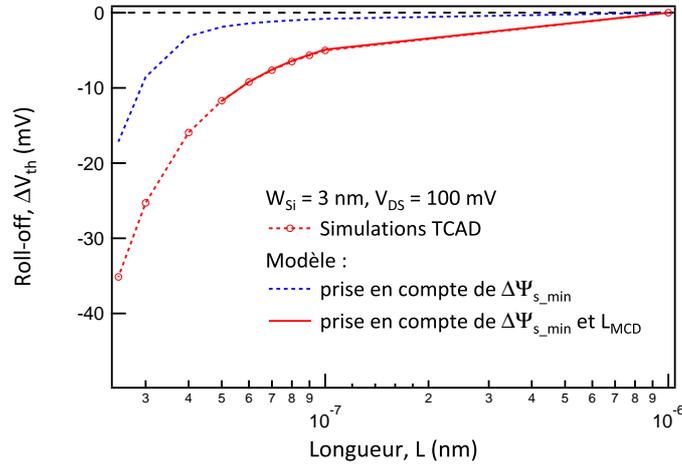


Figure 3.10 – Comparaison du roll-off des simulations TCAD 3D avec le roll-off du modèle compact pour différents effets physique pris en compte. La longueur de diffusion appliquée dans le modèle est de 6 nm étant données les conditions d'extraction du roll-off au niveau de la tension de grille.

de la partie centrale du canal modifie la condition aux limites de la diffusion dans le canal, et une distance plus faible est nécessaire pour que la concentration des régions source et drain atteigne celle du canal. L'effet canal court induit par cette modulation de longueur de canal a donc principalement lieu en faible inversion.

La tendance de la modulation de la longueur effective par la tension de grille est à même d'expliquer une part de la dégradation de la pente sous-le-seuil. En effet, en faible inversion, plus la tension de grille diminue, plus la longueur effective diminue, et plus le courant de drain a tendance à augmenter, impliquant une diminution moins rapide du courant de drain.

Cependant cette méthode présentée pour extraire la longueur effective n'a pas de sens pour les canaux ultra-courts. En effet, il n'existe plus dans le canal, la région centrale de concentrations de charge constantes (cf les profils de concentration partiels de la Figure 3.7 pour les longueurs $L = 25 \text{ nm}$ et $L = 50 \text{ nm}$). Pour de si courts canaux, nous remarquons que la longueur effective devient dépendante de la longueur de grille, selon une légère augmentation de cette longueur effective avec la réduction de la longueur de grille.

Validité

La prise en compte de la longueur de diffusion dans le modèle est validée à la Figure 3.10 à travers la courbe du roll-off de la tension de seuil pour la largeur W_{Si} de 3 nm . Étant donné que la tension de seuil, qui est extraite selon la méthode du courant constant, est autour de $0,5 \text{ V}$, la Figure 3.8 nous donne une longueur de diffusion de 6 nm , soit une longueur effective de $(L - 12 \text{ nm})$. La courbe de roll-off du modèle est en très bon accord avec celle des simulations TCAD sur la plage de longueur de grille non ultra-courte. Nous remarquons que cet effet de modulation de longueur de canal par la diffusion des charges mobiles provoque une part majeure de la dégradation de la tension de seuil lors de la réduction seule de la longueur de grille.

Perspectives

Nous sommes capables de déterminer par caractérisation la longueur de diffusion, valable

tant que le canal du transistor n'est pas ultra-court, mais par manque de temps nous n'avons pas pu étudier une modélisation analytique de cette longueur avec une dépendance de la tension de grille et de la largeur du canal. Un modèle de ce phénomène maintenant clairement identifié est à établir.

La longueur effective pour les canaux ultra-courts est plus difficile à visualiser en observant les profils de concentrations de charges mobiles. Cependant la diffusion des charges libres module néanmoins toujours la longueur du canal. Son étude qui reste à mener est plus délicate .

3.2.2 Dégradation de la mobilité

Il a été observé que lorsque le film de silicium est réduit à quelques nanomètres de largeur, la mobilité des électrons des couches d'inversions faible et modérée devient dépendante de la largeur W_{Si} [9]. La dégradation de la mobilité en faible inversion est particulièrement forte pour les plus faibles largeurs de silicium telles que 3 ou 5 nm. Ce comportement est traité dans [10] où il est décrit que la dispersion due aux phonons optiques de surface joue un rôle important dans le transport à champ faible pour les transistors SOI à canal ultra mince. Expérimentalement, il a été montré que la dégradation de la mobilité devient réellement significative pour des longueurs de canal inférieures à 100 nm [11].

Nous proposons un modèle simple pour prendre en compte la dégradation de la mobilité dans tous les régimes d'opération, que ce soit en inversions faible, modérée ou forte. Afin de modéliser la mobilité, nous nous basons sur le modèle de mobilité de Lombardi [12]. La dispersion due aux phonons acoustiques et optiques et à la rugosité de surface y est modélisée par une loi de Mathiessen. La dégradation de la mobilité due au champ électrique transverse en forte inversion ($E_{\perp SI}$) est modélisée par une approche conventionnelle basée sur la charge utilisant la charge de grille calculée dans notre modèle. De plus, nous considérons aussi un champ électrique transverse additionnel ($E_{\perp WI}$) résultant des effets canaux courts. Ce nouveau terme est significatif en inversions faible et modérée, et est négligeable en forte inversion. Ce modèle de mobilité est continu de la faible à la forte inversion et pour toutes les longueurs de canal, de l'ultra courte à la très longue.

Dans les simulations par éléments finis TCAD, le champ électrique local y est considéré, contrairement à ce qui est dans les modèles compacts où le courant ne peut pas être déduit de l'ensemble des champ électrique locaux. Nous calculons le champ électrique transverse moyen au milieu du canal ($E_{\perp WI}$). Selon l'approche détaillée dans la section 3.2.1.1, nous obtenons :

$$E_{\perp WI} = -\frac{\pi}{\lambda_G} \cdot \tan\left(\frac{\pi \cdot W_{Si}}{2 \cdot \lambda_G}\right) \cdot \Delta\Psi_{Smin} \quad (3.14)$$

Le champ électrique transverse en régime de forte inversion est obtenu à travers la charge de grille :

$$E_{\perp SI} = -Q_G/\varepsilon_{Si} \quad (3.15)$$

Par conséquent, la mobilité effective transverse est donnée par :

$$\mu_{\perp} = \frac{\mu_0}{1 + \frac{E_{\perp WI}}{e_0} + \frac{E_{\perp SI}}{e_1} + \frac{E_{\perp SI}^2}{e_2}} \quad (3.16)$$

où μ_0 est la mobilité à champ faible, $e_0 = -1 \text{ V}\cdot\text{nm}^{-1}$ est un facteur de normalisation, et, e_1 et e_2 sont des paramètres empiriques à extraire à partir de comparaisons avec des mesures ou des simulations TCAD.

Ensuite, la dépendance de la mobilité sur le champ électrique longitudinal est modélisée de manière précise par :

$$\mu = \frac{\mu_{\perp}}{[1 + (E_{\parallel}/E_c)^{\nu}]^{1/\nu}} \quad (3.17)$$

où $E_c = v_{sat}/\mu_{\perp}$ est le champ électrique à la vitesse de saturation (v_{sat} est la vitesse de saturation des porteurs), $E_{\parallel} = V_{deff}/L$ est le champ électrique latéral effectif moyen (V_{deff} sera donné plus tard) et ν est une constante ($=1.4$). Avec ce modèle, nous avons souhaité introduire le concept de longueur équivalente de canal L_{eq} basé sur la définition de la mobilité moyenne donnée dans (3.17). Puisque nous considérons une mobilité moyenne, il est approprié de définir L_{eq} comme suit :

$$L_{eq} = L \cdot [1 + (E_{\parallel}/E_c)^{\nu}]^{1/\nu} \quad (3.18)$$

Eq. (3.18) montre que les effets de limitation de la vitesse de diffusion des porteurs sont équivalents à l'augmentation de la longueur du canal du dispositif. L'implication de cette remarque pour le modèle sera discutée à la section suivante.

3.2.3 Région de saturation

Pour modéliser précisément le fonctionnement du dispositif en saturation, il convient de le prendre en compte selon l'effet de modulation de longueur de canal (CLM). La nature bi-dimensionnelle de la charge d'espace proche de la région de drain doit être considérée. Selon la définition usuelle, la CLM correspond à une réduction ΔL de la longueur de canal où ΔL est la longueur réelle de la région de saturation. Nous utilisons ici le terme « réel » puisque lorsque la dégradation de la mobilité est prise en compte, les effets de limitation de la vitesse de diffusion des porteurs sont équivalents à une augmentation de la longueur de canal du dispositif.

Tout d'abord, une expression précise de la tension de saturation de drain V_{Dsat} est nécessaire pour modéliser finement la CLM. Le modèle de tension de saturation de drain est basé sur l'approche proposée dans [5], que nous avons améliorée en terme de précision en lui ajoutant un paramètre empirique (θ) afin d'aider à l'ajustement du modèle aux simulations TCAD et aux mesures. De par notre formalisme basé sur la normalisation de toutes les quantités électriques, nous obtenons :

$$v_{dsat} = 2 \cdot q_{seff} \cdot \frac{v_{sat}}{\mu_{\perp} \cdot U_T \cdot q_{seff}/(\theta \cdot L) + v_{sat}} \quad (3.19)$$

avec

$$q_{\text{seff}} = -q_{ms} + \frac{v_{sat}}{v_{sat} - U_T \cdot \mu_{\perp} / (\theta \cdot L)} \quad (3.20)$$

où q_{ms} est la densité de charge mobile normalisée au niveau de la source, et v_{sat} est la vitesse de saturation des porteurs qui n'est pas ici considérée comme un paramètre empirique contrairement à beaucoup de modèles compacts ($= 1.2 \times 10^7$ cm/s). Pour lisser la transition entre les régions linéaire et de saturation, nous définissons une tension de drain effective $V_{D\text{eff}}$ qui est égale à la tension de drain dans la région linéaire et qui converge vers la tension de saturation $V_{D\text{sat}}$ dans le régime saturé [13]. Pour calculer ΔL , nous suivons l'expression développée dans [5] :

$$\Delta L = \lambda_W \cdot \ln \left(\frac{\psi_d + \sqrt{\psi_d^2 - \psi_{\text{sat}}^2 + (2 \cdot v_{sat} \cdot \lambda_W / (U_T \cdot \mu_{\perp}))^2}}{\psi_{\text{sat}} + 2 \cdot v_{sat} \cdot \lambda_W / (U_T \cdot \mu_{\perp})} \right) \quad (3.21)$$

avec

$$\lambda_W = (W_{\text{Si}}/2) \cdot \sqrt{2 \cdot C_{\text{Si}}/C_{\text{ox}} + 1/3} \quad (3.22)$$

où μ_{\perp} est la mobilité transverse calculée avec v_{deff} , ψ_d et ψ_{sat} sont les valeurs du potentiel de surface normalisé au niveau du drain et du point de saturation, respectivement donnée par :

$$\psi_d = v_d + q_{ms} - q_{m\text{deff}} \quad (3.23a)$$

$$\psi_{\text{sat}} = v_{\text{deff}} + q_{ms} - q_{m\text{deff}} \quad (3.23b)$$

où $q_{m\text{deff}}$ est la densité de charge mobile calculée avec v_{deff} . En prenant maintenant en compte les effets à la fois du champ électrique longitudinal et de la saturation de la vitesse, la longueur équivalente du canal L_{eq} définie dans (3.18) est modifiée en une longueur effective de canal L_{eff} donnée par :

$$L_{\text{eff}} = (L - \Delta L) \cdot \left[1 + \left(\frac{\mu_{\perp} \cdot v_{\text{deff}} \cdot U_T}{v_{sat} \cdot (L - \Delta L)} \right)^{\nu} \right]^{1/\nu} \quad (3.24)$$

Finalement, nous obtenons une expression compacte pour le courant de drain qui inclut les effets canaux courts, les effets de saturation et la dégradation de la mobilité. Selon la forme dénormalisée (c'est à dire en ampère), cela mène à la relation suivante :

$$I_D = i \cdot 4 \cdot \frac{\mu_{\perp}}{L_{\text{eff}}} \cdot C_{\text{ox}} \cdot U_T^2 \cdot H_{\text{Si}} \quad (3.25)$$

où le courant normalisé i est calculé à partir de (3.3) incluant les SCEs comme expliqué à la section 3.2.1.1

3.2.4 Effets de mécanique quantique

En forte inversion, un effet de quantification des porteurs de la couche d'inversion prend place dans le puits de potentiel défini entre la barrière d'oxyde de grille et la bande de conduction du silicium. Nous parlons alors de confinement électrique (Figure 3.11). Lorsque la largeur du film

de silicium est réduite à quelques nanomètres, un puits quantique est aussi formé entre les deux couches d'oxyde, soit un confinement structurel (Figure 3.11), en plus du confinement électrique due aux forts champs électriques transverses en forte inversion. Un modèle analytique unifié des effets de mécanique quantique (QMEs) pour des DG MOSFET a été développé dans [14]. Ce modèle considère des largeurs arbitraires de film de silicium, le couplage de charges entre les grilles avant et arrière, et le confinement du fort champ électrique transverse. Les résultats de ce modèle prédisent les QMEs, mais les équations développées sont très compliquées et vraiment non appropriées à la modélisation compacte. Un modèle du courant de drain du DG MOSFET prenant en compte de manière combinées la modélisation des QMEs et celle des effets canaux est présenté dans [15]. L'approche de modélisation des effets canaux courts est similaire à la nôtre. Le courant est calculé à partir d'une expression approchée de la distribution 2D du potentiel dans le canal. Cependant l'expression du courant de drain demande l'intégration de la solution du potentiel sur la largeur et la longueur du film de silicium, rendant le calcul du courant complexe. Concernant la modélisation des QMEs, les deux formes de confinements sont prises en compte dans le calcul de la charge d'inversion. Cependant ce calcul est implicite. Il intègre l'expression du potentiel qui elle-même dépend de la charge d'inversion. L'ensemble des décalages des bandes d'énergie du silicium est pris en compte dans ce modèle, ainsi que la distribution 2D du potentiel dépendante des polarisations de grille et de drain, mais son utilisation en conception et donc en simulation de circuits paraît difficile. Nous proposons une approche plus simple pour prendre en compte les QMEs avec les effets canaux, tout en étant entièrement basé sur des considérations physiques. Aucun paramètre empirique ne sera nécessaire. Dans notre développement actuel, nous nous basons sur le modèle physique présenté dans [16], à partir duquel nous réalisons des approximations raisonnables afin de développer un modèle compact quantique efficace en terme de temps de calcul.

Tout d'abord, il faut préciser que nous considérons que seule la première sous-bande est occupée par les porteurs. Cela est pleinement justifié pour les fines largeurs de canal en faible inversion, et est aussi une bonne approximation en forte inversion [17]. De plus, nous supposons,

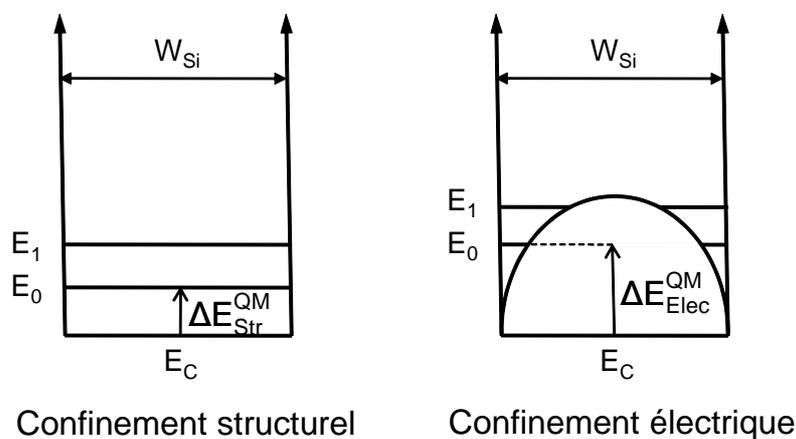


Figure 3.11 – Décalage du premier niveau d'énergie du diagramme de bandes du canal par le confinement structurel et électrique de ses charges.

en forte inversion seulement, que les QMEs peuvent être traités comme dans les dispositifs à une seule grille. Cela signifie dans le cas asymptotique d'un canal épais que la solution peut venir d'un modèle conventionnel utilisé pour les dispositifs à une seule grille [18]. Ensuite, nous proposons de modéliser le décalage quantique du premier niveau d'énergie comme le résultat de la contribution des deux types de confinement, structurel et électrique. Le décalage total de la première sous-bande est défini par :

$$\Delta E_1^{QM} = \frac{(\pi \cdot \hbar)^2}{2 \cdot m_x \cdot W_{Si}^2} + \frac{3 \cdot \hbar^2 \cdot b_1^2}{8 \cdot m_x} \quad (3.26)$$

où le premier terme du côté droit de l'équation est due au confinement structurel (pas de dépendance aux tensions), pendant que le second vient du confinement électrique en forte inversion. Dans (3.26), \hbar est la constante de Planck réduite, m_x est la masse effective des électrons dans la direction x et b_1 est défini comme dans [19] :

$$b_1 = \left(\frac{12 \cdot q^2 \cdot m_x \cdot \frac{n_{inv}}{3}}{\varepsilon_{Si} \cdot \hbar^2} \right)^{1/3} \quad (3.27)$$

où n_{inv} est la densité de charge d'inversion dans chaque canal. Ici, la charge de déplétion est négligée puisque nous supposons un FinFET faiblement dopé.

Le décalage quantique du premier niveau d'énergie ΔE_1^{QM} est finalement modélisé comme une correction du potentiel de surface ($=\Delta E_1^{QM}/q$). En normalisant par la tension thermodynamique, la correction de potentiel devient :

$$\Delta \psi^{QM} = \Delta \psi_1^{QM} + \Delta \psi_2^{QM} \quad (3.28)$$

où

$$\Delta \psi_1^{QM} = \frac{(\pi \cdot \hbar)^2}{2 \cdot q \cdot m_x \cdot W_{Si}^2 \cdot U_T} \quad (3.29)$$

et

$$\Delta \psi_2^{QM} = A^{QM} \cdot q_m^{2/3} \quad (3.30)$$

avec

$$A^{QM} = \frac{3}{2} \cdot \left[\frac{\hbar \cdot C_{ox}}{(q \cdot m_x \cdot U_T)^{1/2} \cdot \varepsilon_{Si}} \right]^{2/3} \quad (3.31)$$

Puisque (3.29) ne dépend pas des tensions aux terminaux du transistor, nous l'insérons simplement comme un décalage dans v_{gN} précédemment donné dans (3.10) :

$$v_{gN}^{QM} = v_{gN} - \Delta \psi_1^{QM} \quad (3.32)$$

$\Delta \psi_2^{QM}$ change la relation (3.2) entre les tensions et charges des terminaux comme suit :

$$v_{gN}^{QM} - v_{to} - v_{ch} = 4 \cdot q_g + \ln q_g + \ln(1 + \alpha \cdot q_g) + A^{QM} \cdot q_m^{2/3} \quad (3.33)$$

Par conséquent, pour calculer le courant de drain nous déterminer d'abord dv_{ch} de (3.33) et le substituons dans l'équation normalisée du courant donné par

$$i = - \int_{v_s}^{v_d} q_m dv_{ch} \quad (3.34)$$

Ainsi avec l'influence des effets de mécanique quantique, la relation (3.3) est réécrite :

$$i = -q_m^2 + 2 \cdot q_m + \frac{2}{\alpha} \cdot \ln \left(1 - \frac{\alpha \cdot q_m}{2} \right) + \frac{2}{5} \cdot A^{QM} \cdot q_m^{5/3} \Big|_{q_{ms}}^{q_{md}} \quad (3.35)$$

3.2.5 Modèle de transcapacités

En inversions modérée et forte, la contribution de la charge d'inversion aux capacités du dispositif est grande. Les transcapacités du transistor FinFET sont représentées à la Figure 3.12. La charge normalisée calculée à partir de deux grilles peut être obtenue en intégrant la densité de charge mobile normalisée sur la surface de la région de grille

$$q_G = -H_{Si} \cdot \int_0^L q_m dx \quad (3.36)$$

Suivant la méthode de partition des charges du canal proposée par Ward [20], la charge de drain est définie selon

$$q_D = H_{Si} \cdot \int_0^L q_m \cdot \frac{x}{L} dx \quad (3.37)$$

Les variables d'intégration dx et x peuvent être exprimées par l'intermédiaire du courant de drain et des tensions aux terminaux [21]. Les équations correspondantes pour q_G et q_D sont

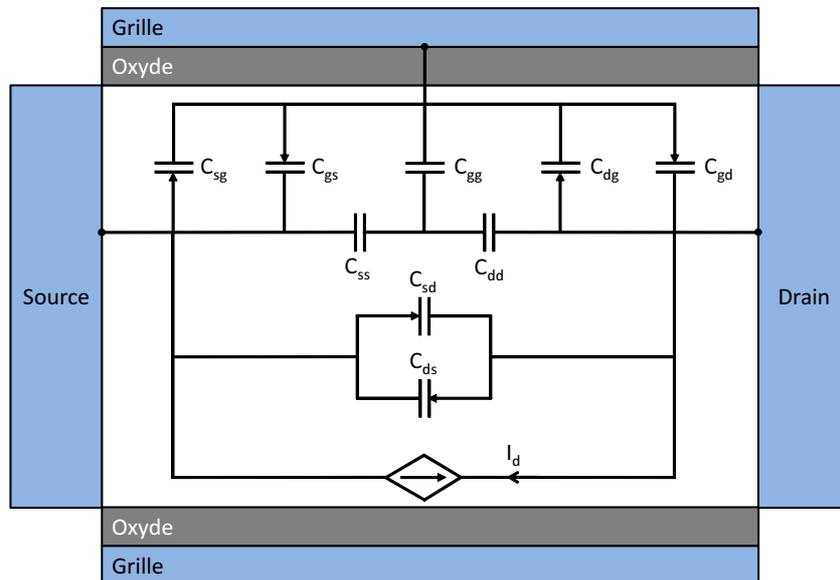


Figure 3.12 – Schéma représentant les 9 transcapacités du transistor FinFET.

développées dans [1] sur la base de notre modèle canal long et sont détaillées ici

$$q_G = \frac{H_{Si} \cdot L}{i} \cdot \left[\frac{2}{3} \cdot q_m^3 - q_m^2 - \frac{2}{\alpha} \cdot q_m - \frac{4}{\alpha^2} \cdot \ln \left(1 - \frac{\alpha}{2} \cdot q_m \right) \right] \Big|_{q_{ms}}^{q_{md}} \quad (3.38)$$

et

$$q_D = \frac{H_{Si} \cdot L}{i^2} (\text{Term1} - \text{Term2} \cdot \text{Term3})$$

où

$$\begin{aligned} \text{Term1} = & \frac{2}{5} \cdot q_m^5 - \frac{3}{2} \cdot q_m^4 + \left[\frac{4}{3} - \frac{2}{9\alpha} - \frac{4}{3\alpha} \cdot \ln \left(1 - \frac{\alpha}{2} \cdot q_m \right) \right] \cdot q_m^3 \\ & + \left[\frac{1}{\alpha} - \frac{2}{3\alpha^2} + \frac{2}{\alpha} \cdot \ln \left(1 - \frac{\alpha}{2} \cdot q_m \right) \right] \cdot q_m^2 \\ & + \left[-\frac{8}{3\alpha^3} + \frac{4}{\alpha^2} \cdot \ln \left(1 - \frac{\alpha}{2} \cdot q_m \right) \right] \cdot q_m \\ & + \frac{304}{9\alpha^4} - \frac{64}{5\alpha^5} - \frac{44}{3\alpha^3} - \frac{16}{3\alpha^4} \cdot \ln \left(1 - \frac{\alpha}{2} \cdot q_m \right) \\ & + \frac{4}{\alpha^3} \cdot \left[\ln \left(1 - \frac{\alpha}{2} \cdot q_m \right) \right]^2 \Big|_{q_{ms}}^{q_{md}} \end{aligned}$$

et

$$\text{Term2} = q_{ms}^2 - 2 \cdot q_{ms} - \frac{2}{\alpha} \ln \left(1 - \frac{\alpha}{2} \cdot q_{ms} \right)$$

et

$$\text{Term3} = \frac{i \cdot q_G}{H_{Si} \cdot L}$$

Les charges aux terminaux dépendent de la relation suivante

$$q_G + q_D + q_S = 0 \quad (3.39)$$

où $q_G > 0$ et $q_D, q_S < 0$. Ainsi, une fois que nous disposons des charges q_G et q_D , la charge de source s'obtient directement à partir de (3.39).

La définition des transcapacités (3.1) mène aux mêmes équations explicites que celles détaillées dans [22]. Nous rappelons ici la capacité normalisée c_{gg}

$$c_{gg} = \frac{H_{Si} \cdot L}{i} \cdot (q_{ms}^2 - q_{md}^2) - \frac{q_G}{i} \cdot (q_{md} - q_{ms}) \quad (3.40)$$

Pour inclure les effets de mécanique quantique dans les calculs des charges aux terminaux, la densité de charge mobile q_m doit être calculée en prenant en compte le confinement structural (cf. Section 3.2.4), et les variables dx et x dans (3.36) et (3.37) doivent être calculées à partir de (3.35) et (3.33). Cela rend les équations des charges des terminaux très compliquées, particulièrement pour q_D . Pour se départir de cette situation, nous appliquons la solution approximée, mais toujours juste, pour les effets de mécanique quantique en forte inversion [23].

Le développement de Taylor au premier ordre (3.30) nous permet d'écrire :

$$\Delta\psi_2^{QM} = A^{QM} \cdot \frac{1}{3} \cdot q_{mo}^{2/3} + A^{QM} \cdot \frac{2}{3} \cdot q_{mo}^{-1/3} \cdot q_m \quad (3.41)$$

où q_{mo} est une valeur particulière de q_m qui sera détaillée plus tard. Cette forme simplifiée de la correction quantique du potentiel de surface dans un fort champ électrique nous permet d'avoir les mêmes équations pour dx et x que dans le cas classique, mais avec une capacité équivalente d'oxyde définie selon

$$C_{ox}^* = \frac{C_{ox}}{1 + A^{QM} \cdot 2/3 \cdot q_{mo}^{-1/3}} \quad (3.42)$$

Notons que le premier terme dans (3.41) ne joue pas de rôle dans le calcul de la dérivée des expressions des charges des terminaux, étant donné que ce terme disparaît dans le calcul de la dérivée de dx par rapport à dv_{ch} . Pour la valeur particulière de q_{mo} , nous calculons la densité de charge mobile au niveau de la source en régime de forte inversion. Ainsi (3.42) n'a pas de dépendance avec W_{Si} . En établissant la valeur de q_{mo} à partir de la condition de forte inversion ($V_g = 1.5$ V) et de $V_{ch} = V_s$, nous réécrivons (3.42) comme : $C_{ox}^* = \alpha_{cp} \cdot C_{ox}$, où $\alpha_{cp} = 0.812$ est indépendant de la technologie et validé pour une grande gamme de largeurs de silicium (de 3 à 20 nm). Ainsi, pour établir les capacités en tenant compte des QMEs, il est nécessaire de calculer la densité de charge mobile par l'intermédiaire de l'emploi de (v_{gN}^{QM}) (cf. (3.32)), et de remplacer C_{ox} par C_{ox}^* dans les expressions des transcapacités.

Avec la réduction de la longueur de canal, le modèle des transcapacités devient très sensible à tous les effets canaux courts, dans les régimes d'inversions modérée et forte. De plus, pour des dispositifs à canal court, nous avons besoin de prendre en compte la capacité inter-électrode dans le régime sous-le-seuil [24]. D'après la loi de Gauss, la charge du couplage inter-électrode est donnée par le champ électrique perpendiculaire se terminant sur l'électrode choisie. Dans notre modèle, le champ électrique normal aux grilles en régime sous-le-seuil est calculé à partir du profil de potentiel que nous utilisons déjà à la section 3.2.1.1. La charge du couplage inter-électrode par unité de longueur du canal est calculée selon la loi de Gauss, et ensuite, par son intégration le long du canal, nous obtenons la charge totale de grille. Enfin, en prenant sa dérivée en fonction de la tension de grille, nous aboutissons à la capacité suivante :

$$C_{gg-in} = 4 \cdot H_{Si} \cdot \varepsilon_{Si} \cdot f_w \cdot \frac{\sin\left(\pi \cdot \frac{W_{Si}}{2 \cdot \lambda_G}\right)}{\sinh(\pi \cdot r)} \cdot [\cosh(\pi \cdot r) - 1] \quad (3.43)$$

Notons que C_{gg-in} est calculée en considérant la contribution des deux grilles. La capacité (3.43) est indépendante des tensions aux terminaux et dépend uniquement de la géométrie du dispositif. Les calculs correspondant montrent que le couplage des charges inter-électrode, respectivement sur les électrodes de drain et de source, sont : $C_{dg-in} = C_{sg-in} = -C_{gg-in}/2$.

En réalité, comme la charge est répartie le long du canal, la capacité est répartie de même, mais elle est modélisée comme étant une capacité localisée entre les deux terminaux. Ainsi, nous pouvons supposer que C_{gg-in} est concentrée sur une portion de la longueur du canal, que

nous définissons comme la longueur L_p

$$L_p = \frac{C_{gg-in}}{2 \cdot H_{Si} \cdot C_{ox}} \quad (3.44)$$

Ainsi, pour calculer la charge mobile du terminal et lui associer la capacité inter-électrode, nous devons remplacer L avec $(L - L_p)$, e.g. dans (3.38) et (3.40). Pour les canaux ultra-courts, il est nécessaire de considérer l'influence de l'effet de modulation de longueur de canal. Ainsi, la longueur du canal L doit être remplacée par L^* définie comme suit

$$L^* = L - \Delta L - L_p \quad (3.45)$$

où ΔL est définie dans (3.21).

Nous avons remarqué dans des simulations numériques que dans la région de saturation ΔL , la charge de couplage associée au fort champ électrique transversal au niveau du drain, a une influence significative sur les caractéristiques C-V en régime de saturation. Cette charge peut être calculée à partir de la dérivée au premier ordre du potentiel dans cette région. En négligeant la charge mobile dans la région de saturation, le profil de potentiel peut être obtenu de l'équation de Laplace. Pour résoudre cette équation, nous faisons une approximation en considérant que la limite de la région ΔL est déjà connue (définie à la section 3.2.3). Nous notons que les conditions aux limites en haut et au centre de l'aileron de silicium, tout comme au niveau du drain, sont les mêmes que celles en régime de faible inversion. Le potentiel à la limite de la région ΔL peut être supposé égal à v_{def} . Après avoir fait cette remarque, le profil de potentiel dans cette région peut être considéré être celui inversé du profil de potentiel que nous avons en faible inversion. Ensuite nous calculons le champ électrique transversal et sa charge de couplage associée. Enfin, le couplage capacitif de la région CLM C_{gg-in2} est défini comme suit

$$C_{gg-in2} = -4 \cdot H_{Si} \cdot \varepsilon_{Si} \cdot f_w \cdot \frac{\sin\left(\pi \cdot \frac{W_{Si}}{2 \cdot \lambda_G}\right)}{\sinh(\pi \cdot \Delta L / \lambda_G)} \cdot [\cosh(\pi \cdot \Delta L / \lambda_G) - 1] \quad (3.46)$$

De par notre approche, C_{gg-in2} a la même forme que C_{gg-in} avec un signe opposé, et avec une dépendance à ΔL au lieu de L . Cela signifie que le couplage capacitif dans la région de saturation dépend des tensions aux terminaux. Ce couplage disparaît en régime linéaire et est significatif en régime de saturation. Ce terme n'a pas de sens en faible inversion, et ainsi il est ajouté seulement dans l'équation de L_p dans (3.44).

Pour être encore plus précis, nous avons également besoin de prendre en compte la charge mobile dans cette région en la calculant comme suit

$$Q_{m-sat} = q_m(L, W_{Si}, v_g, v_{def}) \cdot 4 \cdot C_{ox} \cdot U_T \cdot \Delta L \cdot H_{Si} \quad (3.47)$$

En effet, la charge mobile tout comme celle de couplage en régime de saturation doivent être exprimée à partir de l'équation de Poisson 2D résolue de manière auto-cohérente avec le calcul de ΔL . L'approche que nous introduisons ici est cohérente avec la recherche d'une modélisation

dite compacte. Nous avons ensuite besoin de prendre en compte la capacité de recouvrement :

$$C_{ov} = 4 \cdot H_{Si} \cdot C_{ox} \cdot L_{ov} \quad (3.48)$$

où L_{ov} est la longueur de recouvrement de la grille sur les régions source/drain qui doit être extrait de simulations numériques à tension de grille nulle. De plus nous redéfinissons la région de canal L_p dans (3.45) comme

$$L_{p2} = \frac{C_{gg-in} + C_{gg-in2} + C_{m-sat}}{2 \cdot H_{Si} \cdot C_{ox}} + 2 \cdot L_{ov} \quad (3.49)$$

où C_{m-sat} est calculé comme la dérivée de $-Q_{m-sat}$ en fonction de la tension de grille. L'influence de ce terme sur les caractéristiques électriques du dispositif n'est pas significative, et ce terme peut dans la plupart des cas être ignoré pour éliminer la complexité. Finalement, la capacité totale est donnée par

$$C_{gg-TOT} = 4 \cdot C_{ox} \cdot c_{gg} + C_{gg-in} + C_{ov} \quad (3.50)$$

où c_{gg} est fonction des tensions aux terminaux, de W_{Si} et de L^* puisque C_{gg-in} dépend uniquement des dimensions géométriques W_{Si} et L .

3.2.6 Discussion et résultats

Le modèle analytique et explicite développé a été comparé avec des simulations Atlas 3D. Nous avons simulé des dispositifs à canal de type N disposant de grilles métal mid-gap qui impliquent que la différence de fonction de sortie entre les grilles en métal et le canal de silicium est négligeable. Les épaisseurs d'oxyde latéral t_{ox} sont égales à 1,5 nm, alors que l'épaisseur d'oxyde supérieur est de 50 nm. Le niveau de dopage du canal de silicium est $N_a = 5 \times 10^{14} \text{ cm}^{-3}$, afin de pouvoir négliger le dopage du canal. Le niveau de dopage des régions source et drain est très important ($N_d = 5 \times 10^{21} \text{ cm}^{-3}$) afin de négliger dans notre étude la résistance série. La plus petite hauteur de canal que nous avons simulé est $H_{Si} = 50 \text{ nm}$. Une grande gamme de largeur de film de silicium et de longueurs de canal a été considérée. Le modèle complet a été vérifié ici (tous effets inclus, à la fois dans le modèle analytique et les simulations TCAD). Pour les QMEs dans le simulateur TCAD, nous avons utilisé le modèle BQP où la masse effective isotrope est fixée à $m_x = 0.7 \cdot m_0$ (m_0 est la masse des électrons libres). Le modèle de mobilité utilisé dans les simulations Atlas est le modèle CVT. Soulignons que dans la suite, s'il n'y a pas d'indication spécifique, les simulations tiennent compte à la fois des QMEs et de la dégradation de la mobilité.

Pendant la validation du modèle analytique avec les simulations Atlas, nous avons remarqué qu'en utilisant la tension de pincement conventionnelle v_{po} sur toute la gamme de tension de grille, nous avons une pente sous-le-seuil très précise pour des canaux fins associés à différentes longueurs de canal dont le ratio est $r \geq 1.7$ (e.g., voir Figs. 3.13-3.14). Pour de grandes largeurs ($W_{Si} \geq 20 \text{ nm}$), la nouvelle tension de pincement v_p , celle corrigée par le facteur de pente n ,

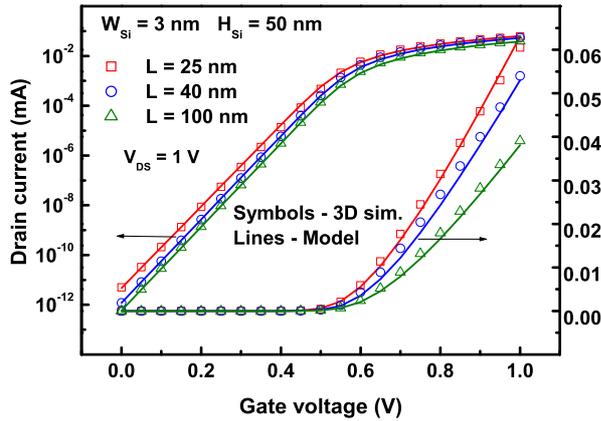


Figure 3.13 – Courant de drain en fonction de la tension de grille pour différentes longueurs de canal. Valeurs des paramètres du modèle : $\theta = 2.5$; pour $L = 25$ nm : $e_1 = 1.9 \times 10^7$ V/m, $e_2 = 150 \times 10^{14}$ (V/m)²; pour $L = 40$ nm : $e_1 = 3.2 \times 10^7$ V/m, $e_2 = 70 \times 10^{14}$ (V/m)²; pour $L = 100$ nm : $e_1 = 3.8 \times 10^7$ V/m, $e_2 = 200 \times 10^{14}$ (V/m)².

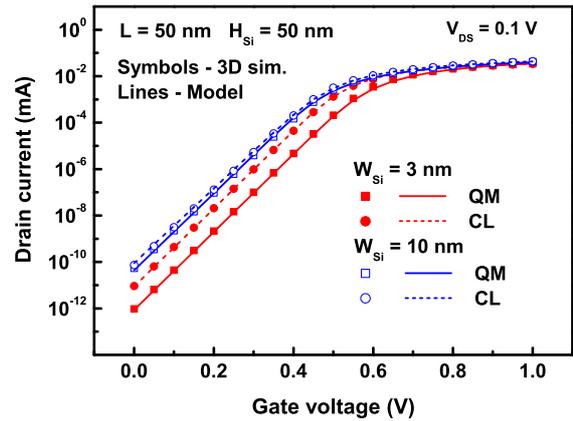


Figure 3.14 – Courant de drain en fonction de la tension de grille pour deux largeurs de canal avec à la fois les modèles classique et quantique. Valeurs des paramètres du modèle : $\theta = 2.5$, $e_1 = 3.3 \times 10^7$ V/m, $e_2 = 150 \times 10^{14}$ (V/m)².

doit être prise en compte. Cela rend évident que les SCEs et particulièrement la dégradation de la pente dans des larges canaux sont plus forts. Dans la simulation du modèle, la mobilité à champ faible (μ_0) est fixée à 1000 cm²/V.s.

Dans la Figure 3.13, les courbes $I_D - V_{GS}$ calculées à partir du modèle analytique pour $W_{Si} = 3$ nm et différentes longueurs de canal sont comparées avec des simulations numériques. Le courant de drain présenté sur l'échelle logarithmique (axe de gauche) montre la précision du modèle en régime de faible inversion. Comme cela peut être vu, la pente sous-le-seuil est bien décrite pour toutes les longueurs de canal. Le bon comportement en forte inversion indique que le modèle prend bien en compte l'effet CLM, la saturation de la vitesse tout comme les QMEs.

Pour montrer l'influence des QMEs en faible inversion, la Figure 3.14 présente la comparaison de deux structures ($W_{Si} = 3$ et 10 nm) selon les deux modèles classique et quantique. Comme cela était attendu, les effets de confinement structurel sont significatifs pour la plus fine structure et négligeables pour des structures telles que $W_{Si} \geq 10$ nm. En forte inversion, le courant de drain présente exactement le même comportement pour les deux largeurs de silicium. Cela met en évidence qu'en choisissant bien la valeur de q_{mo} en forte inversion, nous évitons la dépendance selon W_{Si} dans (3.42).

Le roll-off de la tension de seuil est défini comme le décalage de tension de seuil requis pour obtenir le même niveau de courant (en faible inversion) pour des dispositifs tant à canaux courts que longs ($L = 1$ μ m). La Figure 3.15 montre le roll-off de la tension de seuil due au SCEs pour différents W_{Si} . Le roll-off prédit par le nouveau modèle compact est en excellent accord avec les simulations numériques. Pour les mêmes structures, l'effet DIBL en fonction de la longueur de

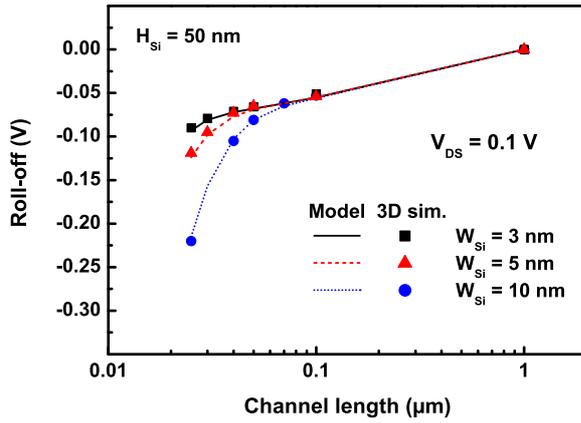


Figure 3.15 – Roll-off de la tension de seuil en fonction de la longueur de canal.

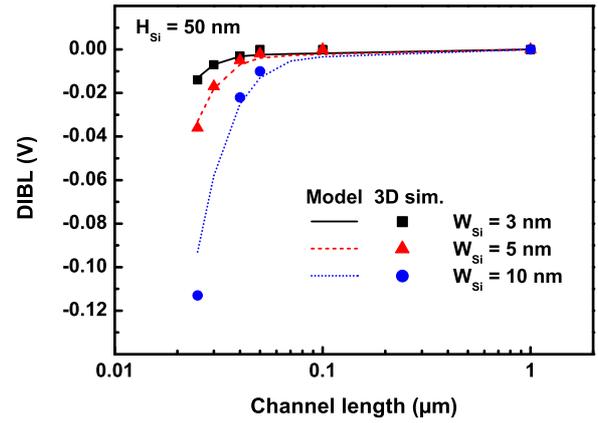


Figure 3.16 – DIBL en fonction de la longueur de canal.

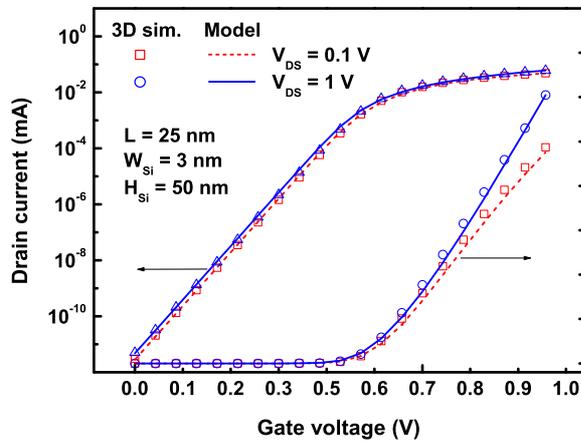


Figure 3.17 – Courant de drain en fonction de la tension de grille en régime linéaire et saturé.

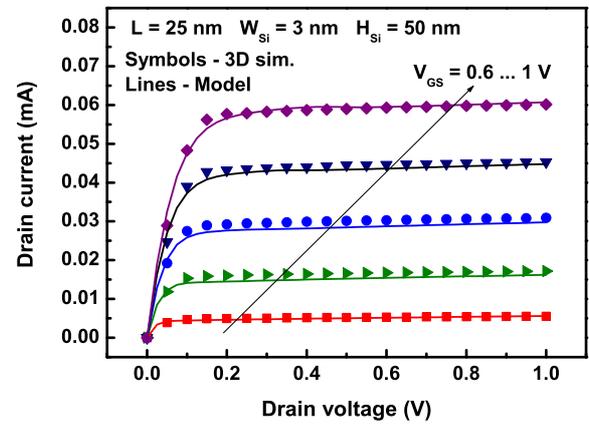


Figure 3.18 – Courant de drain en fonction de la tension de drain en régime linéaire et saturé.

canal est présenté à la Figure 3.16. Le DIBL est calculé comme la différence entre les tensions de grille requises pour obtenir le même courant à $V_{DS} = 0.1$ et 1 V. La comparaison avec les simulations numériques montre le nouveau modèle prédit bien les SCEs jusqu'au ratio $r \simeq 1.7$.

Les Figures 3.17 et 3.18 présentent le courant de drain pour un dispositif de très faibles dimensions avec $W_{Si} = 3$ nm et $L = 25$ nm. Nous pouvons voir que le modèle analytique simule très bien le courant de drain tant en région linéaire que de saturation. Ce bon accord des simulations 3D avec le modèle en régime de saturation provient de la modélisation précise des QMEs, de la CLM, de la dégradation de la mobilité et de la tension de saturation. La conductance de sortie g_{ds} calculée pour cette structure est montrée à la Figure 3.19. Comme nous pouvons le voir, le modèle compact est en bon accord avec les caractéristiques quasi-statiques. La transconductance g_m calculée pour le même dispositif est présentée à la Figure 3.20. Pour la géométrie des Figures 3.17 à 3.20, nous fixons les paramètres du modèle aux valeurs suivantes : $\theta = 3.0$, $e_1 = 2.3 \times 10^7$ V/m et $e_2 = 120 \times 10^{14}$ (V/m)². La légère différence en forte inversion entre

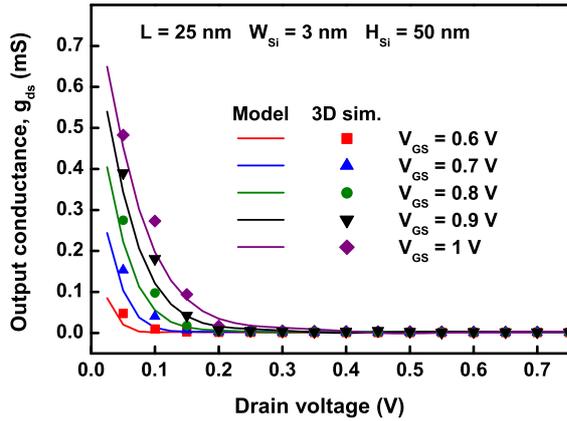


Figure 3.19 – Conductance de sortie en fonction de la tension de drain.

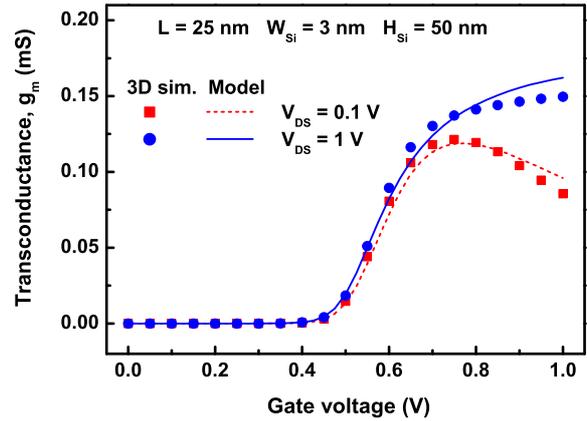


Figure 3.20 – Transconductance en fonction de la tension de grille.

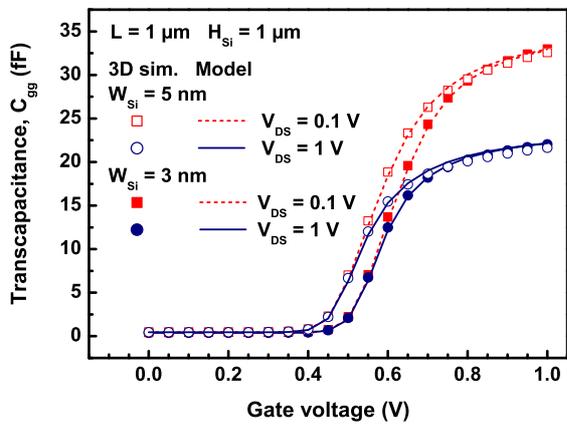


Figure 3.21 – Capacité C_{gg} obtenue du modèle et des simulations pour deux différents W_{Si} . $L_{ov} = 2$ nm.

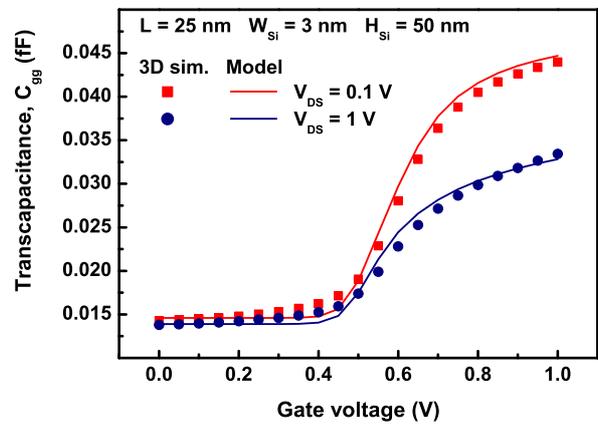


Figure 3.22 – Capacité C_{gg} obtenue du modèle et des simulations pour $W_{Si} = 3$ nm et $L = 25$ nm. $L_{ov} = 0.8$ nm.

le modèle et les simulations numériques provient du couplage entre le modèle de dégradation de la mobilité et celui des QMEs. Une optimisation des valeurs des paramètres du modèle de mobilité doit permettre d'améliorer les résultats.

Maintenant, afin de nous concentrer sur la validation de la modélisation des effets quasi-statique, nous maintenons la mobilité constante dans le modèle analytique et les simulations 3D. Nous validons tout d'abord le modèle de capacités pour un FinFET canal long. La Figure 3.21 montre une comparaison de la capacité de grille C_{gg} entre celle modélisée et celle simulée en 3D, pour $W_{Si} = 3$ et 5 nm, et avec $L = 1$ μ m. Les QMEs sont pris en compte dans le modèle par l'intermédiaire de C_{ox}^* . Nous pouvons voir que les résultats du modèle s'accordent bien avec les simulations 3D tant en faible qu'en fort champs latéraux.

La capacité C_{gg} pour un dispositif de dimensions très réduites est montrée à la Figure 3.22. Etant donné le bon accord du modèle avec des simulations 3D, nous considérons les QMEs et les SCEs bien modélisés. C_{gg} dans le modèle est déterminée par (3.50) qui prend en compte les

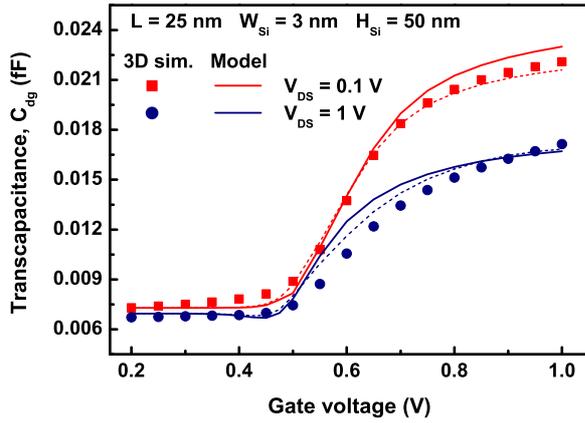


Figure 3.23 – Capacité C_{dg} obtenue du modèle et des simulations pour $W_{Si} = 3$ nm et $L = 25$ nm. $L_{ov} = 0.8$ nm.

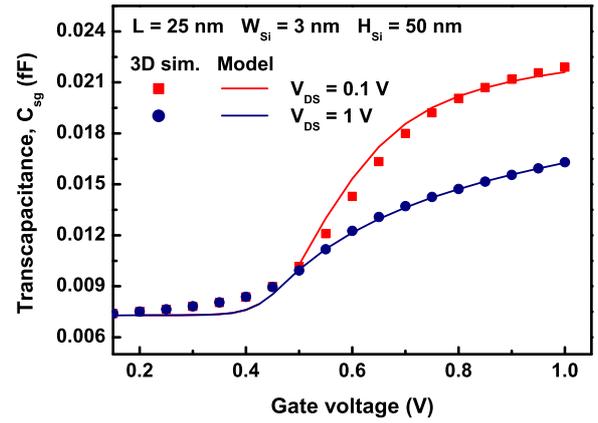


Figure 3.24 – Capacité C_{sg} obtenue du modèle et des simulations pour $W_{Si} = 3$ nm et $L = 25$ nm. $L_{ov} = 0.8$ nm.

QMEs tels que décrit précédemment. Il faut noter que les QMEs ne sont pas inclus dans les capacités parasites (second et troisième termes de (3.50)). La capacité C_{m-sat} dans (3.49) est négligée dans le tracé des courbes de cette figure.

La capacité entre drain et grille C_{dg} est montrée en Figure 3.23. La charge de drain est très sensible à tous les effets que nous décrivons dans la région de saturation. Les lignes pleines Figure 3.23 sont obtenues en négligeant le terme C_{m-sat} dans (3.49), alors que les lignes pointillées prennent en compte la correction de ce terme. Pour obtenir des équations explicites, nous négligeons la dépendance de ΔL aux tensions lors du calcul des dérivées des charges aux terminaux par rapport aux tensions des terminaux, et ce qui entraîne des erreurs visibles dans la région de saturation comme le montre la Figure 3.23.

Pour finir, la capacité entre source et grille C_{sg} est montrée en Figure 3.24. Comme la charge de source est principalement définie proche de la source, cela n'avait aucun sens d'inclure C_{gg-in2} et C_{m-sat} dans le modèle analytique. Ainsi, les résultats de la Figure 3.24 sont obtenus en calculant L_p à partir de (3.44) et en incluant la correction de la longueur de recouvrement (L_{ov}).

Conclusion

Nous avons présenté dans ce chapitre la modélisation compacte du transistor FinFET ultime. Nous nous sommes basés sur la modélisation du transistor MOSFET double grille, dont l'utilisation reste valide pour une géométrie du transistor FinFET aussi petites que $H_{Si} = 50$ nm, $W_{Si} = 3$ nm et $L = 25$ nm, pour une épaisseur d'oxyde de grille de $t_{ox} = 1.5$ nm. Pour cela, nous prenons en compte les effets canaux courts avec le roll-off de la tension de seuil, le DIBL et la dégradation de la pente sous-le-seuil, sans qu'aucun paramètre empirique ne soit nécessaire grâce par l'évaluation de la variation du potentiel minimum. Le modèle compact du FinFET intègre également l'influence des effets de mécanique quantiques qui sont importants

pour une largeur de silicium de quelques nanomètres. De même, la détermination des QMEs ne requiert aucun paramètre empirique. La modélisation de la dégradation de la mobilité et la modélisation de la modulation de longueur de canal sont prises en compte dans la définition d'une longueur effective du canal. Nous disposons désormais d'un modèle dynamique valide pour les petites géométries du transistor. Seule la longueur de recouvrement est à extraire pour ce modèle qui a été publié dans [7]. Cette prédiction statique et dynamique du courant de drain va nous permettre de réaliser des simulations de circuits à base de transistors FinFET qui rendent compte des retards temporels engendrés par les effets capacitifs du transistor. Les modélisations des différents phénomènes physiques sont réalisées indépendamment les unes des autres, et sont ensuite facilement combinées pour former un modèle complet. Ainsi, l'influence de chaque effet physique est directement visible dans chaque expression du modèle compact. Son faible nombre de paramètres à extraire en fait un modèle orienté conception pour la réalisation de circuits numériques. Une procédure d'extraction des paramètres du modèle est présentée au Chapitre 5.

Références

- [1] M. TANG, "Etude et modélisation compacte du transistor FinFET", thèse de doct., Université de Strasbourg, 2009.
- [2] J.-M. SALLESE, F. KRUMMENACHER, F. PRÉGALDINY, C. LALLEMENT, A. ROY et C. ENZ, "A design oriented charge-based current model for symmetric DG MOSFET and its correlation with the EKV formalism", *Solid-State Electron.*, vol. 49, n° 3, pp. 485–489, 2005.
- [3] F. PRÉGALDINY, F. KRUMMENACHER, B. DIAGNE, F. PÊCHEUX, J.-M. SALLESE et C. LALLEMENT, "Explicit modelling of the double-gate MOSFET with VHDL-AMS", *Int. J. Numer. Model.*, vol. 19, n° 3, pp. 239–256, 2006.
- [4] X. LIANG et Y. TAUR, "A 2-D analytical solution for SCEs in DG MOSFETs", *IEEE Trans. Electron Devices*, vol. 51, n° 9, pp. 1385–1391, 2004.
- [5] F. LIME, B. IÑÍGUEZ et O. MOLDOVAN, "A quasi-two-dimensional compact drain-current model for undoped symmetric double-gate MOSFETs including short-channel effects", *IEEE Trans. Electron Devices*, vol. 55, n° 6, pp. 1441–1448, 2008.
- [6] M. TANG, F. PRÉGALDINY, C. LALLEMENT et J.-M. SALLESE, "Explicit compact model for ultranarrow body FinFETs", *IEEE Trans. Electron Devices*, vol. 56, n° 7, pp. 1543–1547, 2009.
- [7] A. YESAYAN, F. PRÉGALDINY, N. CHEVILLON, C. LALLEMENT et J.-M. SALLESE, "Physics-based compact model for ultra-scaled FinFETs", *Solid-State Electron.*, vol. 62, n° 1, pp. 165–173, 2011.
- [8] G. NIU, J. CRESSLER, S. MATHEW et S. SUBBANNA, "A channel resistance derivative method for effective channel length extraction in LDD MOSFET's", *IEEE Trans. Electron Devices*, vol. 47, n° 3, pp. 648–650, 2000.
- [9] D. ESSENI, A. ABRAMO, L. SELMI et E. SANGIORGI, "Study of low field electron transport in ultra-thin single and double-gate SOI MOSFETs", dans *Proc. IEEE Int. Electron Devices Meeting (IEDM'02)*, 2002, pp. 719–722.

- [10] —, “Physically based modeling of low field electron mobility in ultrathin single- and double-gate SOI n-MOSFETs”, *IEEE Trans. Electron Devices*, vol. 50, n° 12, pp. 2445–2455, 2003.
- [11] A. CROS, K. ROMANJEK, D. FLEURY, S. HARRISON, R. CERUTTI, P. CORONEL, B. DUMONT, A. POUYDEBASQUE, R. WACQUEZ, B. DURIEZ, R. GWOZIECKI, F. BOEUF, H. BRUT, G. GHIBAUO et T. SKOTNICKI, “Unexpected mobility degradation for very short devices : a new challenge for CMOS scaling”, dans *Proc. IEEE Int. Electron Devices Meeting (IEDM’06)*, 2006, pp. 1–4.
- [12] C. LOMBARDI, S. MANZINI, A. SAPORITO et M. VANZI, “A physically based mobility model for numerical simulation of nonplanar devices”, *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 7, n° 11, pp. 1164–1171, nov. 1988.
- [13] C. MCANDREW, B. BHATTACHARYYA et O. WING, “A single-piece C_∞ -continuous MOSFET model including subthreshold conduction”, *IEEE Electron Device Lett.*, vol. 12, n° 10, pp. 565–567, 1991.
- [14] L. GE et J. FOSSUM, “Analytical modeling of quantization and volume inversion in thin Si-film DG MOSFETs”, *IEEE Trans. Electron Devices*, vol. 49, n° 2, pp. 287–294, 2002.
- [15] D. MUNTEANU, J. L. AUTRAN, X. LOUSSIER, S. HARRISON, R. CERUTTI et T. SKOTNICKI, “Quantum short-channel compact modelling of drain-current in double-gate MOSFET”, *Solid-State Electron.*, vol. 50, n° 4, pp. 680–686, 2006.
- [16] G. BACCARANI et S. REGGIANI, “A compact double-gate MOSFET model comprising quantum-mechanical and nonstatic effects”, *IEEE Trans. Electron Devices*, vol. 46, n° 8, pp. 1656–1666, 1999.
- [17] F. STERN et W. E. HOWARD, “Properties of semiconductor surface inversion layers in the electric quantum limit”, *Phys. Rev.*, vol. 163, pp. 816–835, 1967.
- [18] F. PRÉGALDINY, C. LALLEMENT, R. van LANGEVELDE et D. MATHIOT, “An advanced explicit surface potential model physically accounting for the quantization effects in deep-submicron MOSFETs”, *Solid-State Electron.*, vol. 48, n° 3, pp. 427–435, 2004.
- [19] T. ANDO, A. B. FOWLER et F. STERN, “Electronic properties of two-dimensional systems”, *Rev. Mod. Phys.*, vol. 54, pp. 437–672, 1982.
- [20] D. WARD et R. DUTTON, “A charge-oriented model for MOS transistor capacitances”, *IEEE Journ. Solid-State Circuits*, vol. 13, n° 5, pp. 703–708, oct. 1978.
- [21] N. ARORA, *MOSFET Models for VLSI Circuit Simulation. Theory and Practice*. New York : Springer-Verlag, 1993, ISBN 3-211-82395-6.
- [22] H. LU et Y. TAUR, “An analytic potential model for symmetric and asymmetric dg mosfets”, *IEEE Trans. Electron Devices*, vol. 53, n° 5, pp. 1161–1168, 2006.
- [23] C. LALLEMENT, J.-M. SALLESE, M. BUCHER, W. GRABINSKI et P. FAZAN, “Accounting for quantum effects and polysilicon depletion from weak to strong inversion in a charge-based design-oriented MOSFET model”, *IEEE Trans. Electron Devices*, vol. 50, n° 2, pp. 406–417, 2003.
- [24] H. BORLI, K. VINKENES et T. FJELDLY, “Physics based capacitance modeling of short-channel double-gate MOSFETs”, *Physica Status Solidi (c)*, vol. 5, n° 12, pp. 3643–3646, 2008.

Chapitre 4

Vers le modèle « universel » de MOSFET multigrille

Sommaire

4.1	Etat de l'art	74
4.2	Modélisation de transistors multigrilles selon un concept de paramètres géométriques équivalents	75
4.2.1	Définition de la largeur équivalente pour les transistors à effets de champs de section de grille rectangulaire	75
4.2.2	Généralisation de la largeur équivalente à toutes géométries à grille enrobante	80
4.2.3	Un cas d'étude : le GAA MOSFET à grille cylindrique	81
4.2.4	Extension de la modélisation à des transistors MOSFET à grille partiellement enrobante	84
4.2.5	Le cas du MOSFET bulk triple grille	90
4.3	Modélisation de l'effet de la température	93
4.4	Paramètres du modèle multigrille	94
4.5	Validation	94
4.5.1	Validation TCAD	95
4.5.2	Validation par rapport aux mesures expérimentales	98
	Conclusion	100
	Références	100

Dans ce chapitre, nous proposons de généraliser le concept de largeur équivalente, défini pour la modélisation du dopage à la section 2.3, afin de modéliser des structures arbitraires de MOSFET multigrilles, dites à grilles entièrement enrobantes, à canaux faiblement dopés, sans avoir besoin d'introduire de paramètres non physiques. Au concept de largeur équivalente, nous associons la notion de capacité d'oxyde de grille équivalente afin d'être capable de modéliser des structures multigrilles dont la grille n'enrobe que partiellement le canal de silicium. Ces définitions, qui permettent de transformer simplement une géométrie multigrille en une struc-

ture MOSFET double grille symétrique, ont été validées par de nombreuses comparaisons avec des simulations numériques 3D de MOSFET quadruple grille (QG), triple grille (TG), à grille triangulaire et à grille cylindrique (GAA), et de FinFETs double grille. Basé sur cette approche de modélisation, toute architecture multigrille fait ainsi appel aux relations fondamentales qui ont été développées pour le DG MOSFET planaire, incluant la normalisation de toutes les quantités électriques ce qui simplifie considérablement l'analyse de la structure multigrille. Ce chapitre traite également de la dépendance à la température du transistor MOSFET multigrille. Enfin, le modèle a pu être validé sur des mesures électriques d'un TG MOSFET, faisant de ce modèle générique un candidat intéressant pour un modèle orienté conception pour des géométries arbitraires de MOSFET multigrilles.

4.1 Etat de l'art

En ce qui concerne la modélisation compacte, le plus simple transistor MOS multigrille à modéliser est sans aucun doute le MOSFET double grille (DG MOSFET) symétrique entièrement déplété. Différentes approches ont été proposées. Taur et al. [1] a initialement développé une solution exacte qu'il est encore aujourd'hui difficile à utiliser dans le contexte de la modélisation compacte où de plus simples relations sont préférées. Par la suite, une formulation plus compacte a été proposée par Sallese et al. [2] dans laquelle une solution approchée a été cherchée, et qui mène à une formulation simple et précise. Cependant ces modèles concernent le DG MOSFET symétrique dont la technologie est difficile à mettre au point et encore non adoptée par l'industrie de la micro-électronique.

Inversement, les transistors MOS multigrilles tels que les FinFETs et Ω -FETs [3] ont prouvé leur compatibilité avec la technologie CMOS standard et les techniques de lithographie, menant à des circuits basiques [4]. Aujourd'hui, les FinFETs sont prévus en tant que dispositifs de prochaine génération pour les microprocesseurs [5]. Cependant, la modélisation d'architectures multigrille est assez complexe, leur structure étant à deux dimensions, aucune solution simple n'existe pour résoudre l'équation de Boltzmann-Poisson.

Une approche empirique a été proposée par Yu et al. [6] qui introduisent des fonctions de lissage et des paramètres d'ajustement pour modéliser les charges et le courant dans les topologies multigrille non planaires. Dans leur approche, les auteurs étendent leur modèle de MOSFET à grille enrobante (SG) cylindrique aux MOSFET quadruple grille (QG). En condition d'inversion volumique, le courant sous-le-seuil est proportionnel à la surface de silicium de la coupe transversale du canal. Par contre, pour une tension de grille bien supérieure au seuil, le courant devient proportionnel au périmètre de silicium de la coupe transversale du canal. La transition entre ces deux comportements asymptotiques pour le QG, a été modélisée selon un SG en introduisant une fonction linéaire empirique qui dépend de la tension de grille. Ainsi, les auteurs proposent un modèle pour MOSFET multigrille de section rectangulaire à grille partiellement enrobante, qui est une combinaison linéaire des modèles DG et QG. Par exemple, le transistor MOS triple grille (TG) est considéré comme la « moyenne » d'un DG

et d'un QG ayant la même section de silicium et la même épaisseur d'oxyde de grille que le TG. En effet, le courant en faible inversion est identique pour les trois structures, et celui en forte inversion est précisément la moyenne des courants du DG et du QG. Pour modéliser un Ω -FET ou un Π -FET, une autre fonction empirique a été introduite, en association avec des paramètres empiriques extraits de simulations TCAD et qui ne peuvent pas être prédits par une simple analyse.

Une modélisation analytique intéressante et complète des MOSFET multigrille en régime sous-le-seuil a aussi été proposé par Ritzenthaler et al. [7]. Le potentiel dans le volume du canal a été obtenu par la résolution de l'équation de Laplace à trois dimensions. L'intégration du potentiel électrostatique le long du chemin le plus favorable au courant, de la source au drain, a été utilisée pour obtenir une expression analytique du courant sous-le-seuil. Comme dans [6], les caractéristiques de courant de drain de la faible à la forte inversion ont été obtenues par l'intermédiaire de fonctions de lissage.

4.2 Modélisation de transistors multigrilles selon un concept de paramètres géométriques équivalents

L'approche de modélisation du transistor MOSFET double grille dopé à la Section 2.3 utilise le concept d'une largeur équivalente pour étendre le modèle non dopé à toute valeur de dopage du film de silicium. Pour rappel, nous avons résolu l'équation de Poisson tenant compte du niveau de dopage selon la même méthode que le cas non dopé. Puis, par identification des expressions de la constante d'intégration obtenue, nous avons pu établir une relation entre la largeur physique du modèle non dopé et une expression du dispositif dopé dépendant de sa largeur physique et du niveau de dopage. A partir de ce travail, nous proposons d'étendre le modèle du transistor MOSFET double grille non dopé pour prendre en compte les géométries non planaires, et aboutir à un modèle compact générique des dispositifs multigrilles sans l'aide de relations ou de paramètres empiriques.

4.2.1 Définition de la largeur équivalente pour les transistors à effets de champs de section de grille rectangulaire

Pour introduire l'idée du concept de largeur équivalente de silicium W_{eq} du DG MOSFET (Figure 4.1(a)) dans le cas des géométries à grille totalement enrobante, nous proposons d'illustrer une telle définition sur un MOSFET quadruple grille tel que décrit à la Figure 4.1(b).

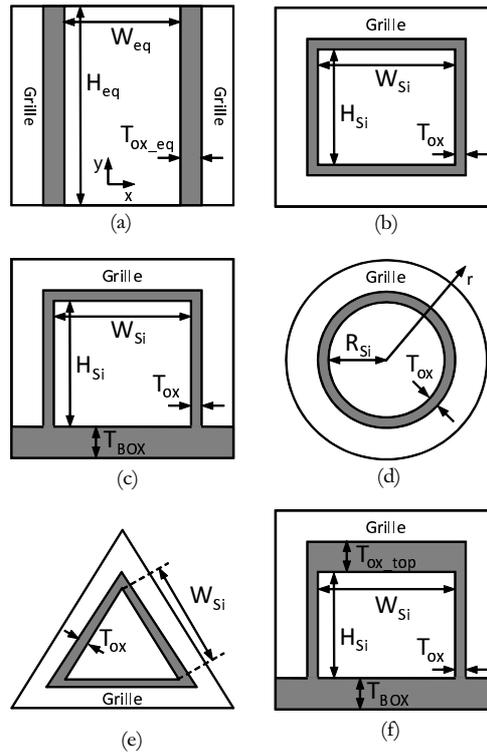


Figure 4.1 – Coupes schématiques de MOSFET multigrilles. (a) DG MOSFET, (b) QG MOSFET, (c) TG MOSFET, (d), GAA MOSFET, (e) MOSFET triangulaire, (f) DG FinFET

4.2.1.1 Expression de la densité de charge mobile

Contrairement à la méthode de résolution de l'équation de Boltzmann-Poisson pour les DG MOSFET, cette équation doit ici être résolue en deux dimensions selon cette formulation :

$$\frac{d^2\Psi(x,y)}{dx^2} + \frac{d^2\Psi(x,y)}{dy^2} = \frac{e \cdot n_i}{\varepsilon_{Si}} \cdot e^{\left(\frac{\Psi(x,y) - V_{ch}}{U_T}\right)} \quad (4.1)$$

A la différence de l'équation de Boltzmann-Poisson à une dimension, cette équation différentielle non linéaire n'a pas de solution analytique exacte. Il devient ainsi utile d'introduire quelques simplifications basées sur deux hypothèses simples. Comme pour le DG MOSFET, les charges mobiles en forte inversion sont localisées aux interfaces Si/SiO₂, puisque le volume de silicium a un faible impact sur la densité de charges totales. Cela signifie que le transistor se ramène à un système à quasiment une dimension. Nous pouvons ainsi suggérer que la solution obtenue pour le DG MOSFET canal long peut être encore suffisamment juste pour estimer la densité de charge en forte inversion pour le QG MOSFET.

Maintenant, en régime de faible inversion (le terme régime "sous-le-seuil" serait mieux adapté lorsque nous considérons des dispositifs non dopés, mais nous proposons de toujours utiliser l'expression "faible inversion" pour plus de simplicité), le côté droit de l'équation 4.1 s'annule, et une solution simple est donnée en considérant $\Psi(x,y)$ comme une constante de valeur Ψ_0 . La densité de charge est donc uniforme et le courant dépend uniquement de la surface de silicium de la coupe transversale du canal. Le transistor est en inversion volumique.

Dans une certaine mesure, ces arguments justifient le fait que le MOSFET multigrille rec-

tangulaire puisse être planarisé. Il peut ainsi être considéré comme un DG MOSFET avec le même périmètre d'interface Si/SiO₂, assurant la cohérence en forte inversion, et le même volume de silicium pour correspondre aux caractéristiques électriques sous-le-seuil. Dans cette transformation géométrique, nous ignorons simplement les effets de coin, qui sont connus pour être négligeables dans les canaux faiblement dopés [8].

A partir de ces considérations, nous proposons de montrer comment le QG MOSFET peut être modélisé comme un simple DG MOSFET. Dans [2], nous avons montré que la relation charge-potentiel dans un DG MOSFET symétrique est donnée par :

$$V_G - \Delta\Phi - V_{ch} = \frac{Q_G}{C_{ox}} + U_T \cdot \ln \left(\frac{Q_G^2}{2 \cdot \varepsilon_{Si} \cdot e \cdot U_T \cdot n_i} - C_1 \right) \text{ avec } C_1 = -e \frac{\Psi_0 - V_{ch}}{U_T} \quad (4.2)$$

où C_1 est une constante d'intégration qui joue un rôle principalement en faible inversion et qui est approximée dans [2].

La constante d'intégration C_1 est l'élément clé de l'extension de la modélisation du DG à des structures multigrille plus complexes. C_1 est reliée au potentiel du centre du canal du DG MOSFET, i.e. Ψ_0 , rendant sa valeur dépendante de la densité de charge mobile *par unité de surface* Q_{mDG}^{WI} . En effet, dans le DG MOSFET, tout comme dans le QG MOSFET puisque le potentiel est uniforme à travers tout le silicium en faible inversion, la concentration d'électrons dans le canal est $n_i \cdot e \frac{\Psi_0 - V_{ch}}{U_T}$. La quantité de charges dans la section transversale du canal du DG MOSFET est donc $-e \cdot n_i \cdot e \frac{\Psi_0 - V_{ch}}{U_T} \cdot (H_{eq} \cdot W_{eq})$, selon les notations de la Figure 4.1(a). Ceci mène à la densité de charge mobile *par unité de surface* pour le DG faisant apparaître l'expression de C_1 . Cette densité est donnée par :

$$Q_{mDG}^{WI} = \frac{-e \cdot n_i \cdot e \frac{\Psi_0 - V_{ch}}{U_T} \cdot (W_{eq} \cdot H_{eq})}{H_{eq}} = -e \cdot n_i \cdot e \frac{\Psi_0 - V_{ch}}{U_T} \cdot W_{eq} \quad (4.3)$$

Dans notre cas du QG MOSFET, la densité de charge mobile *par unité de surface* est donnée par :

$$Q_{mQG}^{WI} = \frac{-e \cdot n_i \cdot e \frac{\Psi_0 - V_{ch}}{U_T} \cdot (H_{Si} \cdot W_{Si})}{H_{Si} + W_{Si}} \quad (4.4)$$

En reprenant la relation charge-potentiel (4.2), issue de l'équation de Poisson à une dimension dans le DG, pour le QG en faible inversion qui est considéré comme un système à une dimension, nous pouvons établir une relation charge-potentiel pour le QG. Pour cela, nous exprimons la constante C_1 en fonction de la densité de charge mobile du QG en faible inversion :

$$C_1 (Q_{mQG}^{WI}) = \frac{Q_{mQG}^{WI}}{e \cdot n_i \cdot \frac{H_{Si} \cdot W_{Si}}{H_{Si} + W_{Si}}} \quad (4.5)$$

En exprimant C_1 en fonction de la densité de charge de grille, étant donné $Q_m = -2Q_G$, nous obtenons :

$$C_1 (Q_G^{WI}) = \frac{-Q_G^{WI}}{e \cdot n_i \cdot \frac{H_{Si} \cdot W_{Si}}{2 \cdot (H_{Si} + W_{Si})}} \quad (4.6)$$

Finalement, en réintroduisant C_1 dans la relation (4.2), nous obtenons la relation charge-potentiel pour le QG MOSFET :

$$V_G - \Delta\Phi - V_{ch} = \frac{Q_G}{C_{ox}} + U_T \cdot \ln \left(\frac{Q_G^2}{2 \cdot \varepsilon_{Si} \cdot e \cdot U_T \cdot n_i} + \frac{Q_G}{q \cdot n_i \cdot \frac{W_{eq}}{2}} \right) \quad (4.7)$$

avec la définition d'une largeur équivalente de DG MOSFET

$$W_{eq} = \frac{H_{Si} \cdot W_{Si}}{H_{Si} + W_{Si}} \quad (4.8)$$

En forte inversion, la constante C_1 étant négligeable, et le QG MOSFET pouvant se rapporter à un système à une dimension, la relation entre charges et potentiels est considérée pour les régimes de fonctionnement de faible à forte inversion.

Comme dans [9], nous introduisons des normalisations qui seront très utiles pour la cohérence du modèle et dans l'objectif de la conception de circuits. Ainsi, tensions et charges sont respectivement normalisées selon $v = V/U_T$ et $q = Q/Q_{SP}$, où $Q_{SP} = 4 \cdot C_{ox} \cdot U_T$. De là nous obtenons une relation fondamentale sans dimension :

$$v_g - \Delta\phi - v_{to} - v_{ch} = 4 \cdot q_g + \ln(q_g) + \ln \left(1 + q_g \cdot \frac{C_{ox}}{C_{Si_{eq}}} \right) \quad (4.9)$$

où v_{to} est la tension de seuil canal long normalisée définie comme :

$$v_{to} = \ln \left(\frac{8 \cdot C_{ox} \cdot U_T}{e \cdot n_i \cdot W_{eq}} \right) \quad (4.10)$$

et $C_{Si_{eq}} = \varepsilon_{Si}/W_{eq}$ est la capacité équivalente multigrille de silicium.

4.2.1.2 Courant normalisé et hauteur équivalente

Puisque la relation (4.9) a la même forme que la relation (2.43) du modèle canal long, nous pouvons utiliser la même expression de courant que dans [2] et donc la même convention que pour le DG, pour définir une hauteur équivalente H_{eq} . Cette dernière est en effet la moitié du périmètre de silicium, i.e. $H_{eq} = H_{Si} + W_{Si}$. Il est intéressant de remarquer que la largeur équivalente de silicium W_{eq} apparaît dans l'expression du courant (4.12). Nous rappelons que la densité de charge mobile normalisée est donnée par $q_m = -2 \cdot q_g$. Ainsi, les paramètres géométriques (W_{eq}, H_{eq}) du DG MOSFET se retrouvent en tant que paramètres géométriques équivalents pour le QG MOSFET, dans le facteur de normalisation du courant. Le courant spécifique du QG MOSFET est maintenant défini comme :

$$I_{SP} = 4 \cdot \mu \cdot C_{ox} \cdot U_T^2 \cdot \frac{H_{eq}}{L} \quad (4.11)$$

et le courant normalisé correspondant devient :

$$i = -q_m^2 + 2 \cdot q_m + 2 \cdot \frac{C_{Si_{eq}}}{C_{ox}} \cdot \ln \left(1 - q_m \cdot \frac{C_{ox}}{2 \cdot C_{Si_{eq}}} \right) \quad (4.12)$$

où la capacité physique de silicium est remplacée par sa valeur équivalente $C_{Si_{eq}} = \varepsilon_{Si}/W_{eq}$.

4.2.1.3 Comparaison avec des simulations numériques 3D

Nous proposons maintenant d'évaluer le modèle analytique pour QG MOSFET basé sur la largeur et la hauteur équivalente, par rapport à des simulations numériques 3D définies au chapitre 5. Afin de minimiser le nombre de paramètres et ainsi nous concentrer sur la validation de l'approche de la largeur équivalente, une mobilité constante de $1000 \text{ cm}^2/\text{V}\cdot\text{s}$ est ici utilisée. Les paramètres physiques et équivalents sont obtenus à partir du même ensemble de paramètres géométriques. Des simulations ont été réalisées pour un QG MOSFET avec une section carré de silicium de 40 nm de côté. La longueur du canal est fixée à $1 \mu\text{m}$. Dans cette configuration géométrique, la largeur et la hauteur équivalente de silicium sont respectivement 20 nm et 80 nm. La Figure 4.2 montre le courant de drain en fonction de la tension de grille à une faible et à une forte tension de drain, et la Figure 4.3 montre le courant de drain en fonction de la tension de drain pour différentes tensions de grille. Pour ces deux types de caractéristiques, la comparaison du modèle (relations (4.11) et (4.12)) avec des simulations 3D confirme la validité de l'approche de modélisation présentée précédemment. La dépendance du courant de drain avec la tension de grille (Figure 4.2) est modélisée avec une très bonne précision, que ce soit selon l'échelle linéaire ou logarithmique, et ce sans avoir besoin d'introduire de paramètre empirique. De même, l'appariement des caractéristiques de sorties montrées à la Figure 4.3 est également bon. Ainsi, la transformation du QG MOSFET en un DG MOSFET basée sur la définition de paramètres géométriques équivalent est totalement justifiée et validée.

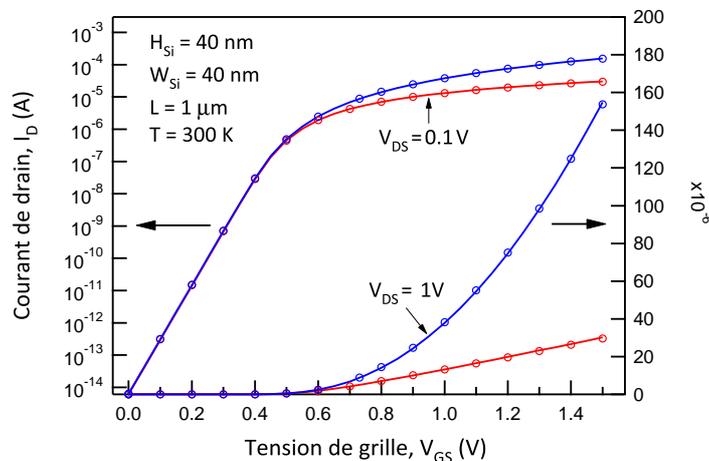


Figure 4.2 – Courant de drain d'un QG MOSFET avec une section de silicium carrée en fonction de la tension de grille en régimes linéaire et saturé. Simulations 3D : symboles, modèle : lignes

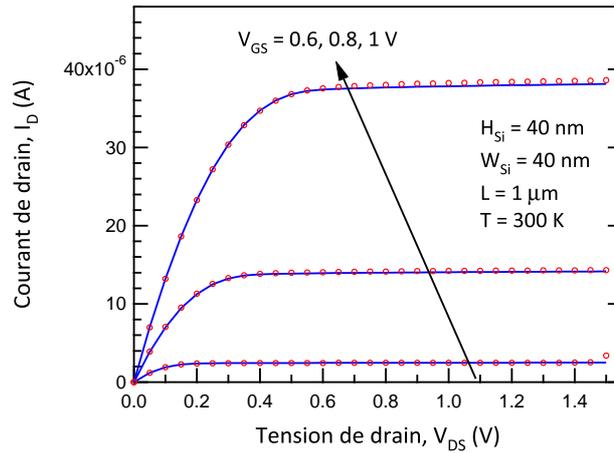


Figure 4.3 – Courant de drain d'un QG MOSFET avec une section de silicium carrée en fonction de la tension de drain en régimes linéaire et saturé. Simulations 3D : symboles, modèle : lignes

4.2.2 Généralisation de la largeur équivalente à toutes géométries à grille enrobante

Jusqu'ici, nous avons été capable de simuler les charges et les courants de MOSFET de section carrée en nous appuyant sur une relation générique et la définition d'une largeur équivalente de silicium. En supposant une capacité d'oxyde de grille constante le long du périmètre de la section de silicium du canal, nous pouvons aller plus loin et proposer une généralisation de cette approche à toutes géométries à grille enrobante, même non rectangulaires.

Nous considérons maintenant un dispositif avec une section de silicium du canal de surface S et de périmètre P . Lorsque le potentiel est uniforme dans le canal, i.e. en faible inversion, $\Psi(x, y) = \Psi_0$. Ainsi, en suivant la même approche que dans la section 4.2.1, nous obtenons une relation générique entre charges et potentiels qui est simplement la relation (4.9) incluant la largeur équivalente définie comme suit :

$$W_{eq} = \frac{2 \cdot S}{P} \quad (4.13)$$

Du point de vue du courant, nous pouvons toujours définir un courant normalisé comme :

$$i = -q_m^2 + 2 \cdot q_m + 2 \cdot \frac{C_{Si_{eq}}}{C_{ox}} \cdot \ln \left(1 - q_m \cdot \frac{C_{ox}}{2 \cdot C_{Si_{eq}}} \right) \quad (4.14)$$

avec le courant spécifique maintenant donné par $I_{SP} = 4 \cdot \mu \cdot C_{ox} \cdot U_T^2 \cdot H_{eq}/L$, avec $H_{eq} = P/2$.

Notons que lorsque les interfaces Si/SiO₂ du film de silicium ne sont plus planaires, la capacité d'oxyde de grille n'est plus donnée par la plus courte distance entre la grille et le silicium. Il est alors nécessaire d'utiliser une expression plus appropriée de cette capacité. Cela s'illustre dans le cas du dispositif GAA MOSFET (Figure 4.1(d)). La capacité de grille doit être évaluée avec attention, particulièrement pour la forte inversion où ce paramètre a un gros impact sur la dépendance de la charge avec la tension de grille. N'oublions pas que la capacité

affecte aussi la faible inversion à travers la tension de seuil.

Pour conclure cette partie de généralisation, nous proposons de vérifier la validité de l'épaisseur équivalente pour un MOSFET à grille enrobante de section triangulaire (Figure 4.1(e)), dont une étude est rapportée dans [10]. Le choix de cette géométrie est motivé par le fait qu'elle combine à la fois une géométrie 2D et des angles vifs, susceptibles de compliquer la modélisation. Le dispositif consiste en un MOSFET de section de silicium triangulaire équilatéral dont la longueur W_{Si} du côté vaut 10 nm. Selon la relation (4.13), la largeur équivalente est alors $W_{eq} = \frac{1}{3} \cdot \sqrt{W_{Si}^2 - \left(\frac{W_{Si}}{2}\right)^2}$ et la hauteur équivalente du courant spécifique est $H_{eq} = \frac{3}{2} \cdot W_{Si}$. La Figure 4.4 montre des simulations 3D du courant de drain en fonction de la tension de grille pour ce MOSFET triangulaire. Nous observons toujours un très bon appariement entre les simulations et le modèle DG utilisant le concept de la largeur équivalente. Cela soutient le fait que notre approche est robuste même si les angles entre grilles deviennent aigus.

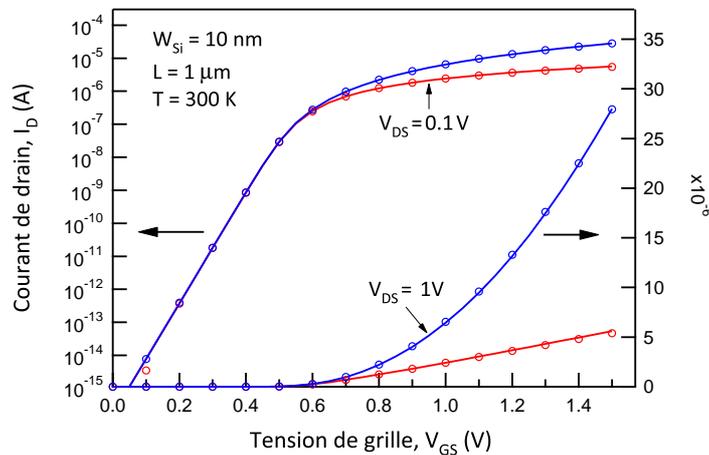


Figure 4.4 – Courant de drain d'un MOSFET à grille triangulaire en fonction de la tension de drain en régimes linéaire et saturé. La longueur d'arrêt du silicium de la section du canal est W_{Si} . Simulations 3D : symboles, modèle : lignes

4.2.3 Un cas d'étude : le GAA MOSFET à grille cylindrique

4.2.3.1 La solution approximée du DG, une solution exacte pour le GAA

L'architecture multigrille non plane ultime est sans aucun doute le GAA MOSFET à grille cylindrique (Figure 4.1(d)). De façon intéressante, une solution analytique pour les relations de la charge et du courant en fonction des potentiels a été proposée dans [11].

L'équation différentielle de Boltzmann-Poisson est maintenant exprimée en coordonnées cylindriques :

$$\frac{d^2\Psi(r)}{dr^2} + \frac{1}{r} \cdot \frac{d\Psi(r)}{dr} = \frac{e \cdot n_i}{\varepsilon_{Si}} \cdot e^{\left(\frac{\Psi(r) - V_{ch}}{U_T}\right)} \quad (4.15)$$

Contrairement au DG MOSFET plane, cette équation a une solution exacte [12] donnée

par :

$$\Psi(r) = V_{ch} + U_T \cdot \ln \left(\frac{-8 \cdot B}{(1 + B \cdot r^2)^2} \cdot \frac{U_T \cdot \varepsilon_{Si}}{e \cdot n_i} \right) \quad (4.16)$$

Ici, le terme B est relié à la condition du potentiel de surface. En considérant l'expression de la charge, à partir de la capacité d'oxyde de grille, $Q_G = C_{ox} \cdot (V_G - \Delta\Phi - \Psi(R_{Si}))$, et à partir de l'application du théorème de Gauss à l'interface Si/SiO₂, i.e. $Q_G = \varepsilon_{Si} \cdot \frac{d\Psi(r)}{dr} \Big|_{r=R_{Si}}$, nous pouvons montrer que nous obtenons une relation implicite entre les potentiels et la densité de charge de grille :

$$\frac{V_G - \Delta\Phi - V}{U_T} - \ln \left(\frac{8}{\delta \cdot R_{Si}^2} \right) = \frac{Q_G}{C_{ox} \cdot U_T} + \ln \left(\frac{Q_G}{Q_0} \right) + \ln \left(1 + \frac{Q_G}{Q_0} \right) \quad (4.17)$$

avec $Q_0 = \frac{4 \cdot \varepsilon_{Si} \cdot U_T}{R_{Si}}$

où R_{Si} est le rayon du nanofil de silicium, $\delta = \frac{e \cdot n_i}{\varepsilon_{Si} \cdot U_T}$ est un paramètre dépendant uniquement des propriétés du semi-conducteur et C_{ox} est la capacité d'oxyde de grille *par unité de surface* pour le GAA donnée par :

$$C_{ox} = \frac{\varepsilon_{ox}}{R_{Si} \cdot \ln \left(1 + \frac{T_{ox}}{R_{Si}} \right)} \quad (4.18)$$

La relation 4.17 a de fortes similarités avec la relation générique 4.2 que nous avons utilisée jusqu'ici pour le DG MOSFET : en effet la relation charge-potential pour un DG MOSFET symétrique, avec les quantités normalisées, est :

$$v_g - v - \Delta\phi + \ln \left(\frac{q_{int}}{2} \right) = 4 \cdot q_g + \ln(q_g) + \ln \left(1 + q_g \cdot \frac{C_{ox}}{C_{Si}} \right) \quad (4.19)$$

Maintenant, en utilisant le même facteur de normalisation des charges que pour le DG MOSFET, la relation charge-potentiels normalisée du GAA prend la forme :

$$v_g - v - \Delta\phi - v_{to} = 4 \cdot q_g + \ln(q_g) + \ln \left(1 + q_g \cdot \frac{C_{ox}}{C_{Si}} \right) \quad (4.20)$$

où la densité de charge de grille normalisée est donnée par $q_g = Q_G/Q_{SP}$, la tension de seuil normalisée $v_{to} = \ln \left(\frac{8 \cdot C_{ox} \cdot U_T}{e \cdot n_i \cdot R_{Si}} \right)$ et la capacité de silicium est définie par $C_{Si} = \varepsilon_{Si}/R_{Si}$.

Ainsi, nous obtenons une relation charge-potential normalisée qui est strictement identique à la solution approximée du transistor double grille planaire [voir relation (4.19)]. Cependant, étant donné que (4.20) est une solution exacte, nous pouvons nous demander si la largeur équivalente de silicium s'applique toujours.

D'après notre concept, une telle largeur équivalente devrait être donnée par la surface divisée par le demi-périmètre du fil de silicium. De là, la largeur équivalente serait $W_{eq} = \frac{\pi \cdot R_{Si}^2}{(2 \cdot \pi \cdot R_{Si})/2} = R_{Si}$, ce qui est en effet ce que donne (4.20).

Ainsi, la définition de la largeur équivalente est pleinement cohérente avec la solution exacte de l'équation de Poisson-Boltzmann pour le GAA. Ceci est une propriété remarquable à laquelle nous ne nous attendions pas.

Ensuite, nous pouvons adopter la même procédure que pour le DG MOSFET et calculer le courant en fonction des potentiels appliqués, directement à partir de la relation (4.14). Dans ce cas, la hauteur équivalente sera la moitié du périmètre du fil de silicium, $H_{eq} = \pi \cdot R_{Si}$, impliquant que le facteur de normalisation du courant devienne :

$$I_{SP} = 4 \cdot \mu \cdot C_{ox} \cdot U_T^2 \cdot \frac{\pi \cdot R_{Si}}{L} \quad (4.21)$$

et le courant normalisé est toujours donné par :

$$i = \frac{I_D}{I_{SP}} = -q_m^2 + 2 \cdot q_m + 2 \cdot \frac{C_{Si}}{C_{ox}} \cdot \ln \left(1 - q_m \cdot \frac{C_{ox}}{2 \cdot C_{Si}} \right) \text{ avec } q_m = -2 \cdot q_g \quad (4.22)$$

Ces fortes similarités entre le DG et le GAA MOSFET mènent à un unique formalisme tant pour la relation charge-potentiels et que pour la relation courant-potentiels. Ainsi, en considérant la définition appropriée de la capacité d'oxyde de grille, il ressort que les caractéristiques électriques d'un GAA avec un rayon R_{Si} peuvent être prédites d'un modèle de DG planaire, en utilisant les paramètres géométriques équivalents $W_{eq} = R_{Si}$ et $H_{eq} = \pi \cdot R_{Si}$.

Il est intéressant de remarquer que cet apparemment « pire cas » en terme d'éloignement d'un système 1D idéal, peut être exactement résolu en adoptant le formalisme générique lorsqu'il s'appuie sur la définition de la largeur équivalente.

4.2.3.2 Comparaison avec des simulations TCAD 3D

Comme pour le QG MOSFET, nous exécutons des simulations TCAD pour le courant en fonction des tensions de grille et de drain. Puisque la solution utilisée dans (4.22) est maintenant exacte, nous nous attendons à avoir de très bonnes prédictions du modèle. Ceci est en effet le cas pour les rayons de canal de silicium de 10, 20 et 30 nm, pour lesquels le modèle est vraiment précis (Figures 4.5 et 4.6). Une légère sous-estimation du courant à relativement grands V_{DS} est à noter. Ceci est du à la nature 2D de la distribution du champ électrique proche du drain.

En conclusion, pour autant que l'épaisseur d'oxyde de grille soit uniforme, le concept de largeur équivalente nous a permis de simuler des géométries multigrilles non planaires avec une grande justesse. Il existe cependant des structures multigrilles pour lesquelles la grille n'entoure pas totalement le canal de silicium avec un oxyde fin (exemple : le FinFET). Dans ce cas, la capacité d'oxyde est non constante le long du périmètre du canal. Nous proposons d'adresser ce cas non idéal dans la section suivante.

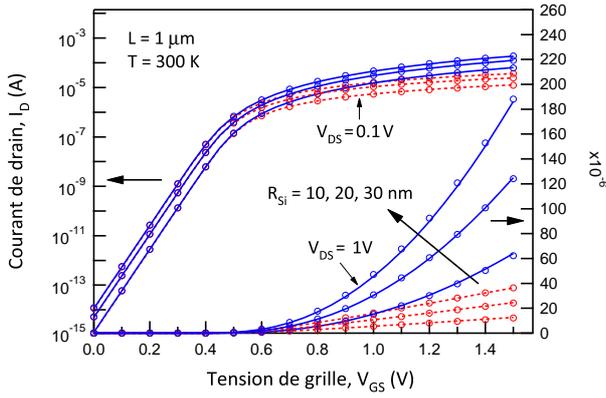


Figure 4.5 – Courant de drain d’un GAA MOSFET en fonction de la tension de grille en régimes linéaire et saturé pour différents rayons de silicium. Simulations 3D : symboles, modèle : lignes.

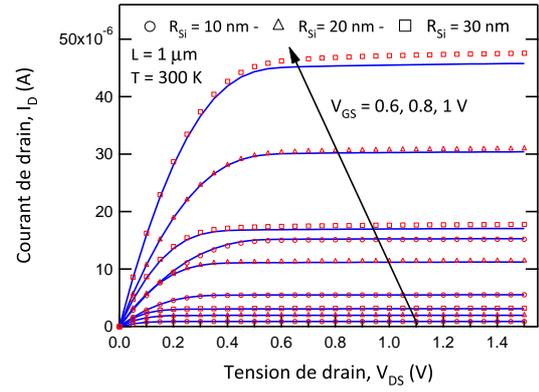


Figure 4.6 – Courant de drain d’un GAA MOSFET en fonction de la tension de drain en régimes linéaire et saturé pour différents rayons de silicium. Simulations 3D : symboles, modèle : lignes.

4.2.4 Extension de la modélisation à des transistors MOSFET à grille partiellement enrobante

4.2.4.1 Épaisseur d’oxyde de grille non uniforme

Lorsque la grille n’est pas uniforme tout autour du canal, comme pour le DG FinFET (Figure 4.1(f)) ou le TG MOSFET (Figure 4.1(c)), il faut faire attention puisque la capacité d’oxyde de grille n’est plus constante le long de l’interface Si/SiO₂ de la section transversale du film de silicium. En régime de faible inversion, si nous supposons l’inversion volumique et donc que le potentiel est constant dans tout le volume du silicium, nous nous attendons à ce que ces topologies donnent le même courant de drain que pour le QG puisque le courant dépend uniquement de la surface de silicium de la section du film. Cependant les transistors DG FinFET et TG MOSFET ont de plus petits périmètres de grille autour du canal, ce qui implique que le courant du transistor en forte inversion sera diminué en comparaison avec le QG. Si nous supposons que l’inversion des charges n’apparaît pas aux interfaces Si/SiO₂ non recouvertes par la grille, la largeur équivalente et la hauteur équivalente peuvent être définies en prenant le périmètre de silicium couvert par la fine couche d’oxyde de grille au lieu du périmètre entier, dans la relation (4.4).

Nous pouvons ainsi définir un modèle de MOSFET triple grille en utilisant le concept de la largeur équivalente. En considérant que le périmètre géométrique P est le périmètre de silicium couvert par la « fine » couche d’oxyde de grille, alors la hauteur équivalente et la largeur équivalente sont respectivement :

$$\begin{aligned} H_{eq} &= H_{Si} + \frac{W_{Si}}{2} \\ W_{eq} &= \frac{2 \cdot H_{Si} \cdot W_{Si}}{2 \cdot H_{Si} + W_{Si}} \end{aligned} \quad (4.23)$$

Comparé au TG MOSFET, le DG FinFET a une plus grande épaisseur d’oxyde de grille

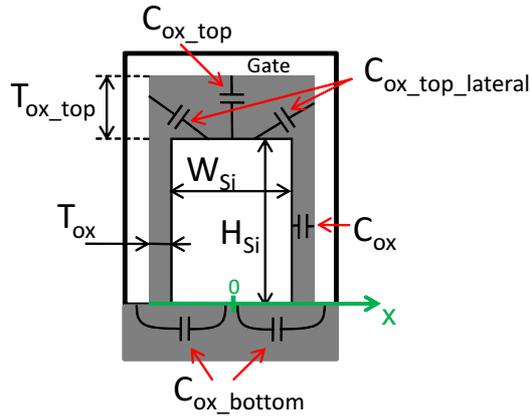


Figure 4.7 – Réseau de capacités équivalentes d'oxyde de grille du DG FinFET.

supérieure nommée T_{ox_top} . Si nous appliquons le même raisonnement sur le DG FinFET, lorsque cette épaisseur T_{ox_top} est assez large, nous pouvons considérer ce dispositif comme un DG MOSFET planaire, et le modéliser en tant que tel, ce que nous faisons dans le chapitre 3 du FinFET ultime.

Cependant, même pour une couche d'oxyde plus épaisse que la valeur nominale, les interfaces à l'oxyde épais peuvent tout de même voir apparaître un canal lorsque la tension de grille dépasse la tension de seuil. Ceci est en effet supporté par les simulations 3D. Ainsi, des courants additionnels localisés aux interfaces d'oxydes épais, devraient aussi être pris en compte lorsqu'on veut réaliser une modélisation très précise des architectures à grille non enrobante.

4.2.4.2 Définition d'une épaisseur équivalente d'oxyde de grille

Le cas du DG FinFET

Nous nous rapportons maintenant à la Figure 4.7 pour définir la section transversale du transistor DG FinFET. Basiquement, si nous considérons qu'en forte inversion chaque canal est « indépendant », c'est-à-dire que le champ électrique normal aux interfaces Si/SiO₂ domine, le courant total du DG FinFET peut être considéré comme la somme de tous les courants des canaux des interfaces Si/SiO₂. De plus, si l'épaisseur d'oxyde est différente à chaque interface, cela affectera à la fois la tension de seuil correspondante et le courant normalisé [voir (4.14)].

A ce point, il est intéressant d'introduire une simple hypothèse qui consiste à considérer que les tensions de seuil aux interfaces d'oxyde de grille fin et épais, seront presque les mêmes. Cette hypothèse signifie que le courant de forte inversion à chaque interface apparaîtra pour la même tension de grille. Ainsi, nous considérons le transistor FinFET comme un transistor MOSFET quadruple grille dont les quatre courants de forte inversion n'ont pas tous la même intensité. La modélisation du transistor FinFET par le modèle multigrille implique donc d'utiliser la largeur équivalente (4.8) du MOSFET QG.

En forte inversion, le courant normalisé i^{SI} peut être approximé par $-q_m^2 + 2 \cdot q_m$ et de ce fait il ne dépend pas de la capacité d'oxyde de grille. Cela signifie que le courant total en forte

inversion I_D^{SI} pour le DG FinFET peut être exprimé via la somme des courants spécifiques I_{SP} pour chaque canal :

$$I_D^{SI} = i^{SI} \cdot (2 \cdot I_{SP_{Lateral}} + I_{SP_{Top}} + I_{SP_{Bottom}}) \quad (4.24)$$

Nous proposons de ramener le courant en forte inversion ayant lieu à une interface au courant en forte inversion développé par un transistor double grille. Cette identification est correcte lorsque la longueur d'une interface du FinFET (ou pseudo MOSFET QG) est égale à deux fois la largeur du double grille, et l'expression du courant spécifique du transistor double grille peut être utilisé. Ainsi, le courant spécifique pour une interface donnée est $I_{SP} = 4 \cdot \mu \cdot C_{ox_{interface}} \cdot U_T^2 \cdot L_{interface}/L$, avec $L_{interface}$ la demi-longueur de l'interface ($W_{Si}/2$ ou $H_{Si}/2$). A partir de la relation (4.24), nous obtenons alors :

$$I_D^{SI} = i^{SI} \cdot \frac{4 \cdot \mu \cdot U_T^2}{L} \cdot \left(2 \cdot C_{ox} \cdot \frac{H_{Si}}{2} + C_{ox_{top_{eq}}} \cdot \frac{W_{Si}}{2} + C_{ox_{bottom}} \cdot \frac{W_{Si}}{2} \right) \quad (4.25)$$

ou de manière équivalente

$$I_D^{SI} = i^{SI} \cdot 4 \cdot \mu \cdot C_{ox_{eq}} \cdot U_T^2 \cdot \frac{H_{eq}}{L} \quad (4.26)$$

où $H_{eq} = H_{Si} + W_{Si}$, et $C_{ox_{eq}}$ est une capacité équivalente d'oxyde de grille donnée par :

$$C_{ox_{eq}} = \frac{C_{ox} \cdot H_{Si} + C_{ox_{top_{eq}}} \cdot W_{Si}/2 + C_{ox_{bottom}} \cdot W_{Si}/2}{H_{Si} + W_{Si}} \quad (4.27)$$

$C_{ox_{bottom}}$ et $C_{ox_{top_{eq}}}$ sont définies comme les capacités équivalentes d'oxyde *par unité de surface* pour les interfaces inférieure et supérieure, et $C_{ox} = \varepsilon_{ox}/T_{ox}$ pour, respectivement, les interfaces latérales. Ainsi, en adoptant ces nouvelles définitions, nous pouvons maintenant considérer le DG FinFET comme un MOSFET QG avec un courant spécifique donné par $I_{SP} = 4 \cdot \mu \cdot C_{ox_{eq}} \cdot U_T^2 \cdot \frac{H_{Si} + W_{Si}}{L}$.

Le courant normalisé garde la même forme et devient :

$$i = -q_m^2 + 2 \cdot q_m + 2 \cdot \frac{C_{Si_{eq}}}{C_{ox_{eq}}} \cdot \ln \left(1 - q_m \cdot \frac{C_{ox_{eq}}}{2 \cdot C_{Si_{eq}}} \right) \quad (4.28)$$

avec la largeur équivalente $W_{eq} = \frac{H_{Si} \cdot W_{Si}}{H_{Si} + W_{Si}}$ et la capacité équivalente de silicium $C_{Si_{eq}} = \frac{\varepsilon_{Si}}{W_{eq}}$.

Finalement, le problème de la modélisation des MOSFET à grille partiellement enrobante revient à trouver une expression analytique des capacités de chaque interface Si/SiO₂.

Généralisation

Ce résultat peut être généralisé à n'importe quel type de transistor à effet de champ ayant une grille commune avec n interfaces canal/oxyde de grille de longueur L_i et de capacité C_i . Dans ce cas, si nous supposons que le dispositif peut être planarisé afin d'être ramené à un DG

MOSFET, chaque longueur des interfaces représentera deux fois une largeur en topologie DG, la largeur du DG étant la longueur d'une seule de ses interfaces. De plus si nous gardons la même hypothèse qui est une tension de seuil commune, alors le courant spécifique correspondant sera la somme de tous les courants spécifiques I_{SP_i} :

$$I_{SP} = \sum_{i=1}^n I_{SP_i} = 4 \cdot \mu \cdot U_T^2 \cdot \sum_{i=1}^n C_i \cdot \frac{L_i/2}{L} = 4 \cdot \mu \cdot C_{ox.eq} \cdot U_T^2 \cdot \frac{H_{eq}}{L} \quad (4.29)$$

où la hauteur équivalente est donnée par $H_{eq} = \sum_{i=1}^n L_i/2$, qui est pleinement cohérente avec la définition introduite précédemment. La capacité équivalente d'oxyde de grille *par unité de surface* (dans le sens du DG) est définie par :

$$C_{ox.eq} = \frac{\sum_{i=1}^n C_i \cdot L_i/2}{\sum_{i=1}^n L_i/2} = \frac{\sum_{i=1}^n C_i \cdot L_i}{\sum_{i=1}^n L_i} \quad (4.30)$$

Il est à noter que la capacité équivalente d'oxyde de grille se ramène à la valeur moyenne de la capacité *par unité de surface* dans le contexte du formalisme du DG.

Évaluation précise des capacités supérieure et inférieure des DG FinFETs

Alors que les capacités d'oxyde de grille fin sont bien décrites par le modèle de capacité semi-infini, cette vue idéale ne tient plus lorsque l'épaisseur de la capacité devient comparable à sa longueur. Nous retrouvons cette configuration dans la modélisation de la capacité de la grille supérieure du DG FinFET. Un traitement spécial est alors nécessaire. Dans ce but, nous avons évalué cette capacité en partant du travail de Roy et al. [13] qui modélise les capacités frangeantes de grille. En effet, la configuration géométrique des capacités d'oxyde frangeantes est presque la même. Les correspondances des paramètres géométriques avec [13] sont listées dans le tableau 4.1.

Tableau 4.1 – Correspondance entre la géométrie latérale de la grille et la géométrie de l'oxyde supérieur.

Paramètres géométriques de Roy [13]	Paramètres géométriques du DG FinFET (Figure 4.7)
Épaisseur de grille T_g	Épaisseur d'oxyde supérieur $T_{ox.top}$
Épaisseur d'oxyde T_{ox}	0
Longueur d'underlap L_u	Épaisseur d'oxyde latéral T_{ox}
Longueur de source/drain L_{sd}	Moitié de la largeur de silicium $W_{Si}/2$

Suivant [13], le facteur d'échelle devient :

$$\lambda = \frac{T_{ox.top}^2 \cdot \left(\frac{W_{Si}^2}{4} + T_{ox} \cdot W_{Si} \right)}{T_{ox}^2 \cdot \left(T_{ox.top}^2 + T_{ox}^2 + \frac{W_{Si}^2}{4} + T_{ox} \cdot W_{Si} \right)} \quad (4.31)$$

et la capacité approximée d'oxyde latéral *par unité de surface*, de l'interface supérieure, qui

prend uniquement en compte le champ électrique normal à l'interface, est donnée par :

$$C_{ox_top_lateral} = \varepsilon_{ox} \cdot \frac{1}{\pi} \cdot \ln\left(\frac{1}{k}\right) \cdot \frac{2}{W_{Si}} \quad (4.32)$$

avec $k = 1 + 2 \cdot (\lambda - \sqrt{\lambda^2 + \lambda})$.

A partir de cette analyse, la capacité équivalente d'oxyde de grille de l'interface supérieure peut être calculée selon :

$$C_{ox_top_eq} = \frac{W_{Si} \cdot C_{ox_top} + W_{Si} \cdot C_{ox_top_lateral}}{W_{Si}} \quad (4.33)$$

Concernant la capacité d'oxyde inférieure (sans électrode de grille arrière) liée à l'oxyde enterré (BOX), voir la Figure 4.7, une expression de la capacité frangeante à l'interface canal/BOX, fonction de la distance entre la grille et une position sur l'interface, a été proposée par Ernst et al. [14] :

$$C_{BG}(x) = \frac{\varepsilon_{ox}}{T_{ox}} \cdot \frac{1}{\exp\left[\frac{-\pi}{T_{BOX}} \cdot \left(x - \left(\frac{W_{Si}}{2} + T_{ox}\right)\right)\right] - 1} \quad (4.34)$$

La capacité totale peut alors être obtenue en intégrant (4.34) de 0 à $W_{Si}/2$, aboutissant à une capacité d'oxyde inférieure *par unité de surface* donnée par :

$$C_{ox_bottom} = \varepsilon_{ox} \cdot \frac{2}{W_{Si}} \cdot \left[-\frac{W_{Si}}{2 \cdot T_{BOX}} + \frac{1}{\pi} \cdot \ln\left(\frac{\exp\left(\frac{\pi \cdot (W_{Si} + 2 \cdot T_{ox})}{2 \cdot T_{BOX}}\right) - 1}{\exp\left(\frac{\pi \cdot T_{ox}}{T_{BOX}}\right) - 1}\right) \right] \quad (4.35)$$

Ayant défini toutes les capacités, la capacité total équivalente d'oxyde de grille pour le DG FinFET $C_{ox_eq_{DG_FinFET}}$ peut être calculée d'après la relation (4.27).

Dans le cas particulier d'un TG MOSFET, puisque la grille supérieure a la même épaisseur d'oxyde que les grilles latérales, selon la relation (4.30), la capacité équivalente d'oxyde devient :

$$C_{ox_eq_{TG}} = \frac{(H_{Si} + W_{Si}/2) \cdot C_{ox} + (W_{Si}/2) \cdot C_{ox_bottom}}{H_{Si} + W_{Si}} \quad (4.36)$$

Il est intéressant de remarquer que lorsque les capacités supérieure et inférieure sont négligeables, la capacité équivalente se ramène à la capacité d'oxyde C_{ox} d'un DG MOSFET planaire, prouvant ainsi la cohérence de la définition de la capacité équivalente.

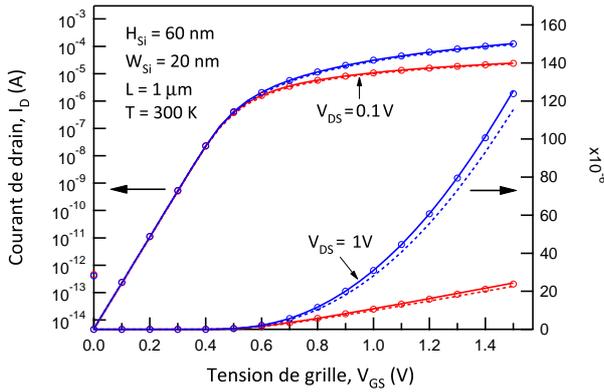


Figure 4.8 – Courant de drain d'un DG FinFET en fonction de la tension de grille en régime linéaire et saturé avec une épaisseur d'oxyde supérieure de 50 nm. Simulations 3D : symboles, modèle DG : lignes pointillées, modèle QG avec la capacité équivalente d'oxyde : lignes pleines.

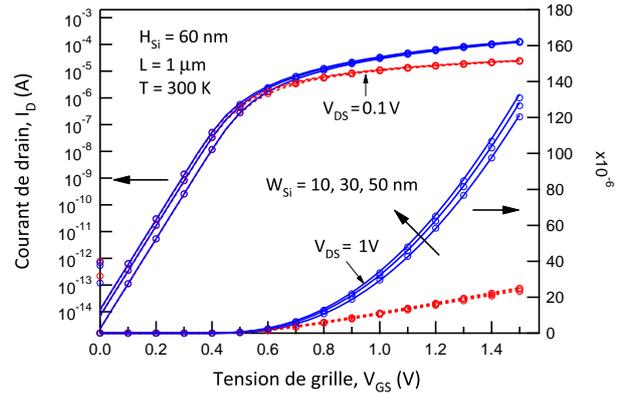


Figure 4.9 – Courant de drain de DG FinFETs en fonction de la tension de grille en régime linéaire et saturé pour différentes largeurs de silicium W_{Si} avec une épaisseur d'oxyde supérieure de 50 nm. Simulations 3D : symboles, modèle QG avec la capacité équivalente d'oxyde : lignes.

4.2.4.3 Simulations TCAD 3D de DG FinFETs et de TG MOSFET

La Figure 4.8 montre le courant d'un DG FinFET en fonction de la tension de grille, à faible (0,1 V) et forte (1 V) tensions de drain, pour une longueur de canal L de 1 μm , une hauteur de silicium H_{Si} de 60 nm, une largeur de silicium W_{Si} de 20 nm, une épaisseur d'oxyde supérieure T_{ox_top} de 50 nm, une épaisseur d'oxyde latérale T_{ox} de 1,5 nm, et une épaisseur d'oxyde enterré T_{BOX} de 30 nm. Pour cette structure en particulier, nous trouvons les quantités équivalentes suivantes : $W_{eq} = 15$ nm, $H_{eq} = 80$ nm, $T_{ox_eq} = 1,88$ nm, et $I_{SP} = 4 \cdot \mu \cdot \epsilon_{ox} / T_{ox_eq} \cdot U_T^2 \cdot H_{eq} / L$.

Deux différents modèles sont comparés aux simulations 3D de courant. Si nous considérons le DG FinFET seulement comme un DG MOSFET planaire, i.e. nous ignorons les interfaces d'oxyde épais, nous observons que la densité de courant est légèrement sous-estimée par le modèle DG en forte inversion. Cette différence vient du courant s'établissant aux interfaces supérieure et inférieure que nous devrions prendre en compte. D'autre part, le modèle incluant une capacité équivalente d'oxyde dans une topologie de QG (qui est proche de celle du FinFET) ne montre pas cette déviation. Ce modèle est en très bon accord avec les simulations TCAD 3D. Cela confirme que les courants additionnels attribués aux interfaces d'oxyde épais doivent être intégrés au modèle. Il est à noter que ces courants sont pris en compte sans introduire de paramètre empirique ou fonction de lissage. Il est également intéressant de remarquer que l'épaisseur équivalente d'oxyde est impliquée dans tous les régimes d'opération du dispositif, mais sans affecter le régime de faible inversion comparativement avec le modèle DG. En effet, la capacité d'oxyde de grille n'influe pas sur la densité de charges en faible inversion, mais seulement sur celle en forte inversion au-dessus de la tension de seuil.

La Figure 4.9 montre les résultats obtenus avec le modèle du DG FinFET pour des géométries avec un film de silicium qui passe d'étroit à large, i.e. au plus large ayant une largeur de film proche de la hauteur de silicium. Pour le TG MOSFET, le courant augmenté des courants aux

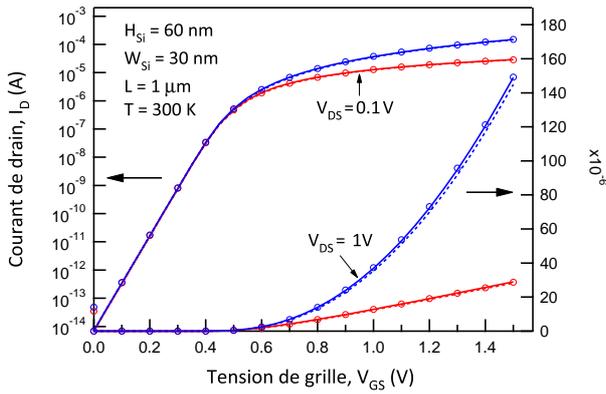


Figure 4.10 – Courant de drain d'un TG MOSFET en fonction de la tension de grille en régimes linéaire et saturé. Simulations 3D : symboles, modèle TG : lignes pointillées, modèle QG avec la capacité équivalente d'oxyde : lignes pleines.

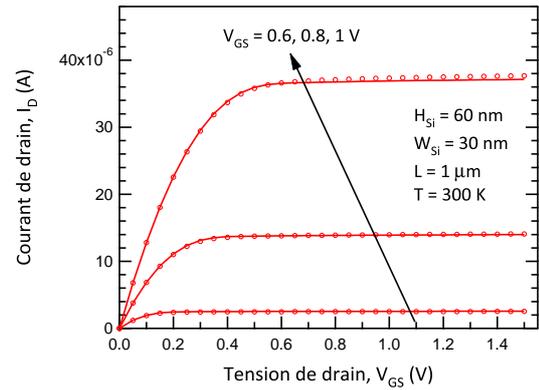


Figure 4.11 – Courant de drain d'un TG MOSFET en fonction de la tension de drain en régimes linéaire et saturé. Simulations 3D : symboles, modèle QG avec la capacité équivalente d'oxyde : lignes.

interfaces d'oxyde épais, est bien pris en compte par la capacité de grille frangeante proposée ci-dessus (Figures 4.10 et 4.11).

Pour le DG FinFET et pour le TG MOSFET, la combinaison des capacités planaires et frangeantes est ainsi pleinement justifiée et valide le concept d'une capacité équivalente d'oxyde pour la modélisation compacte des MOSFET multigrilles de géométries quelconques.

4.2.5 Le cas du MOSFET bulk triple grille

Le transistor FinFET peut également être fabriqué à partir de wafers silicium bulk [15]. Comparativement à la technologie SOI, la technologie bulk réduit le coût du wafer, et donc de la fabrication du transistor, diminue la densité de défauts du dispositif, permet un plus haut taux

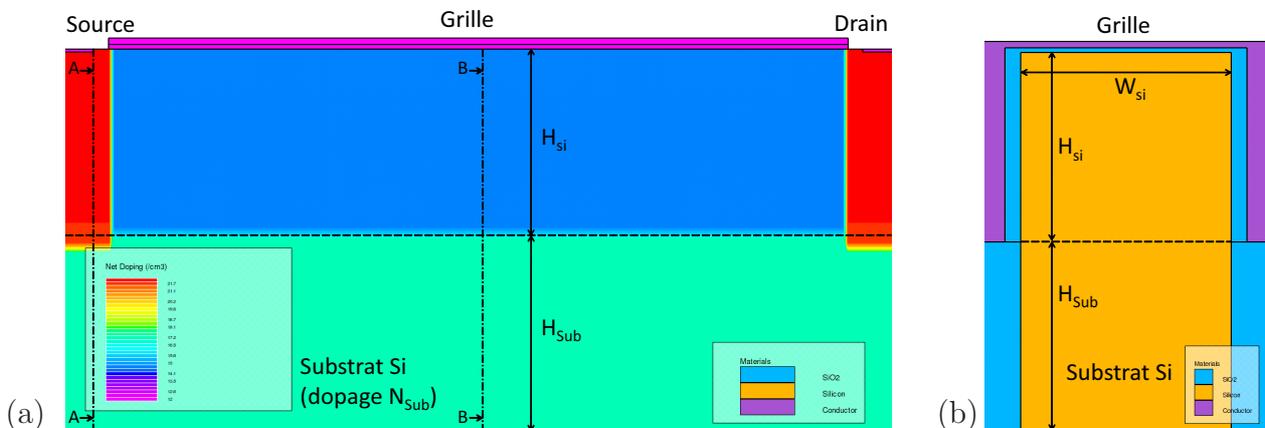


Figure 4.12 – (a) Coupe longitudinale du TG MOSFET bulk portant l'information de dopage du dispositif, (b) Coupe transversale du TG MOSFET bulk, pour la géométrie $H_{Si} = 60 \text{ nm}$, $W_{Si} = 20 \text{ nm}$, $L = 1 \mu\text{m}$, $H_{Sub} = 60 \text{ nm}$ et $t_{ox} = 1.5 \text{ nm}$ (Remarque : les échelles verticale et horizontale des coupes sont différentes, l'épaisseur d'oxyde de grille est bien la même sur les trois grilles).

de transfert de la chaleur, élimine l'effet de potentiel flottant dans le canal, tout en permettant une réduction d'échelle du transistor tout aussi importante [16]. L'inconvénient majeur est la perte de courant de source à drain à travers le substrat (« Punch-through » en anglais), sur lequel se concentre les efforts de recherche sur le processus de fabrication du transistor FinFET bulk.

La Figure 4.12 représente une section longitudinale et une section transversale d'un MOSFET TG bulk. Le principe de fonctionnement du transistor MOSFET TG bulk est le même que celui du MOSFET TG SOI. Il est constitué d'une grille métal mid-gap et d'un canal non dopé ou faiblement, impliquant un canal complètement déplété. Le substrat peut lui aussi être faiblement dopé, ou l'être fortement dans le but de réduire le courant du substrat entre source et drain. Ainsi, nous pouvons envisager d'adapter le modèle compact multigrille présenté jusqu'ici, pour modéliser le transistor MOSFET TG bulk.

4.2.5.1 Etude du courant du MOSFET TG bulk pour différents substrats

A la Figure 4.13, la simulation d'un TG sans substrat, ni BOX, est comparée à deux types de simulations TCAD 3D de TG MOSFET bulk, l'un avec un substrat faiblement dopé (du même niveau de dopage que le canal) et l'autre avec un fort dopage du substrat. Nous constatons que le niveau de dopage du substrat influence le courant de drain, tout autant en faible inversion qu'en forte inversion. A faible dopage de substrat, le courant en inversion volumique s'étend dans le substrat. En effet, selon les profils de concentration d'électrons de la Figure 4.14, le réservoir d'électrons de la source s'étend dans le substrat, et la concentration d'électrons au centre du canal ne diminue que légèrement au fur et à mesure que l'on descend dans le volume du substrat. Cela est confirmé par la constatation à la Figure 4.13, d'un courant à faibles tensions de grille plus important pour la hauteur de substrat $H_{Sub} = 500$ nm, par rapport au cas d'une hauteur plus faible à 200 nm, et encore plus par rapport au cas sans substrat. En terme de modélisation, nous pouvons imaginer dans un premier temps, fixer la hauteur H_{Si} du modèle comme la somme de la hauteur des électrodes S/D avec celle du substrat. Cependant, l'augmentation de la hauteur de substrat implique une dégradation de la pente sous-le-seuil observé sur la Figure 4.13. La hauteur H_{Si} du modèle devrait alors dépendre de la géométrie du substrat.

4.2.5.2 Modélisation par des corrections géométriques des paramètres équivalents

Les constatations faites dans ce paragraphe sont toutes basées sur les résultats de simulations numériques de la Figure 4.13. En forte inversion, si nous considérons un dopage du substrat identique à celui du canal, nous constatons un surplus de courant du TG bulk, comparé au dispositif purement TG, c'est-à-dire sans substrat, ni BOX. Tout comme dans le TG SOI, un champ électrique entre la grille et le silicium à travers l'oxyde enterré, crée un canal d'inversion dans le substrat à faible profondeur. Nous pouvons ainsi assimiler cette augmentation du courant de forte inversion à une prolongation du canal d'inversion le long de l'oxyde enterré, soit en terme de modélisation, une augmentation de la hauteur équivalente de valeur ΔH_{eq} .

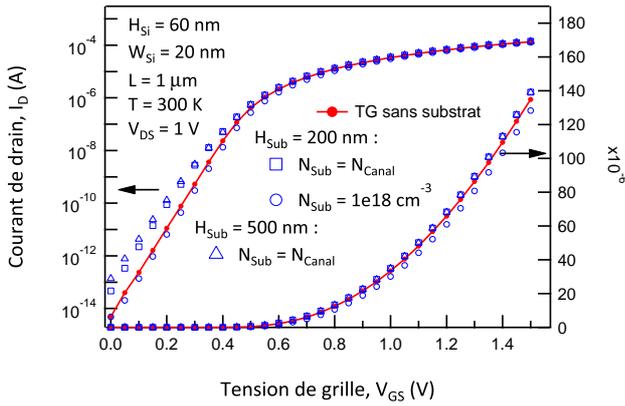


Figure 4.13 – Courant de drain d’un TG MOSFET bulk en fonction de la tension de grille, pour deux hauteurs de substrat faiblement dopé, et, pour deux niveaux de dopage, un faible et un fort, à une même hauteur de substrat. Ces courants de drain de TG bulk sont comparés au courant de drain d’un TG sans substrat.

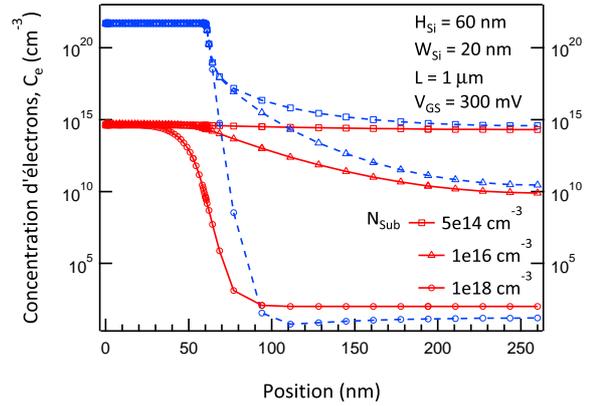


Figure 4.14 – Influence du dopage du substrat sur le profil de concentration d’électrons, de la grille supérieure à l’extrémité inférieure du substrat. Lignes bleues : profil selon la coupe A-A, lignes rouges : profil selon la coupe B-B.

A fort dopage du substrat, au contraire, la concentration d’électrons en régime de faible inversion est proche de zéro dans tout le volume du substrat. Elle diminue même dans une partie du canal, à l’interface avec la région du substrat. La pénétration de la forte concentration de trous du substrat fortement dopé P, réduit la région dans laquelle peut s’établir une inversion du semi-conducteur. Ainsi, en faible inversion la hauteur du canal sur laquelle s’établit le courant de drain est réduite comparée à la hauteur des régions dopées formant les électrodes S/D. Cette conséquence du dopage du substrat, peut être prise en compte dans le modèle TG par une variation de la hauteur H_{Si} que nous nommons ΔH_{Si} .

En forte inversion, nous constatons une réduction du courant par rapport au pur TG, suggérant une réduction de la hauteur équivalente, contrairement au cas faiblement dopé. Même à de grandes tensions de grille, là où l’inversion est forte, il ne peut pas s’établir d’inversion des charges à l’interface canal/oxyde de grille, au niveau de l’interface avec le substrat. En terme de modélisation, cela implique que la correction ΔH_{eq} doit être de valeur négative.

4.2.5.3 Résultats de modélisation

La Figure 4.15 nous montre la comparaison du courant de drain en fonction de la tension de grille, entre les simulations 3D de TG MOSFET bulk et le modèle TG dont les paramètres géométriques H_{Si} et H_{eq} ont été corrigés en fonction du dopage du substrat. A substrat faiblement dopé, nous simulons le modèle TG avec $\Delta H_{Si} = H_{Sub}$ et $\Delta H_{eq} = 3$ nm. Nous observons une tension de seuil similaire entre la simulation et le modèle, mais un léger désaccord au niveau de la pente sous-le-seuil, celle du modèle restant à la valeur idéale de 60 mV/dec d’un canal long. En forte inversion, la correction de la hauteur équivalente permet un très bon accord du modèle compact. Pour un substrat dopé à $N_{Sub} = 10^{18}$ cm⁻³, $\Delta H_{Si} = -25$ nm et $\Delta H_{eq} = -3,5$ nm. Avec cette correction géométrique, la tension de seuil est correctement modélisée, ainsi

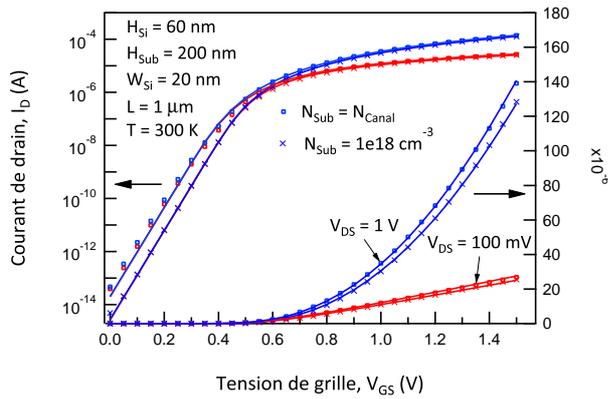


Figure 4.15 – Courant de drain d'un TG MOSFET bulk en fonction de la tension de grille en régime linéaire et saturé, en échelle linéaire et logarithmique, pour une profondeur de substrat de 200 nm et deux niveaux de dopage. Simulations 3D : symboles, modèle TG avec corrections des paramètres géométriques : lignes pleines.

que la pente sous-le-seuil. Au delà d'un dopage de substrat relativement faible de 10^{16} cm^{-3} , la pente sous-le-seuil n'est pas affectée par rapport à un TG SOI, seule la tension de seuil varie avec le dopage. En forte inversion, la simple correction de la hauteur équivalente est suffisante pour très bien prendre en compte le courant de la simulation numérique.

Synthèse des corrections géométriques en fonction du niveau de dopage du substrat

La figure 4.16 donne les valeurs de ΔH_{Si} et ΔH_{eq} sur la gamme de dopage à laquelle le modèle compact TG corrigé peut s'appliquer. Ces valeurs de corrections sont extraites des comparaisons entre les simulations numériques et le modèle, de telle sorte que le modèle concorde avec les simulations. L'extraction se fait indépendamment, en comparant le courant en faible inversion pour ΔH_{Si} , et en comparant le courant en forte inversion pour ΔH_{eq} .

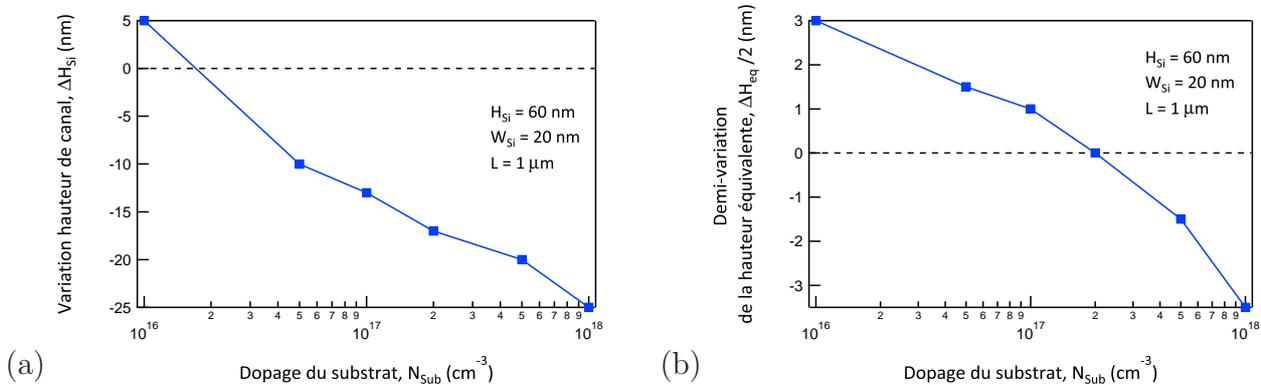


Figure 4.16 – Correction dans le modèle purement triple grille du modèle multigrille de (a) la hauteur du canal et de (b) la demi-variation de la hauteur équivalente en fonction du niveau de dopage du substrat.

4.3 Modélisation de l'effet de la température

Dans cette partie nous proposons une extension vers la dépendance à la température du modèle multigrille. Nous avons montré à la Section 2.2.3 que le modèle du transistor MOSFET double grille, sur lequel est basé le modèle multigrille, est valide lorsque la température varie si nous considérons une mobilité constante. Notre but est ici de lui ajouter un modèle de mobilité dépendant du champ électrique transverse et de la température.

Le modèle de mobilité que nous proposons est adapté du modèle de Jeon et Burk [17]. Ce modèle est basé sur la somme des réciproques des trois mécanismes de dispersions (loi de Matthiessen), que sont les dispersions par les phonons, coulombienne et de rugosité de surface. Le modèle dépend explicitement de la température T par les phénomènes de dispersions par les phonons et les interactions coulombiennes, et du champ électrique transversal E_{eff} par les phénomènes de dispersions par les phonons et de rugosité de surface. A haute température, la dispersion coulombienne est négligeable [17], ainsi nous gardons seulement les termes de dispersion par les phonons et de rugosité de surface, aboutissant selon [17] à :

$$\mu = \frac{1}{a_2 \cdot T^n \cdot E_{\text{eff}}^{1/\gamma} + a_3 \cdot E_{\text{eff}}^2} \quad (4.37)$$

Cependant, en faible inversion, comme le champ électrique tend vers zéro, la mobilité devrait diverger, alors que nous attendons à ce qu'elle atteigne une valeur asymptotique en champ faible. Pour supprimer cette incohérence, un nouveau paramètre a_0 a été introduit. Après transformation, nous obtenons :

$$\mu = \frac{a_0^{-1} \cdot T^{-n}}{1 + E_0 \cdot E_{\text{eff}}^{1/\gamma} + E_1 \cdot E_{\text{eff}}^2 \cdot T^{-n}} \quad (4.38)$$

où E_{eff} est le champ électrique transverse calculé comme dans [18], E_0 et E_1 sont des paramètres empiriques, la mobilité à champ faible est $\mu_0 = a_0^{-1} \cdot T^{-n}$, avec $\gamma = 3-6$ et $n = 1-1.5$ d'après [17].

Concernant l'effet de saturation, nous adoptons le même modèle que celui que nous avons introduit dans [18] qui dépend de la densité de charge du canal, de la mobilité et de la vitesse de saturation. Afin d'avoir un modèle dépendant complètement de la température, nous suivons l'approche de Quay et al. [19] qui donne une dépendance à la température pour la vitesse de saturation.

4.4 Paramètres du modèle multigrille

Le tableau 4.2 montre tous les paramètres qui doivent être extraits pour les dispositifs multigrilles à canaux longs.

4.5 Validation

Nous présentons une validation du modèle multigrille avec des simulations TCAD 3D de DG FinFETs, de TG et GAA MOSFET, pour des températures variant de 275 K à 425 K. Finalement, le modèle sera évalué avec des mesures électriques d'un TG MOSFET pour des températures entre 273.15 K et 373.15 K.

Tableau 4.2 – Description des paramètres du modèle.

Effet	Nom	Description	Unité
Mobilité	n	Coefficient de température pour la modélisation de l'effet de dispersion par les phonons	-
	γ	Coefficient associé au champ électrique pour la modélisation de l'effet de dispersion par les phonons	-
	a_0	Coefficient de la mobilité à champ faible	$V.s^2.m^{-1}.K^{-n}$
	E_0	Coefficient de réduction de la mobilité pour la modélisation de l'effet de dispersion par les phonons	$(V.m^{-1})^{-\frac{1}{\gamma}}$
	E_1	Coefficient de réduction de la mobilité pour la modélisation de l'effet de dispersion de rugosité de surface	$(V.m^{-1})^{-2}.K^n$
Modulation de longueur de canal	θ	Coefficient de la tension de saturation du drain	-

4.5.1 Validation TCAD

Nous avons étudié le comportement de dispositifs multigrilles pour des températures entre 275 K et 425 K. Dans le modèle de mobilité, les coefficients de la dispersion par les phonons sont $n = 1.3$ et $\gamma = 2$, et la mobilité à champ faible extraite à 300 K est $\mu_0 = 1350 \text{ cm}^2/\text{V.s}$. La Figure 4.17 compare des simulations TCAD 3D avec le modèle pour le courant en fonction de la tension de grille d'un GAA MOSFET de 20 nm de rayon de silicium (pour à la fois $V_{DS} = 0,1 \text{ V}$ et 1 V). Les dégradations de la pente sous-le-seuil, de la tension de seuil et de la mobilité à champ faible lorsque la température augmente, sont prédites avec justesse. En forte inversion, la dégradation de la mobilité avec la température est bien prise en compte pour de faibles et grandes tensions de drain. La Figure 4.18 confirme que le modèle est toujours en bon accord avec les simulations 3D en ce qui concerne le courant de drain en fonction de la tension de drain à 400 K. Les paramètres de mobilité du modèle aux Figures 4.17 et 4.18 sont fixés à : $E_0 = 1.58 \times 10^7 \text{ V/m}$, $E_1 = 1000 \times 10^{14} \text{ (V/m)}^2$.

Les paramètres électriques importants du point de vue du design sont comparés entre les simulations 3D TCAD et le modèle. Ainsi nous avons extrait le courant de drain pour le transistor à l'état OFF I_{off} , celui à l'état ON I_{on} , la tension de seuil V_{th} et la pente sous-le-seuil S_S , selon des méthodes décrites au paragraphe 5.3.3. Ces paramètres électriques sont tracés en fonction de la température aux Figures 4.20, 4.21 et 4.22, respectivement pour des DG FinFETs, des TG MOSFET et des GAA MOSFET. Nous pouvons ainsi constater une très bonne estimation faite par le modèle des paramètres issus des simulations TCAD, et ce sur toute la gamme de température de notre étude.

La précision du modèle pour différentes géométries, i.e. des DG FinFETs, des TG et GAA MOSFET, est résumée dans le tableau 4.3.

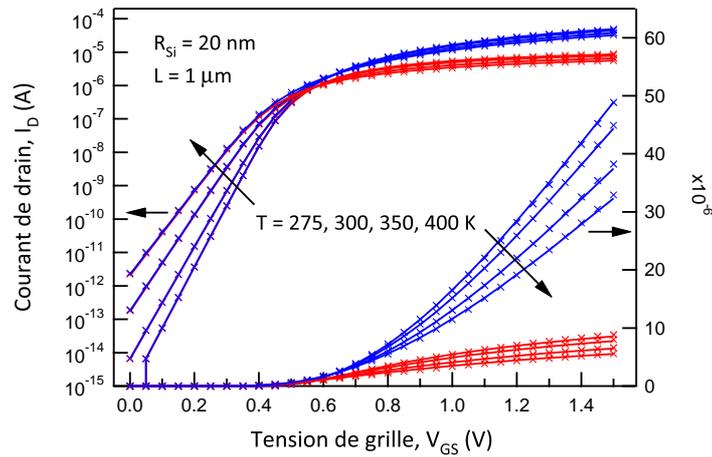


Figure 4.17 – Courant de drain d'un GAA MOSFET en fonction de la tension de grille en régime linéaire et saturé. Simulations 3D : croix, modèle : lignes.

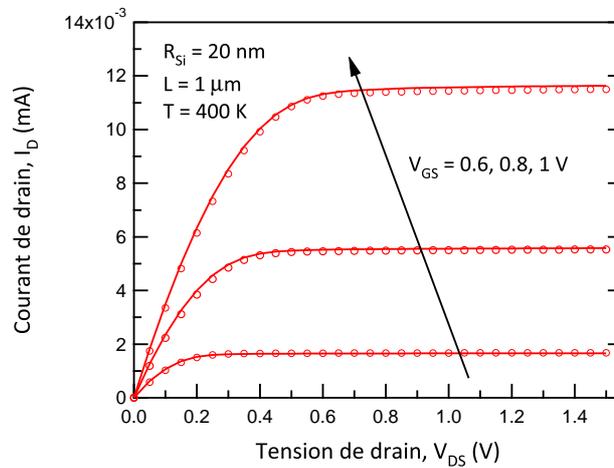


Figure 4.18 – Courant de drain d'un GAA MOSFET en fonction de la tension de drain en régime linéaire et saturé. Simulations 3D : cercles, modèle : lignes.

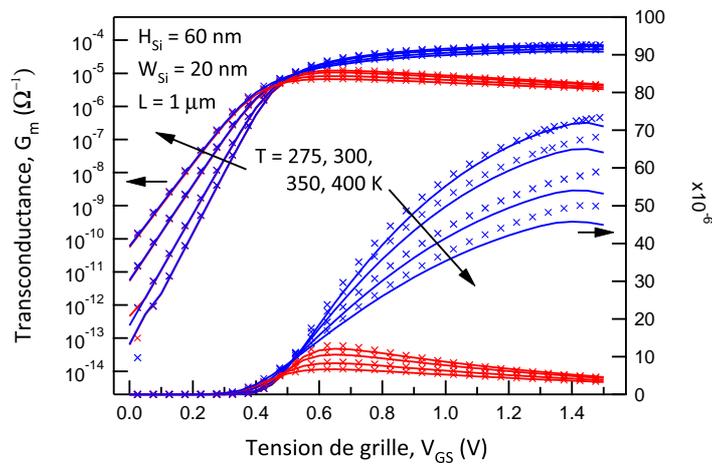


Figure 4.19 – Transconductance d'un TG MOSFET en fonction de la tension de grille en régime linéaire et saturé. Simulations 3D : croix, modèle : lignes. $E_0 = 1.46 \times 10^7$ V/m, $E_1 = 500 \times 10^{14}$ (V/m)².

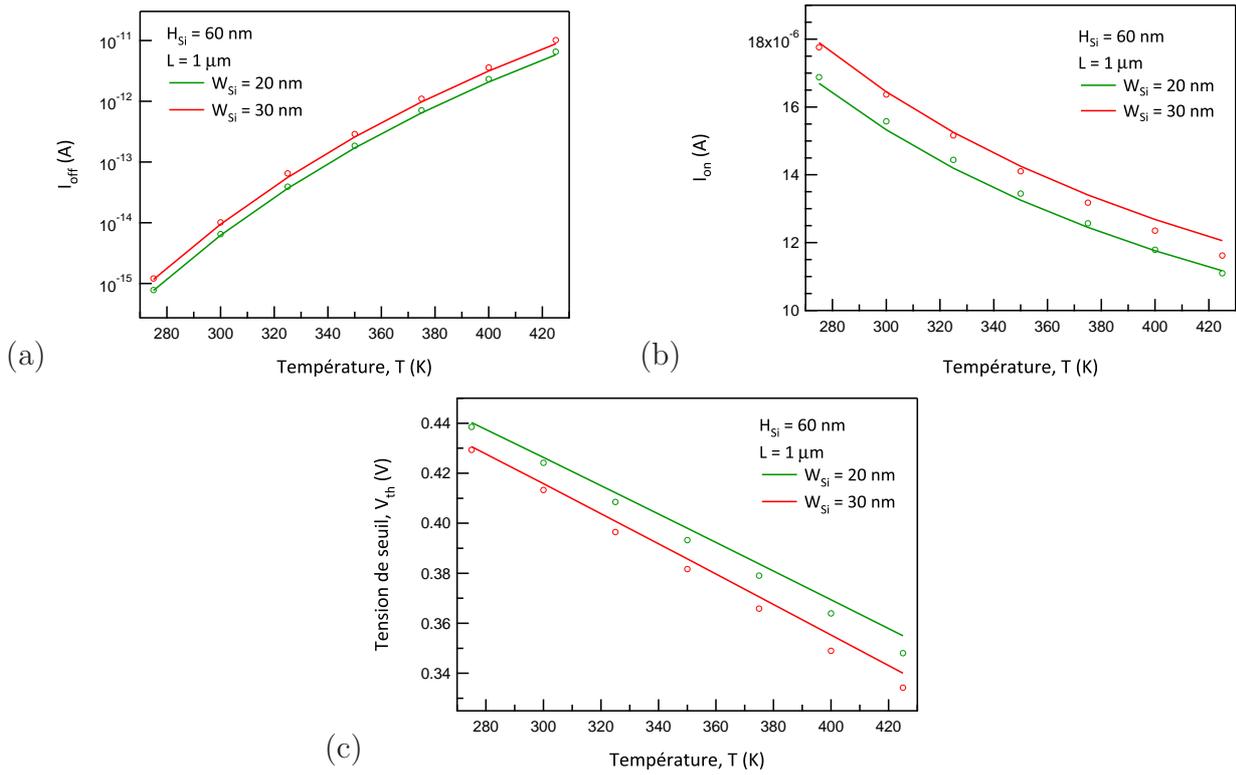


Figure 4.20 – (a) Courant de drain I_{off} en échelle logarithmique, (b) Courant de drain I_{on} et (c) Tension de seuil V_{th} en échelle linéaire, de DG FinFET en fonction de la température pour différentes largeur de canal. Simulations 3D : cercles, modèle : lignes.

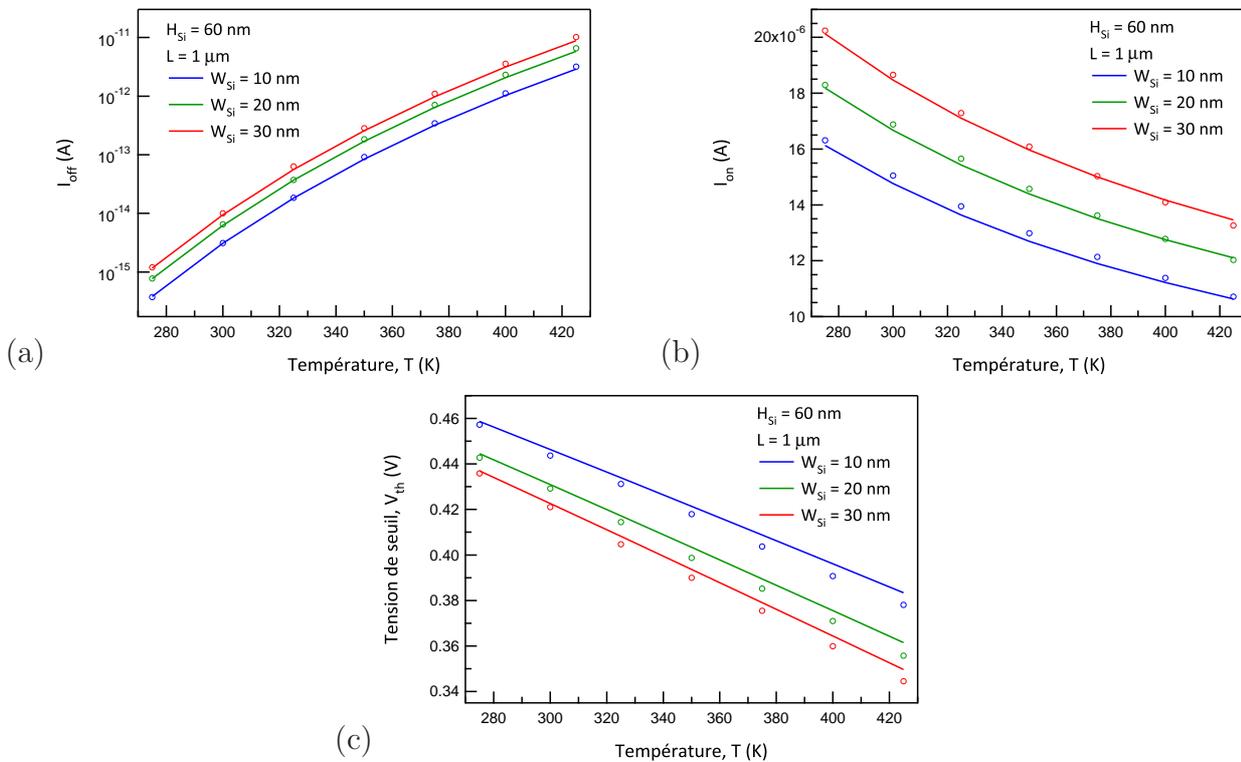


Figure 4.21 – (a) Courant de drain I_{off} en échelle logarithmique, (b) Courant de drain I_{on} et (c) Tension de seuil V_{th} en échelle linéaire, de TG MOSFET en fonction de la température pour différentes largeur de canal. Simulations 3D : cercles, modèle : lignes.

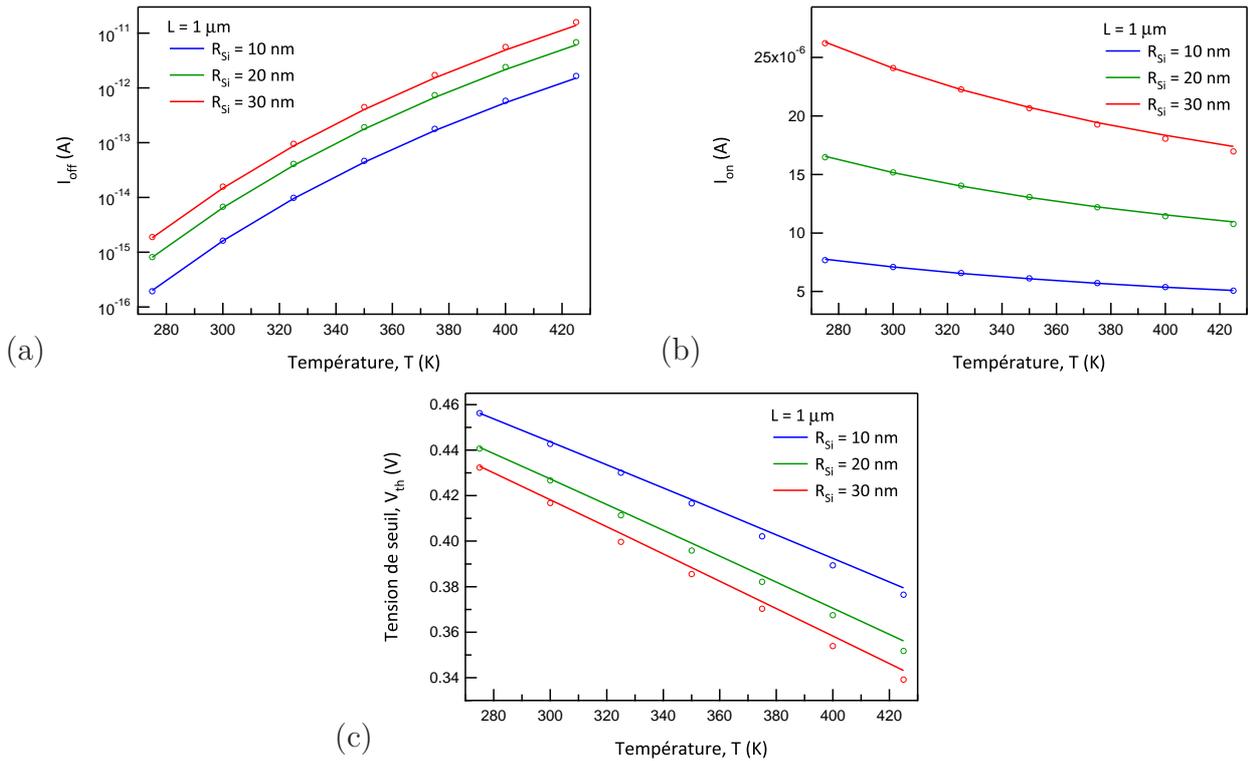


Figure 4.22 – (a) Courant de drain I_{off} en échelle logarithmique, (b) Courant de drain I_{on} et (c) Tension de seuil V_{th} en échelle linéaire, de GAA MOSFET en fonction de la température pour différents rayons du canal de silicium. Simulations 3D : cercles, modèle : lignes.

Tableau 4.3 – Erreurs moyenne relative (en %) entre les simulations 3D TCAD et le modèle, de courants I_{on} et I_{off} sur une plage de température de 275 K à 425 K. Ces résultats sont donnés pour différentes géométries de DG FinFET, de TG et GAA MOSFET.

Géométrie		I_{on}	I_{off}	V_{th}
DG FinFET	$W_{Si} = 20$ nm	1.12	0.68	1.11
	$W_{Si} = 30$ nm	1.85	0.76	1.13
Triple grille	$W_{Si} = 10$ nm	1.40	0.52	0.91
	$W_{Si} = 20$ nm	0.89	0.78	0.94
	$W_{Si} = 30$ nm	0.95	0.52	0.87
GAA	$R_{Si} = 10$ nm	0.51	0.32	0.46
	$R_{Si} = 20$ nm	0.86	0.42	0.62
	$R_{Si} = 30$ nm	0.98	0.53	0.77

4.5.2 Validation par rapport aux mesures expérimentales

4.5.2.1 Caractérisation

Des mesures ont été faites dans le cadre du projet Européen COMON sur des TG MOSFET ayant des longueurs de grille entre 70 nm et $1 \mu\text{m}$, une largeur de canal W_{Si} de 30 nm et une hauteur de canal H_{Si} de 60 nm. Les dispositifs ont été fabriqués avec une électrode de grille midgap TiN selon un empilement de grille SiON [4]. L'épaisseur d'oxyde est approximativement de 1,9 nm. Les résultats montrés ci-dessous sont pour une longueur de dispositif de $1 \mu\text{m}$. En suivant la méthode développée dans [20], nous avons extrait une longueur de canal effective L_{eff}

$= 0,9 \mu\text{m}$ et une résistance série $R_{SD_{298K}} = 1300 \Omega$ à température ambiante. Nous établissons une loi de variation linéaire de la résistance série avec la température en extrayant la résistance série pour différentes températures. Nous obtenons ainsi :

$$R_{SD}(T) = R_{SD_{298K}} \cdot (1 + a_T \cdot (T - 298)) \text{ avec } a_T = 0.0108. \quad (4.39)$$

4.5.2.2 Comparaisons entre le modèle et les mesures

Afin de pouvoir comparer les résultats du modèle multigrille aux mesures, il est nécessaire d'extraire la mobilité à champ faible et les coefficients du modèle de mobilité lié à sa dépendance à la température. Nous trouvons $n = 1.3$, $\gamma = 2$ et $\mu_0 = 1350 \text{ cm}^2/\text{V.s}$.

La Figure 4.23 compare les mesures et le modèle multigrille pour le courant de drain en fonction de la tension de grille à $V_{DS} = 0.05 \text{ V}$ et 1 V , et pour des températures entre 273 K et 373 K . La pente sous-le-seuil et la tension de seuil prédites par le modèle sont en bon accord avec les mesures sur tout l'intervalle de température. Pour des tensions de grille proches de 0 V , nous observons par contre des désaccords entre le modèle et les mesures qui sont attribués aux mécanismes de perte de courant non inclus dans le modèle.

Plus probablement, à haute température, le courant I_{off} est dominé par la perte du canal (dans ce cas la perte de courant au drain est la même que celle à la source). Cependant, plusieurs autres mécanismes de perte existent tel que la perte au drain induite par la grille (GIDL) [21], la perte de grille et la perte par effet tunnel assistée par les pièges (TAT) [22]. En forte inversion, les bonnes prédictions du modèle pour une faible et grande tension de drain justifie l'expression de mobilité dépendante de la température adoptée. La Figure 4.24 montre le courant de drain en fonction de la tension de drain pour $V_{GS} = 1 \text{ V}$, à différentes températures. Le régime de saturation est bien modélisé, comme l'est le courant dans la région linéaire, signifiant ainsi que la variation de la résistance série avec la température est raisonnablement prise en compte.

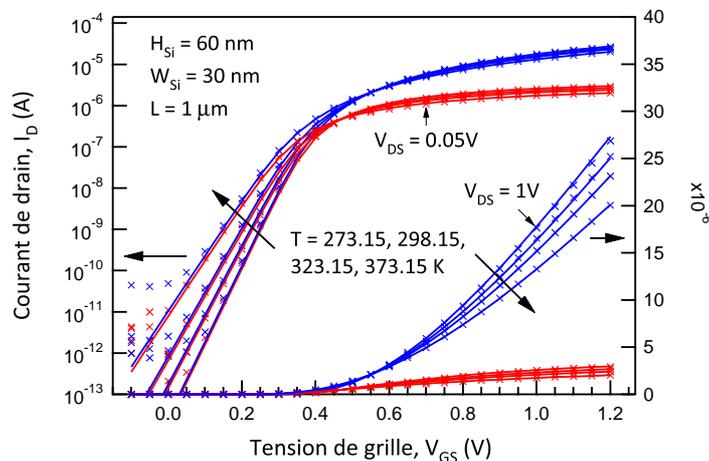


Figure 4.23 – Courant de drain d'un TG MOSFET en fonction de la tension de grille en régime linéaire et saturé. Mesures : croix, modèle : lignes. $E_0 = 1.0 \times 10^7 \text{ V/m}$, $E_1 = 150 \times 10^{14} \text{ (V/m)}^2$.

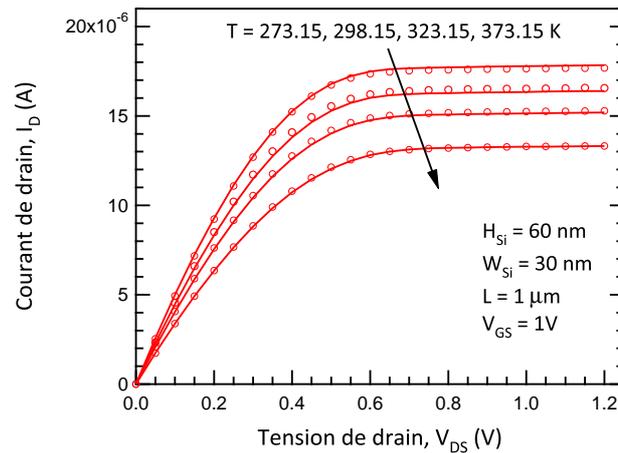


Figure 4.24 – Courant de drain d’un TG MOSFET en fonction de la tension de drain en régime linéaire et saturé. Mesures : cercles, modèle : lignes. $E_0 = 1.0 \times 10^7$ V/m, $E_1 = 150 \times 10^{14}$ (V/m)².

Conclusion

Dans ce chapitre, nous avons démontré comment une topologie de transistors à effet de champ multigrille ayant une unique électrode de grille, peut être simplement considérée comme un DG MOSFET symétrique à travers la définition d’une largeur équivalente de silicium, d’une capacité équivalente d’oxyde de grille et d’une hauteur équivalente de canal. Ce travail a été publié dans [23]. Toutes ces transformations sont explicites et font uniquement appel à des paramètres physiques et technologiques, et à aucune relation empirique. La capacité du modèle à prédire la densité de charges et le courant jusqu’à 425 K, confirme les bases physiques assez fortes de cette approche. De plus, par l’introduction d’un modèle de mobilité dépendant de la température, indispensable pour les dispositifs réels, le modèle est capable de simuler des mesures d’un TG MOSFET en température. De par ces équivalences entre le DG planaire et les transistors multigrilles, toute géométrie peut maintenant être modélisée avec un formalisme commun. Un travail additionnel est maintenant nécessaire pour prendre en compte les effets canaux courts, les courants de grille et effets de quantification dans les transistors multigrilles. Une première étude sur la modélisation des effets canaux courts du transistor MOSFET triple grille a été effectuée et publiée dans [24].

Références

- [1] B. YU, M. LU Huaxin Liu et Y. TAUR, “Explicit continuous models for double-gate and surrounding-gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 54, n° 10, pp. 2715–2722, 2007.
- [2] J.-M. SALLESE, F. KRUMMENACHER, F. PRÉGALDINY, C. LALLEMENT, A. ROY et C. ENZ, “A design oriented charge-based current model for symmetric DG MOSFET and its correlation with the EKV formalism”, *Solid-State Electron.*, vol. 49, n° 3, pp. 485–489, 2005.
- [3] J.-P. COLINGE, “Multiple-gate SOI MOSFETs”, *Solid-State Electron.*, vol. 48, n° 6, pp. 897–905, 2004.

-
- [4] K. von ARNIM, E. AUGENDRE, A. C. PACHA, T. SCHULZ, K. T. SAN, F. BAUER, A. NACKAERTS, R. ROOYACKERS, T. VANDEWEYER, B. DEGROOTE, N. COLLAERT, A. DIXIT, R. SINGANAMALLA, W. XIONG, A. MARSHALL, C. R. CLEAVELIN, K. SCHRUFER et M. JURCZAKTRIVEDI, “Low-power multi-gate FET CMOS technology with 13.9 ps inverter delay, large-scaled integrated high performance digital circuits and SRAM”, dans *Proc. IEEE Symposium on VLSI Technology*, 2007, pp. 106–107.
- [5] *Announce d’intel pour la production de microprocesseurs avec la technologie FinFET 22 nm.* adresse : http://newsroom.intel.com/community/intel_newsroom/blog/2011/05/04/intel-reinvents-transistors-using-new-3-d-structure.
- [6] B. YU, J. SONG, Y. YUAN, W.-Y. LU et Y. TAUR, “A unified analytic drain-current model for multiple-gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 55, n° 8, pp. 2157–2163, 2008.
- [7] R. RITZENTHALER, F. LIME, B. IÑIGUEZ, E. MIRANDA, F. MARTINEZ, F. PASCAL, M. VALENZA, O. FAYNOT et S. CRISTOLOVEANU, “Analytical modeling of multiple-gate MOSFETs”, dans *Proc. 8th Spanish Conference on Electron Devices (CDE)*, 2011, pp. 1–4.
- [8] W. XIONG, J. PARK et J. COLINGE, “Corner effect in multiple-gate SOI MOSFETs”, dans *Proc. IEEE Int. SOI Conference*, 2003, pp. 111–113.
- [9] J.-M. SALLESE, N. CHEVILLON, F. PRÉGALDINY, C. LALLEMENT et B. IÑIGUEZ, “The equivalent-thickness concept for doped symmetric DG MOSFETs”, *IEEE Trans. Electron Devices*, vol. 57, n° 11, pp. 2917–2924, 2010.
- [10] K. E. MOSELUND, D. BOUVET, L. TSCHUOR, V. POTT, P. DAINESI et A. IONESCU, “Local volume inversion and corner effects in triangular gate-all-around MOSFETs”, dans *Proc. European Solid-State Device Research conference (ESSDERC)*, 2006, pp. 359–362.
- [11] B. IÑIGUEZ, D. JIMÉNEZ, J. ROIG, H. A. HAMID, L. F. MARSAL et J. PALLARÈS, “Explicit continuous model for long-channel undoped surrounding gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 52, n° 8, pp. 1868–1873, 2005.
- [12] P. CHAMBRÉ, “On the solution of the poisson-boltzmann equation with application to the theory of thermal explosions”, *The Journal of Chemical Physics*, vol. 20, n° 11, 1952.
- [13] A. ROY, C. ENZ et J.-M. SALLESE, “Compact modeling of gate sidewall capacitance of DG-MOSFET”, *IEEE Trans. Electron Devices*, vol. 53, n° 10, pp. 2655–2657, 2006.
- [14] T. ERNST, R. RITZENTHALER, O. FAYNOT et S. CRISTOLOVEANU, “A model of fringing fields in short-channel planar and triple-gate SOI MOSFETs”, *IEEE Trans. Electron Devices*, vol. 54, n° 6, pp. 1366–1375, 2007.
- [15] T.-S. PARK, H. J. CHO, J. D. CHOE, S. Y. HAN, D. PARK, K. KIM, E. YOON et J.-H. LEE, “Characteristics of the full CMOS SRAM cell using body-tied TG MOSFETs (bulk FinFETs)”, *IEEE Trans. Electron Devices*, vol. 53, n° 3, pp. 481–487, 2006.
- [16] D. SHAMIRYAN, A. REDOLFI et W. BOULLART, “Dry etching process for bulk finFET manufacturing”, *Microelectronic Engineering*, vol. 86, pp. 96–98, 2009.
- [17] D. JEON et D. BURK, “MOSFET electron inversion layer mobilities - A physically based semi-empirical model for a wide temperature range”, *IEEE Trans. Electron Devices*, vol. 36, n° 8, pp. 1456–1463, 1989.
- [18] A. YESAYAN, F. PRÉGALDINY, N. CHEVILLON, C. LALLEMENT et J.-M. SALLESE, “Physics-based compact model for ultra-scaled FinFETs”, *Solid-State Electron.*, vol. 62, n° 1, pp. 165–173, 2011.

- [19] R. QUAY, C. MOGLESTU, V. PALANKOVSKI et S. SELBERHERR, “A temperature dependent model for the saturation velocity in semiconductor materials”, dans *Materials Science in Semiconductor*, vol. 3, 2000, pp. 149–155.
- [20] G. NIU, J. CRESSLER, S. MATHEW et S. SUBBANNA, “A channel resistance derivative method for effective channel length extraction in LDD MOSFET’s”, *IEEE Trans. Electron Devices*, vol. 47, n° 3, pp. 648–650, 2000.
- [21] J.-H. CHEN, S.-C. WONG et Y.-H. WANG, “An analytic three-terminal band-to-band tunneling model on GIDL in MOSFET”, *IEEE Trans. Electron Devices*, vol. 48, n° 7, pp. 1400–1405, 2001.
- [22] T.-E. CHANG, C. HUANG et T. WANG, “Mechanisms of interface trap-induced drain leakage current in off-state n-MOSFET’s”, *IEEE Trans. Electron Devices*, vol. 42, n° 4, pp. 738–743, 1995.
- [23] N. CHEVILLON, J.-M. SALLESE, C. LALLEMENT, F. PRÉGALDINY, M. MADEC, J. SELDMER et J. AGHASSI, “Generalization of the concept of equivalent thickness and capacitance to multigate MOSFETs modeling”, *IEEE Trans. Electron Devices*, vol. 59, n° 1, pp. 60–71, 2012.
- [24] N. CHEVILLON, P. F., C. LALLEMENT et J.-M. SALLESE, “Standardization of multigate MOSFET modeling”, dans *IEEE Int. Mixed Design of Integrated Circuits and Systems (MIXDES’12)*, 2012, pp. 78–83.

Chapitre 5

Environnement de travail logiciel

Sommaire

5.1	Simulation numérique de dispositifs MOSFET multigrilles	104
5.1.1	Méthode des éléments finis	104
5.1.2	Modèles et méthodes utilisés	105
5.1.3	Constitution du maillage	107
5.2	Extraction de paramètres du modèle FinFET	109
5.2.1	Caractéristiques électriques pour l'extraction de paramètres	109
5.2.2	Méthodologie de la procédure d'extraction de paramètres de modèle .	111
5.2.3	Procédure complète automatisée d'extraction de paramètres du modèle FinFET	112
5.3	Environnement de travail	113
5.3.1	Plate-forme de simulation numérique et d'extraction des paramètres du modèle	113
5.3.2	Outils pour l'étude du comportement physique des transistors.	116
5.3.3	Méthodes d'extraction de paramètres électriques	116
	Conclusion	117
	Références	117

Tout modèle compact se doit d'être rigoureusement validé par rapport à des données expérimentales. Pendant cette thèse, nous avons pu obtenir expérimentalement des mesures électriques afin de valider le modèle compact du transistor multigrille au Chapitre 4. De telles mesures sont coûteuses, rares (peu de géométries différentes) et difficilement accessibles. La réalisation des mesures sur des transistors MOSFETs triple grille du fondeur allemand Infineon (désormais Intel) à Munich, Allemagne, ont seulement été possibles grâce à notre participation au projet européen COMON. En l'absence de mesures expérimentales, nous faisons appel à la simulation numérique qui se révèle tout aussi intéressante, moins chère et plus souple pour permettre un développement rapide du modèle. Les simulations nous permettent d'étudier les phénomènes physiques qui ont lieu au sein des transistors, et qui nous sont inaccessibles par les mesures (les aspects extrinsèques cachent ou modifient l'observation du comportement in-

trinsèque du dispositif). Tout modèle compact doit également être associé à une procédure d'extraction automatisée de ses paramètres. L'exécution de la procédure d'extraction permet d'obtenir un jeu de paramètres du modèle prédisant les caractéristiques électriques d'une technologie donnée de transistor. Pour être automatisée, l'extraction de paramètres doit être associée à différents programmes informatiques qui manipulent les données nécessaires à la procédure. L'ensemble des logiciels constitue un environnement logiciel que nous décrivons dans ce chapitre.

La première partie de ce chapitre traite de la mise en oeuvre des simulations numériques électriques de transistors MOSFET multigrilles par l'outil de TCAD Silvaco[®]. Nous détaillerons ensuite l'environnement logiciel exploitant la simulation numérique et l'extraction de paramètres, dont les programmes de traitement des données sont des scripts en langage Python.

5.1 Simulation numérique de dispositifs MOSFET multigrilles

La simulation électrique de transistors MOSFET multigrilles, comme pour de nombreux systèmes physiques, fait appel à la méthode des éléments finis pour résoudre numériquement les équations aux dérivées partielles qui régissent le fonctionnement des dispositifs. Sous l'outil Atlas qui est un simulateur électrique de dispositifs chez Silvaco, nous avons réalisé la simulation 2D de transistors MOSFET double grille avec et sans jonctions, et la simulation 3D des transistors FinFET et multigrilles (triple grille, quadruple grille, gate-all-around cylindrique, à section triangulaire).

5.1.1 Méthode des éléments finis

Comme cela a déjà été rappelé plus haut, cette méthode doit permettre la résolution d'un ensemble d'équations aux dérivées partielles décrivant la physique du transistor. On retrouve parmi ces équations l'équation de Poisson.

$$\frac{\delta^2\psi}{\delta x^2} + \frac{\delta^2\psi}{\delta y^2} + \frac{\delta^2\psi}{\delta z^2} = -\frac{\rho}{\varepsilon_r\varepsilon_0}$$

Une solution approchée du système d'équations est obtenue par l'application d'un algorithme mathématique discret sur un domaine géométrique défini et des conditions électriques aux bords de ce domaine. Cette méthode numérique de résolution d'équations continues demande donc une discrétisation qui se matérialise par un maillage du domaine en fragments quelconques, les éléments finis (Figure 5.1). Ils sont usuellement de forme triangulaire ou rectangulaire. Cela pose cependant des questions de qualité des résultats. La solution calculée sera toujours entachée d'une erreur. Elle est d'ailleurs d'autant plus grande que le maillage ne comporte pas assez de points. De plus la convergence vers la solution exacte n'est pas toujours assurée. Il est nécessaire de jouer sur le choix de l'algorithme de calcul et sur sa configuration.

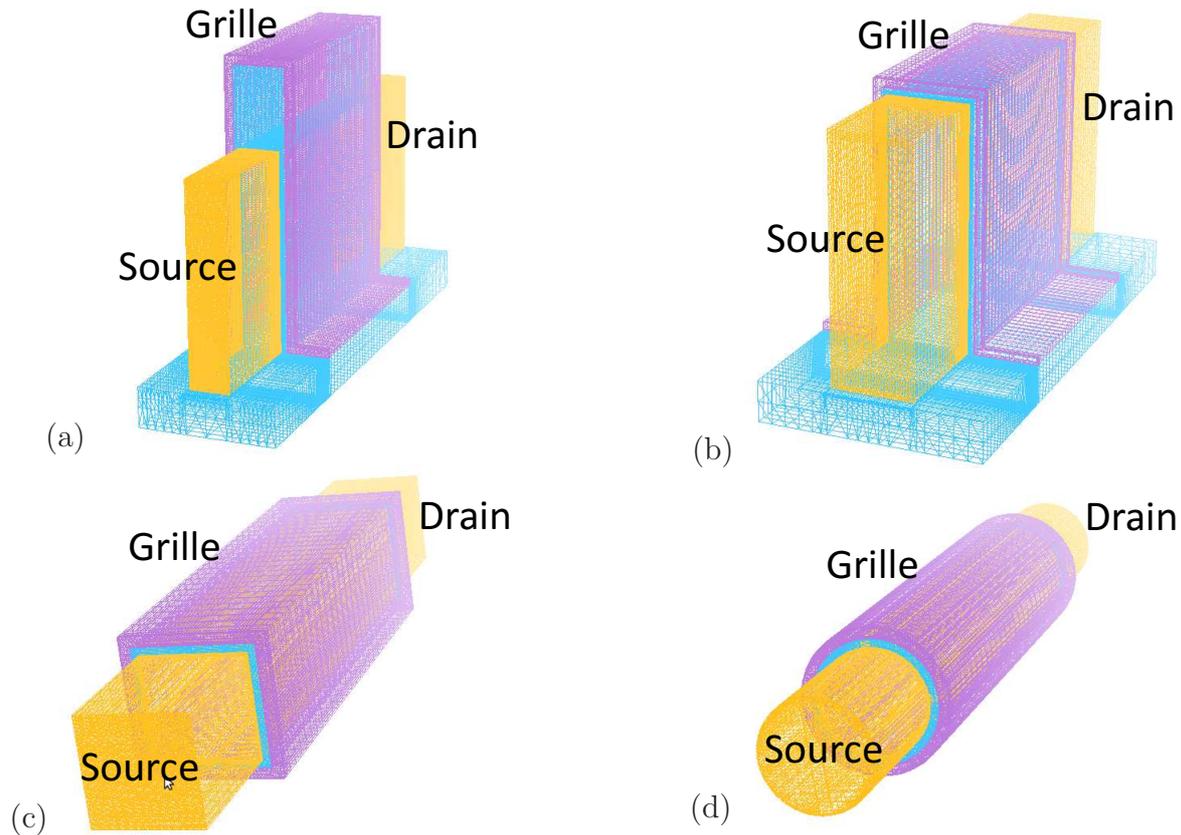


Figure 5.1 – Exemples de vues 3D du maillage (a) du transistor FinFET, et des transistors MOSFET (b) triple grille, (c) quadruple grille et (d) cylindrique.

5.1.2 Modèles et méthodes utilisés

- ▶ Les **modèles physiques** représentent les équations physiques qui sont résolues lors de la simulation. Le simulateur Atlas utilise une configuration de modèles par défaut, donnant la base du système d'équations à résoudre.
- ▷ Le modèle de concentration de porteurs de charges défini est **Fermi-Dirac**. La statistique de Boltzmann utilisé par défaut n'est pas valable pour les régions des électrodes de source et drain dont le dopage est très élevé, et situé à $5 \times 10^{21} \text{ cm}^{-3}$.
- ▷ L'échelle nanométrique de certaines dimensions du canal rend non négligeable l'effet de confinement quantique, ou effet de quantification, dans les valeurs de potentiel électrique au sein du silicium. Les épaisseurs de canal simulées sont entre 3 et 10 nm. L'utilisation du modèle **BQP (Bohm Quantum Potential)** permet une correction quantique de ce potentiel dans toutes les équations mises en jeu dans la simulation. Les raisons du choix de ce modèle sont expliquées dans la thèse de Tang [1]. Le modèle des effets quantiques le plus performant est Schrödinger-Poisson (SP), mais celui-ci n'est pas applicable dans le cas de simulations de courant. L'incompatibilité de ce modèle avec un modèle de transport de charges impose l'appel au modèle BQP qui ne souffre pas de ce problème. SP est tout de même employé à la calibration du modèle BQP qui compte deux paramètres, γ et α . Il faut pour cela comparer les courbes $C_g(V_g)$ d'une capacité MOS entre le résultat du modèle SP, et des résultats du modèle BQP pour déterminer

le bon couple de valeurs γ et α .

- ▷ Le modèle de transport de charge est le modèle “**drift-diffusion**”. Celui-ci est le plus simple existant, prenant uniquement en compte les courants de conduction et de diffusion des porteurs. La vitesse de saturation, le transport balistique ou l’ionisation par impact, des porteurs de charges, sont négligées.
 - ▷ La spécification dans le fichier de simulation des modèles “incomplete” et “ioniz” comme inactifs permet de considérer une ionisation de tous les dopants. Tous libèrent leur porteur de charge, la température n’est pas prise en compte.
 - ▷ Un phénomène de rétrécissement du gap apparaît aux forts dopages, tels qu’ils le sont aux régions de source et de drain ($5 \cdot 10^{21} \text{ cm}^{-3}$). L’utilisation du modèle **BGN (Band-gap Narrowing)** devient nécessaire. Cependant, dans la simulation des transistors à canaux dopés, le modèle BGN n’est pas employé ici étant donné que nos modèles compacts pour ces transistors ne prennent pas en compte ce phénomène.
 - ▷ Des défauts du silicium sont pris en compte dans un modèle de recombinaisons des porteurs, **CONSRH (Shockley-Read-Hall)** dépendant de la concentration). Le modèle de recombinaison **Auger** est également utilisé.
 - ▷ La mobilité dans le silicium peut être définie comme constante. Dans ce cas nous imposons la mobilité pour les électrons $\mu_n = 1000 \text{ cm}^2/V.s$, et pour les trous $\mu_p = 500 \text{ cm}^2/V.s$. Cela facilite l’interprétation du comportement électrique du transistor du à d’autres phénomènes physiques. Nous pouvons aussi choisir d’utiliser un modèle de dégradation de mobilité. Dans ce cas, nous avons opté pour le modèle **CVT** de Lombardi qui est adapté à la simulation des transistors mutigrilles. Il tient compte de l’effet de dispersion des porteurs de charges libres du aux phonons optiques et acoustiques, et à la rugosité de surface.
- ▷ Ligne de définition des modèles sous Silvaco/Deckbuild :
- ```
models ^incomplete ^ioniz fermi bgn consrh auger bqp.n bqp.ngamma=4.5
bqp.nalpha=0.016 bqp.qdir=1 print
```
- Une fois les modèles définis, il convient de mettre en place l’algorithme mathématique de résolution du système des équations physiques. Elles sont un ensemble d’équations aux dérivées partielles et d’équations non-linéaires qui, compte tenu de leur discrétisation et du maillage, donne lieu à un problème algébrique non-linéaire. Il existe plusieurs méthodes pour résoudre ce type de problème. La documentation du simulateur Atlas fournit un mode d’emploi des méthodes dont il dispose. On y apprend que la **méthode itérative de Newton** est la mieux adaptée à notre cas de simulation 3D possédant un nombre important de points de calcul et un modèle de transport de charge *drift-diffusion*. La méthode effectivement utilisée est une variante appelée Newton-Richardson qui apporte une accélération de la convergence, lorsque celle-ci est possible. La méthode est spécifiée par le terme *AUTONR* dans le script de simulation. Chaque itération de cette

méthode demande la résolution d'un sous-problème linéaire du problème initial. En trois dimensions, le simulateur Atlas utilise un résolveur linéaire itératif. La méthode choisie est nommée *BICGST*, comme bi-conjugate gradient stabilized.

Ligne de définition des méthodes de résolution sous Silvaco/Deckbuild :

```
method maxtrap=10 autonr nblockit=45 bicgst dvlimit=1.0
```

### 5.1.3 Constitution du maillage

La constitution d'un maillage n'est pas une science exacte. Elle demande principalement de l'expérience, mais quelques règles de base suffisent pour commencer à mailler :

- ▷ Le maillage doit contenir suffisamment de points, autrement dit de mailles, pour atteindre une justesse acceptable des résultats de la simulation.
- ▷ La forme des mailles a également une grande incidence sur la qualité des résultats. Le motif de maillage le plus couramment utilisée est le triangle. On le retrouve ainsi dans les logiciels Silvaco. Sa forme triangulaire y est contrainte automatiquement selon certaines règles.
  - Les angles des triangles ne doivent pas être obtus.
  - La finesse des triangles ne doit pas dépasser une certaine limite. On impose ainsi un ratio maximal entre la plus grande dimension du triangle et la plus petite en considérant les trois dimensions du volume du transistor simulé.

Des règles supplémentaires, propres à l'objet de la simulation, peuvent être ajoutées. Dans notre cas les calculs portent sur la concentration des porteurs de charge. En fonction de leur vitesse de variation dans le volume, le maillage n'a pas besoin d'avoir la même densité. Ces différences apparaissent selon les matériaux utilisés et leur structuration dans le composant.

- ▷ Influence de la nature des matériaux :

Les grilles du transistor sont en métal. Tous ses potentiels électriques sont ainsi considérés comme égaux et un maillage large est suffisant. De même dans les régions constituées d'isolants où les potentiels varient linéairement avec la position dans le matériau, le maillage est peu dense. Cependant dans le cas des transistors MOS SOI multigrilles, il est nécessaire de densifier davantage l'oxyde enterré pour correctement simuler les lignes de champ électrique qui implique un canal d'inversion à l'interface avec le film de silicium. Cela est d'autant plus nécessaire que la grille du transistor est enterrée dans l'oxyde (cas des MOSFET II-Gate et  $\Omega$ -Gate).

- ▷ Influence des hétérojonctions :

Il est nécessaire d'affiner le maillage aux interfaces de deux matériaux différents. Une continuité de concentration de porteurs doit être assurée aux interfaces Si/SiO<sub>2</sub>.

- ▷ Influence des zones de fort champ électrique :

On trouve notamment cette situation aux jonctions source et drain, et au niveau des interfaces Si/SiO<sub>2</sub>, dans lesquelles la concentration de porteurs évolue fortement. Le maillage doit donc être densifié dans ces zones.

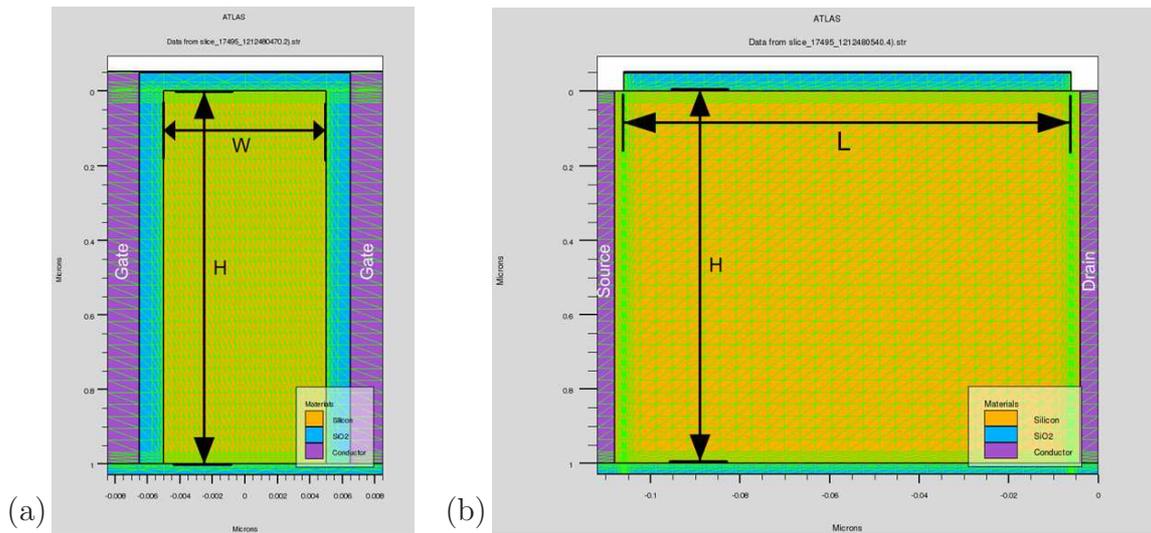


Figure 5.2 – Coupe (a) transversale et (b) longitudinale du transistor FinFET maillé.

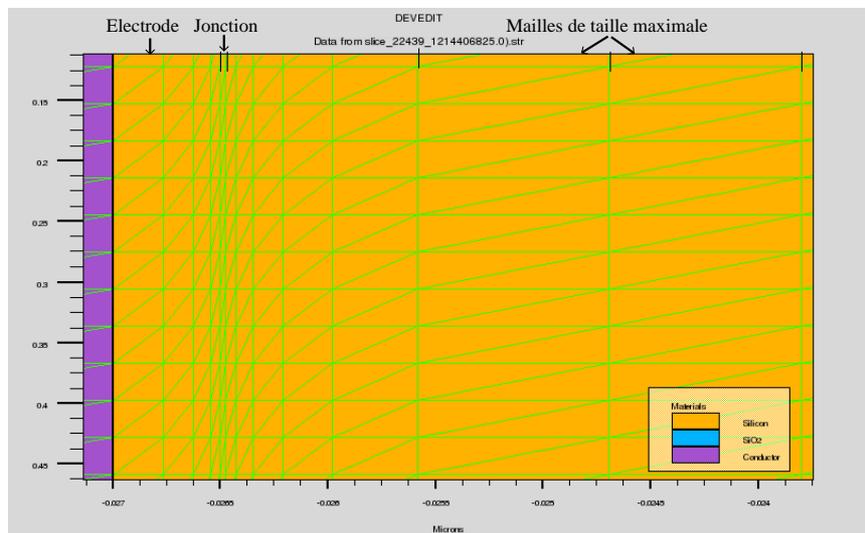


Figure 5.3 – Détail du maillage longitudinal à une jonction électrode - canal

Dans le cas des simulations 3D, deux plans de coupe du maillage doivent être définis pour discrétiser le volume du transistor. Une première coupe transversale laisse apparaître la section du canal et les deux grilles latérales (Figure 5.2 (a)). Une coupe longitudinale rend compte des jonctions au source et drain, et de la longueur du canal (Figure 5.2 (b)).

Le maillage dans le sens de la longueur ( $L$ ) du canal est défini différemment de celui de sa section. La taille des mailles n'est pas explicitement définie sur toute la longueur. Il est nécessaire de mailler finement aux jonctions du canal avec les électrodes. Dans la définition de la structure du FinFET, la longueur de ces jonctions apparaît comme étant le millième de la longueur de grille. Le maillage est plus fin à cet endroit, puis se relâche dans les électrodes et dans le centre du canal jusqu'à une taille de maille maximale. Ainsi deux paramètres seulement définissent ce maillage, l'un étant le taux de relaxation de l'espacement entre les plans de mailles transversaux, et l'autre l'espacement maximal entre deux de ces plans (voir Figure 5.3).

Pour tous les transistors que nous simulons, chacun des paramètres de leur maillage est défini en fonction de la géométrie du transistor. Ainsi, nous pouvons nous permettre de ne réaliser qu'une seule définition du maillage pour chaque structure. La géométrie du transistor simulé peut-être changée de manière transparente, tout en gardant toujours le nombre total de mailles de la structure proche du nombre maximale supportable par les ressources de notre ordinateur.

## 5.2 Extraction de paramètres du modèle FinFET

Au chapitre 3, nous avons présenté notre modèle compact du transistor FinFET ultime. L'intégration future du modèle dans un flot de conception (« Design flow ») nécessite trois aspects supplémentaires :

1. Le modèle compact du FinFET doit être écrit dans un langage dédié à la conception. Il est entièrement implémenté en Verilog-A, un des deux langages HDL les plus utilisés avec le VHDL-AMS [2–4]. Cette implémentation, présentée au Chapitre 6, permet la simulation de circuits à base de transistors FinFET. Le modèle du transistor MOSFET double grille est implémenté en VHDL-AMS [5] et une telle version du modèle FinFET fera l'objet de futurs travaux au sein de notre équipe de recherche.
2. Pour que le modèle soit complet, il doit être couplé à une méthodologie d'extraction de paramètres rigoureuse. Nous avons développé une telle méthodologie pour le modèle compact du transistor FinFET ultime. L'extraction de paramètres est automatisée. Cela permettra notamment des analyses statistiques sur les variations des paramètres (informations indispensables pour les concepteurs, et à insérer dans un flot de conception « design flow »).
3. Le modèle doit être testé au travers de simulations de circuits spécifiques permettant d'évaluer son niveau de performance (convergence – rapidité). Le chapitre 6 y est consacré.

Nous détaillerons dans cette partie le deuxième aspect cité précédemment, l'extraction de paramètres du modèle compact du transistor FinFET.

### 5.2.1 Caractéristiques électriques pour l'extraction de paramètres

Certains paramètres du modèle FinFET développé dans le cadre de cette thèse sont supposés connus ou fixés par le concepteur, et d'autres devront être extraits au travers de caractéristiques électriques.

La tâche d'extraction de paramètres est typiquement réalisée ici dans le logiciel de caractérisation IC-CAP d'Agilent, avec deux parties distinctes :

- la mesure de caractéristiques électriques,
- l'exécution de méthodes d'extraction des paramètres à partir des caractéristiques électriques.

Ces deux parties définissent notre procédure d'extraction de paramètres. Une première version de la procédure d'extraction de paramètres du modèle FinFET avait été développée au début de la thèse [6].

Les paramètres du modèle sont obtenus au travers de différentes caractéristiques électriques. Typiquement, les mesures nécessaires à l'extraction de paramètres sont les caractéristiques courant-tension :

- $I_d$  vs.  $V_{gs}$ , pour différentes valeurs de tension  $V_{ds}$ , de longueurs de grille et de largeurs de canal, bien spécifiques,
- $I_d$  vs.  $V_{ds}$ , pour différentes valeurs de tension  $V_{gs}$ , de longueurs de grille et de largeurs de canal, bien spécifiques,

et les caractéristiques capacitives :

- $C_{gg}$ ,  $C_{dg}$  et  $C_{sg}$  vs.  $V_{gs}$  pour des valeurs de tensions  $V_{ds}$ , de longueurs de grille et de largeurs de canal, bien spécifiques.

La Figure 5.5 illustre la mesure des caractéristiques électriques par le logiciel IC-CAP (Agilent). Celui-ci pilote les appareils de mesure tels que spécifiés dans la procédure d'extraction, pour obtenir des données de mesures. La Figure 5.4 est une photo prise dans la salle de caractérisation du groupe « modélisation compacte » du fondeur allemand Infineon (devenu Intel). Elle présente le banc de mesure utilisé pour obtenir les caractéristiques électriques du transistor triple grille validant le modèle multigrille au Chapitre 4. Dans le cadre de notre étude du transistor FinFET, la procédure peut également s'appliquer à des résultats de simulations 3D issus d'un simulateur de dispositifs (Atlas associé à Devedit de Silvaco par exemple).

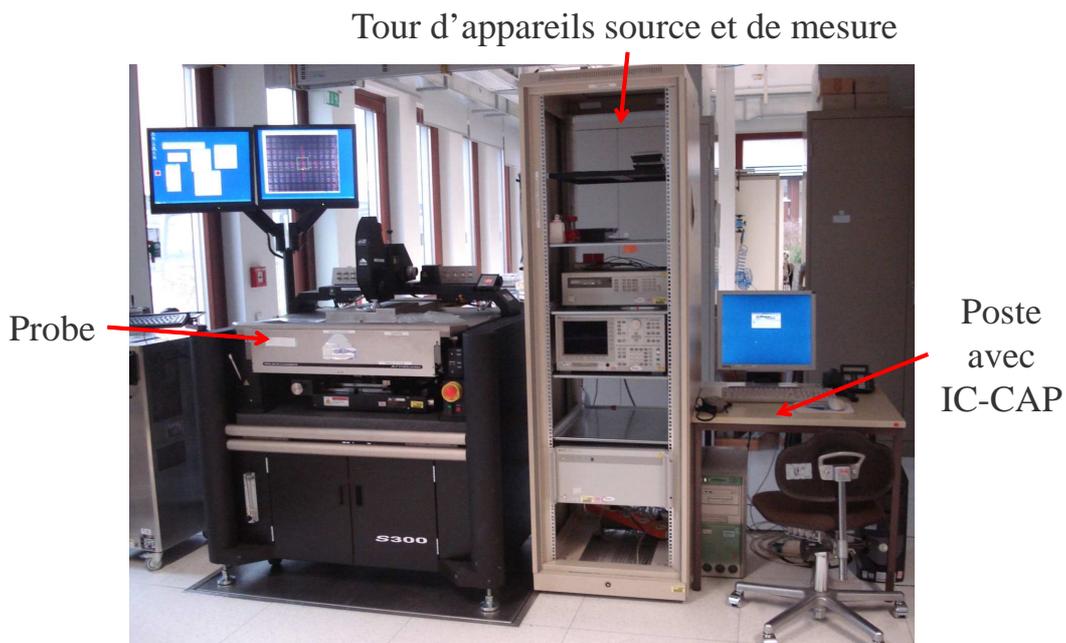


Figure 5.4 – Banc de mesures électriques du fondeur Infineon.

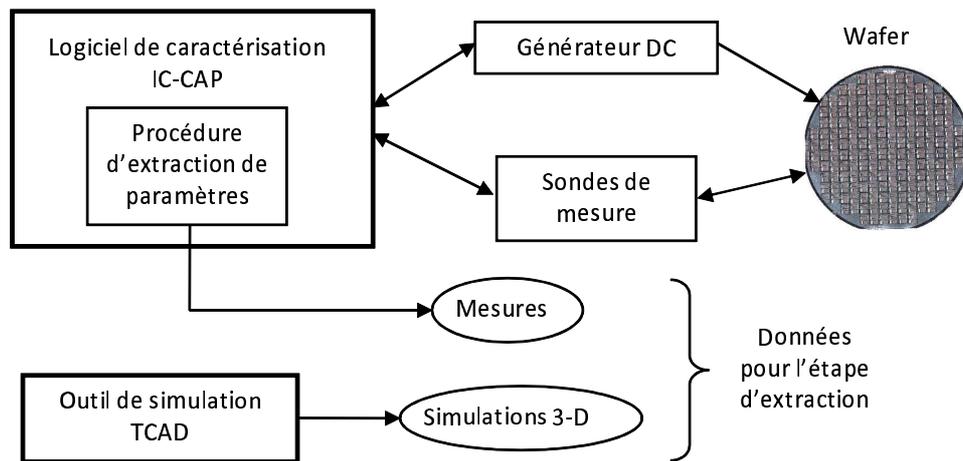


Figure 5.5 – Schématisation du banc de mesure et simulations 3D, nécessaire à l'étape d'extraction de paramètres.

## 5.2.2 Méthodologie de la procédure d'extraction de paramètres de modèle

### Méthodes d'extraction

L'extraction des paramètres peut se faire :

- ▶ soit par l'analyse directe d'une caractéristique électrique,
- ▶ soit à l'aide d'une méthode d'optimisation.

La première méthode exploite des équations du modèle pour obtenir analytiquement la valeur d'un paramètre. Mais une telle méthode n'est pas toujours possible. Ainsi il peut être nécessaire d'utiliser l'approche par optimisation. Elle consiste à comparer les résultats de caractéristiques électriques simulées selon le modèle compact, avec les caractéristiques de référence, que sont les mesures ou les simulations 3D. Chaque paramètre du modèle est optimisé pour obtenir un minimum d'écart entre des résultats du modèle et ceux de référence, définis dans une procédure d'extraction.

### Méthodologie d'extraction par optimisation

Pour extraire efficacement les paramètres, on recherche pour chacun une région de fonctionnement des caractéristiques électriques, qui est influencée par l'effet physique auquel le paramètre est associé. Ce paramètre est ensuite optimisé sur cette région. Malheureusement plusieurs paramètres sont souvent influant dans une même région, et leur optimisation doit être simultanée. D'une manière générale, ces méthodes d'optimisation demandent que des estimations initiales appropriées des paramètres soient faites. Une fois que certains paramètres sont extraits, ils sont ensuite gelés, et d'autres paramètres sont alors déterminés sur leur région de fonctionnement des caractéristiques électriques. Ces régions et l'ordre d'optimisation des paramètres définissent la procédure d'extraction.

La méthode par optimisation a des avantages et des inconvénients en ce sens que nous trou-

verons toujours une solution, mais que celle-ci pourra être très éloignée de tout sens physique. Il convient donc d'être prudent lors de la définition des procédures d'extraction de paramètres par optimisation ; c'est-à-dire qu'il est essentiel d'avoir une bonne connaissance de la valeur des paramètres devant être extraits, et de la meilleure zone (parfaitement délimitée) où ces paramètres sont les plus influents.

### Mise en oeuvre de la procédure d'extraction

La procédure d'extraction est programmée par l'utilisateur dans le logiciel de caractérisation (Figure 5.7). Elle exploite les données des caractéristiques électriques de mesures ou de simulations 3D. Pour l'application des méthodes d'optimisation, le logiciel est couplé à un simulateur de circuit qui s'occupe d'exécuter le modèle compact pour obtenir les caractéristiques simulées. Le modèle est ainsi implémenté dans un langage de description matériel (HDL), tel que Verilog-A ou VHDL-AMS.

### 5.2.3 Procédure complète automatisée d'extraction de paramètres du modèle FinFET

Le modèle compact du transistor FinFET a seulement 4 paramètres qui doivent être extraits (Table 5.1). Deux paramètres,  $E_1$  et  $E_2$ , sont utilisés dans la modélisation de la dégradation de la mobilité. Ils sont extraits pour une largeur de silicium  $W_{Si}$  donnée et pour un canal  $L$  long, à partir de la caractéristique du courant de drain en fonction de la tension de grille en échelle linéaire, dans la région de forte inversion et haute tension de drain. L'extraction est effectuée par l'optimisation du courant de drain du modèle comparé aux mesures expérimentales ou aux simulations numériques. Ensuite nous extrayons le paramètre  $\theta$ , utilisé dans le modèle de longueur de canal. Il est extrait une fois d'un dispositif canal court, à partir de la caractéristique du courant de drain en fonction de la tension de drain dans la région de saturation, par optimisation. Le dernier paramètre, la longueur de recouvrement  $L_{OV}$ , est calculé à partir de la valeur de la capacité de grille  $C_{GG}$  à  $V_{GS} = 0V$  pour un dispositif canal long à faible tension de drain. La Figure 5.6 représente cette méthodologie sous la forme d'un synoptique. La Table 5.2 synthétise les conditions électriques pour l'extraction de chaque paramètre.

Tableau 5.1 – Description des paramètres du modèle FinFET.

| Objectif                                           | Nom      | Description                                              | Unité          | Exemple      |
|----------------------------------------------------|----------|----------------------------------------------------------|----------------|--------------|
| Paramètre du process                               | $L_{ov}$ | Longueur de recouvrement                                 | m              | $10^{-8}$    |
| Paramètres liés à la mobilité                      | $e_1$    | Coefficient de réduction de la mobilité au premier ordre | $V.m^{-1}$     | $2.10^5$     |
|                                                    | $e_2$    | Coefficient de réduction de la mobilité au second ordre  | $(V.m^{-1})^2$ | $-6.10^{11}$ |
| Paramètre lié à la modulation de longueur de canal | $\theta$ | Coefficient de la tension de saturation du drain         | -              | 2.5          |

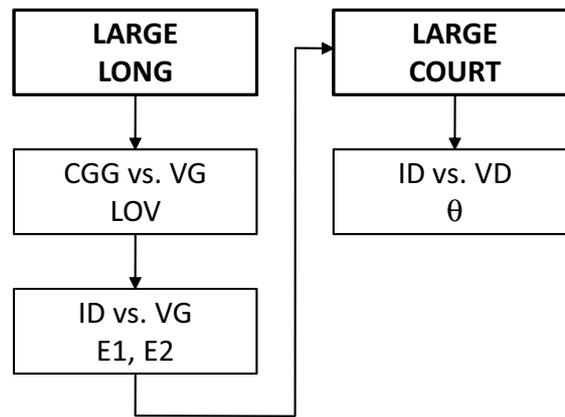


Figure 5.6 – Synoptique de la procédure d'extraction des paramètres du modèle FinFET.

Tableau 5.2 – Résumé des conditions électriques pour l'extraction de paramètres.

| Dimensions du dispositif | Caractéristiques         | Conditions                 | Paramètres |
|--------------------------|--------------------------|----------------------------|------------|
| Large / Long             | $C_{GG}$ vs. $V_G$ @ LIN | $V_G = 0V$ , co            | $L_{OV}$   |
|                          | $I_D$ vs. $V_G$ @ LIN    | SI, sat                    | $E_1, E_2$ |
| Large / Court            | $I_D$ vs. $V_D$ @ LIN    | MI, sat( $V_D \simeq 1V$ ) | $\theta$   |

SI = forte inversion, MI = inversion modérée, WI = faible inversion

sat = Saturation, co = Conduction

LOG = en échelle logarithmique, LIN = en échelle linéaire

La mise en oeuvre de la procédure d'extraction de paramètres nécessite des logiciels annexes à IC-CAP. D'une manière générale, un environnement de travail logiciel s'est mis en place au cours de la thèse, regroupant tous les logiciels et programmes informatiques développés pour faciliter le travail de recherche.

## 5.3 Environnement de travail

L'exploitation efficace des logiciels de simulation et de caractérisation a nécessité le développement de différents programmes informatiques. Ils réalisent notamment la conversion de données entre différents formats pour permettre l'échange de données entre les logiciels. Ils peuvent également traiter des caractéristiques électriques pour en extraire des grandeurs électriques comme la tension de seuil. Nous nous sommes ainsi constituéé tout un environnement de travail logiciel dans le but de développer des modèles compacts de transistors MOSFETs.

### 5.3.1 Plate-forme de simulation numérique et d'extraction des paramètres du modèle

La Figure 5.7 est une représentation schématique des logiciels utilisés lors des tâches de simulations numériques et d'extraction de paramètres ainsi que de leur connexion.

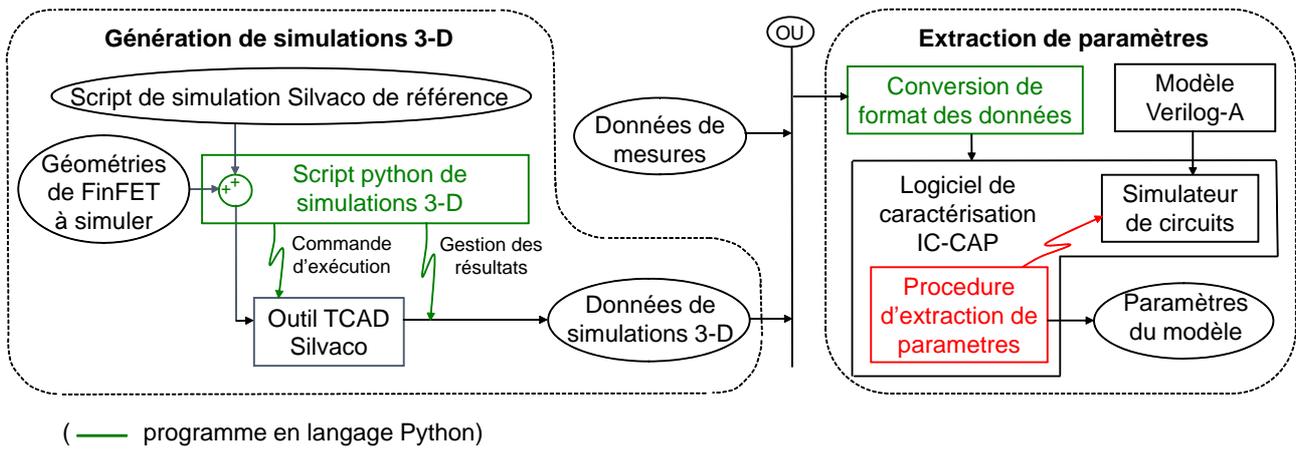


Figure 5.7 – Schéma de la plate-forme logicielle d'extraction de paramètres.

## Campagne de simulations numériques

La suite Silvaco propose des logiciels pour éditer et développer le script (ou programme) qui définit une simulation numérique, et qui l'exécute ensuite. Une fois l'exécution terminée, les fichiers contenant les caractéristiques électriques simulées sont générés au côté du fichier de script à l'origine de la simulation. Si l'utilisateur souhaite simuler un ensemble de géométries, l'exécution simulation par simulation est requise.

Pour faciliter l'obtention de caractéristiques électriques des simulations numériques, nous avons développé un programme en langage Python qui automatise l'exécution d'un ensemble de simulations. Python est un langage de script haut-niveau libre permettant la programmation de tâche simple en peu de temps [7]. Le langage est par bien des côtés semblable à Matlab, notamment dans la programmation scientifique pour laquelle de nombreuses bibliothèques sont disponibles et bien documentées. Une campagne de simulation est paramétrée dans un simple fichier texte ad-hoc au programme Python. Le Listing 5.1 en est un exemple. Une première partie définit les modèles physiques qu'il est possible de choisir parmi ceux que nous utilisons régulièrement. Puis une seconde partie est la configuration des simulations à proprement parlé, avec la spécification des modèles physiques, le nom du répertoire dans lequel sera enregistré les fichiers résultats, le type de transistor entre N et P, et le nombre de simulations que l'utilisateur souhaite exécuter en parallèle si l'ordinateur dispose de plusieurs processeurs. Enfin, les campagnes de simulations sont définies par la liste des paramètres que l'utilisateur souhaite faire varier dans le fichier Silvaco définissant la structure du transistor, sa géométrie, son maillage (dépendant de la géométrie), son niveau de dopage et les caractéristiques électriques à simuler. Dans l'exemple, nous faisons varier la température (*Lattice\_temperature*) de tout le transistor en fixant sa hauteur (*Fin\_H*) et largeur (*Fin\_W*) de film de silicium et sa longueur de grille (*Gate\_H*). Le nom des paramètres doit être celui défini dans le fichier de script de la simulation Silvaco. Si des listes sont constituées pour plusieurs paramètres, le programme Python définit tous les jeux de paramètres possibles et exécute au fur et à mesure toutes les simulations correspondantes de manière automatique. Ainsi, nous avons pu, par exemple, réaliser des campagnes d'une cinquantaine de simulations exécutées en continu pendant plusieurs jours.

Listing 5.1 – *Fichier de configuration d'une série de simulations numériques.*

```

#model definitions
begin_def physic_models
CL: default
QM: bqp.n bqp.ngamma=4.5 bqp.nalpha=0.016 bqp.qdir=1
 output: p.quantum band.temp con.band val.band band.par
end_def

begin_def mobility_models
constantMob: mobility: mun=1000 tmun=1.5 mup=500 tmup=1.5
CVT: cvt
end_def

#simulation configuration
begin_header
rootfolder /home/chevillo/Simulations/PMS04
models CL CVT
mostype N
simultaneousSimuNumber 5
end_header

begin_set
Lattice_temperature (300 T_300K), (350 T_350K), (400 T_400K)
Fin_H (0.06 h_60nm)
Fin_W (0.02 w_20nm)
Gate_L (1 l_1um)
end_set

begin_set
Lattice_temperature (300 T_300K), (350 T_350K), (400 T_400K)
Fin_H (0.06 h_60nm)
Fin_W (0.03 w_30nm)
Gate_L (1 l_1um)
end_set

```

### Gestion des fichiers résultats de simulations numériques

Les fichiers de chaque simulation d'une campagne de simulation sont automatiquement enregistrés dans des répertoires séparés et organisés en une arborescence reprenant les paramètres variant de la campagne. Chaque niveau de sous-répertoire de l'arborescence de fichiers correspond à un paramètre. La recherche ensuite d'un résultat précis est vraiment facilitée.

Lors de l'exécution de la procédure d'extraction de paramètres, le logiciel IC-CAP permet

l'importation de caractéristiques électriques dans des fichiers au format *mdm*. Pour importer des caractéristiques simulées dans Silvaco, présentes dans des fichiers textes au format propre à Silvaco, il est nécessaire de convertir les formats de données. Ainsi, différents programmes Python ont été développés pour lire et écrire les formats de fichier d'IC-CAP et de Silvaco, et donc permettre une conversion des fichiers.

### 5.3.2 Outils pour l'étude du comportement physique des transistors.

D'autres programmes Python nous permettent d'extraire différents paramètres électriques de caractéristiques électriques, issues de simulations Silvaco ou de toutes autres sources dont nous sommes capable de lire le format de fichier. Ces programmes peuvent être pilotés par d'autres qui exécutent une extraction de grandeur électrique en fonction d'une tension (de grille ou de drain), d'un paramètre géométrique ( $L, W_{Si}, H_{Si}$ ) ou du niveau de dopage ( $N_d, N_a$ ). Nous pouvons par exemple programmer l'obtention du roll-off d'un transistor en extrayant sa tension de seuil en fonction de sa longueur de grille. Nous pouvons également extraire le DIBL, la pente sous-le-seuil, les courants  $I_{on}$  et  $I_{off}$ .

L'ensemble des données extraites est ensuite automatiquement visualisé dans un graphe qui peut être enregistré en une image pour être facilement exporté. Les données sont également enregistrées dans un fichier texte au format particulier CSV sous forme d'un tableau qui est lisible dans tout logiciel tableur (tel que Excel de Microsoft).

### 5.3.3 Méthodes d'extraction de paramètres électriques

#### Courant de drain pour un MOSFET à l'état OFF et ON

Le courant de drain à l'état OFF ( $I_{off}$ ) est le courant de drain à  $V_g = 0$  V et  $V_d = 1$  V. Le courant de drain à l'état ON ( $I_{on}$ ) est quant à lui le courant de drain à  $V_g = 1$  V et  $V_d = 1$  V.

#### Tension de seuil

La tension de seuil peut être définie de différentes manières. Nous avons choisi d'utiliser la méthode dite « à courant constant » pour nous permettre de comparer facilement la tension de seuil entre deux jeux de caractéristiques électriques (par exemple, la comparaison entre des caractéristiques issues de simulations numériques et d'autres issues du modèle compact simulé). Cette méthode consiste à déterminer la tension de grille à laquelle le courant de drain est égal à un courant de référence  $I_{ref\_Vth}$  dépendant du périmètre de grille  $P$  entourant la section du canal de silicium, soit ici  $I_{ref\_Vth} = P/L \cdot 0.5 \mu A$ . Dans le cas du DG FinFET, le périmètre est  $P = 2 \cdot H_{Si}$ .

#### DIBL

Le DIBL est la différence de tensions de seuil entre celle extraite à  $V_{DS} = 1$  V et celle extraite à  $V_{DS} = 100$  mV.

## Pente sous-le-seuil

La pente sous-le-seuil est définie par :

$$S = \frac{\Delta V_{GS}}{\Delta \log(I_D)} \quad (5.1)$$

autour d'une tension de grille en faible inversion.

## Conclusion

Nous avons présenté dans ce chapitre les différents outils logiciels, qu'ils soient commerciaux ou développés par nos soins, qui nous ont permis de réaliser les travaux de cette thèse. La simulation numérique par éléments finis est possible par le logiciel Silvaco. Au sein de ce logiciel, nous avons défini un fichier par structure de transistor (FinFET, MOSFET double grille, triple grille, quadruple grille, cylindrique, et MOSFET double grille sans jonction). Le programme Python exécutant les campagnes de simulations utilise ensuite le fichier de la structure désirée par l'utilisateur.

Des résultats de mesures effectuées par IC-CAP ou des résultats de simulations numériques peuvent être importés dans la procédure d'extraction de paramètres que nous avons définie dans IC-CAP. Son exécution nous donne le jeu de paramètres du modèle prédisant le comportement précis du transistor FinFET.

Une myriade de programmes Python a été développée au cours de la thèse, chacun exécutant une tâche simple, de l'extraction d'une grandeur électrique sur une caractéristique électrique précise à la conversion de format de donnée. En perspective, il serait très intéressant de regrouper l'ensemble des programmes Python en un seul, si possible, et facilement utilisable par toute personne.

## Références

- [1] M. TANG, "Etude et modélisation compacte du transistor FinFET", thèse de doct., Université de Strasbourg, 2009.
- [2] E. CHRISTEN et K. BAKALAR, "VHDL-AMS a hardware description language for analog and mixed-signal applications", *IEEE Trans. on Circuits and Systems, part I*, vol. 46, n° 10, pp. 1263–1272, 1999.
- [3] F. PECHEUX, C. LALLEMENT et A. VACHOUX, "VHDL-AMS and Verilog-AMS as alternative hardware description languages for efficient modeling of multi-discipline systems", *IEEE Trans. on Computer Aided Design of Integrated Circuits and Systems*, vol. 24, n° 2, pp. 204–225, 2005.
- [4] P. FREY et D. O'RIORDAN, "Verilog-AMS : mixed-signal simulation and cross domain connect modules", dans *IEEE/ACM International Workshop on Behavioral Modeling and Simulation (BMAS)*, 2000, pp. 103–108.

- [5] F. PRÉGALDINY, F. KRUMMENACHER, B. DIAGNE, F. PÊCHEUX, J.-M. SALLESE et C. LALLEMENT, “Explicit modelling of the double-gate MOSFET with VHDL-AMS”, *Int. J. Numer. Model.*, vol. 19, n° 3, pp. 239–256, 2006.
- [6] N. CHEVILLON, M. TANG, F. PRÉGALDINY, C. LALLEMENT et M. MADEC, “FinFET compact modeling and parameter extraction”, dans *16th IEEE International Conference Mixed Design of Integrated Circuits and Systems*, 2009, pp. 55–60.
- [7] *Documentation du langage python*, docs.python.org.

# Chapitre 6

## Simulation de circuits numériques

### Sommaire

---

|            |                                                                       |            |
|------------|-----------------------------------------------------------------------|------------|
| <b>6.1</b> | <b>Simulations niveau-transistor de circuits numériques . . . . .</b> | <b>120</b> |
| 6.1.1      | Généralités . . . . .                                                 | 120        |
| 6.1.2      | Modèle compact du transistor FinFET pour un simulateur de circuits    | 121        |
| 6.1.3      | Simulation de portes logiques . . . . .                               | 124        |
| <b>6.2</b> | <b>Modélisation niveau-porte . . . . .</b>                            | <b>126</b> |
| 6.2.1      | Principe . . . . .                                                    | 126        |
| 6.2.2      | Procédure de paramétrage du modèle niveau-porte . . . . .             | 128        |
| 6.2.3      | Méthode d'extraction des temps de propagation . . . . .               | 130        |
| 6.2.4      | Résultats d'extraction . . . . .                                      | 131        |
| <b>6.3</b> | <b>Simulation de l'influence de variations de process . . . . .</b>   | <b>132</b> |
| 6.3.1      | Etude de la porte NOT . . . . .                                       | 132        |
| 6.3.2      | Etude de circuits CMOS complexes . . . . .                            | 133        |
|            | <b>Conclusion . . . . .</b>                                           | <b>135</b> |
|            | <b>Références . . . . .</b>                                           | <b>136</b> |

---

La conception d'un circuit numérique est une tâche complexe, même si le circuit possède peu de transistors. Pour faciliter sa tâche, le concepteur simule son circuit pour prédire son fonctionnement. La simulation de circuits est un outil indispensable au concepteur de circuits tant numériques qu'analogiques, pour s'assurer de leur performances vis-à-vis du cahier des charges. Les éléments constitutifs des circuits sont ainsi représentés par des modèles qui permettent de prédire rapidement et économiquement le comportement électrique global d'un circuit complet. Cependant, la simulation au niveau transistor d'un circuit entier demande rapidement un long temps de calcul quel que soit le niveau de complexité des circuits développés aujourd'hui. Ainsi, les concepteurs découpent leur circuit en sous-circuits sur plusieurs niveaux d'abstraction, sur lesquels ils peuvent travailler indépendamment les uns des autres avec un temps de simulation acceptable. Chaque sous-circuit doit ainsi disposer de son modèle prédictif qui est ensuite utilisé dans la simulation du sous-circuit de niveau d'abstraction directement supérieur.

La technologie de transistor FinFET est développée pour permettre la poursuite de la réduction d'échelle des circuits CMOS jusqu'à des dimensions ultimes. Le transistor FinFET ne sera donc employé que dans la conception de circuits numériques. Ces circuits sont caractérisés par deux principaux aspects : leur fonction logique et leurs paramètres temporels représentant des temps de propagations des signaux numériques entre les entrées et les sorties du circuit. Seuls ces deux aspects sont pris en compte dans les modèles des sous-circuits CMOS. Les plus simples sous-circuits CMOS sont les portes logiques qui sont modélisées par des modèles dits niveau-porte. En technologie CMOS standard, il existe de nombreuses bibliothèques de modèles niveau-porte pour des circuits pré-caractérisés. Pour de tels circuits, les paramètres technologiques et géométriques des transistors utilisés sont fixés. Par conséquent, les paramètres temporels peuvent être extraits de manière directe par des mesures expérimentales ou des simulations TCAD. A notre connaissance, il n'existe pas de tels modèles pour le transistor FinFET.

Dans ce chapitre, nous traiterons tout d'abord de la simulation niveau-transistor de portes logiques (chaque transistor est considéré dans le circuit). Puis, nous présenterons plus en détail le modèle niveau-porte et son élaboration. Pour permettre des analyses fines des performances des circuits conçus, les modèles niveau-porte intègrent également des informations quant aux variations des grandeurs temporelles autour de leur valeur nominales. Ces variations sont principalement liées aux variations du process de fabrication des transistors. Ainsi dans la dernière section de ce chapitre, nous présenterons un travail statistique sur l'influence des variations de fabrication du FinFET sur les paramètres temporels du modèle niveau-porte de circuits à base de transistors FinFET.

## 6.1 Simulations niveau-transistor de circuits numériques

### 6.1.1 Généralités

La simulation de circuits niveau-transistor demande un modèle compact des transistors qui soit implémenté dans un langage de description matériel (HDL), pour qu'il puisse être simulé dans un simulateur de circuits. Deux langages sont utilisés dans l'industrie, le langage Verilog-A [1] et le langage VHDL-AMS [2]. Nous avons implémenté notre modèle compact du transistor FinFET en Verilog-A. L'implémentation en VHDL-AMS est prévue dans de futurs travaux.

Quel que soit le logiciel simulateur de circuits, une simulation fonctionne toujours selon le même principe. Une netlist type-SPICE de tous les noeuds du circuit est créée indiquant les connexions entre tous les dispositifs le constituant. Cela peut être fait à travers une interface graphique et une édition du schéma du circuit, ou elle peut être directement écrite dans un fichier texte. En plus de la netlist, l'utilisateur spécifie pour chaque dispositif quel modèle lui attribuer. Le modèle en question est constitué d'un ensemble d'équations, explicites ou implicites, qui lient les grandeurs électriques (courant et tension) entre les noeuds interne et/ou externe du circuit. Ainsi, l'opération de simulation d'un circuit consiste en la résolution d'un système d'équations différentielles dont les inconnues sont les courants et les potentiels électriques du circuit. Le

système d'équation est d'autant plus grand que le nombre de noeud est grand et que les modèles sont complexes. Il existe différents types d'analyse. L'analyse statique consiste à résoudre le jeu d'équation en régime permanent (toutes les dérivées sont nulles). Dans ce cas, le système est résolue pour une polarisation donnée du circuit.

L'analyse transitoire permet, quant à elle, de prédire l'évolution temporelle du circuit. Le principe est simple. Le simulateur calcule à  $t=0$  un point de fonctionnement. A partir de ce point et à chaque pas de temps (le plus souvent variable), le système d'équation est linéarisé autour du point de fonctionnement puis résolu de manière matricielle. Ainsi, la complexité de la simulation a tendance à évoluer avec le nombre de noeuds du système selon une loi exponentielle. De plus, l'augmentation du nombre de noeuds dans le circuit tend également à réduire sa stabilité et ainsi augmente la probabilité que le simulateur ne puisse converger. La simulation au niveau transistor est donc limitée à un faible nombre de transistor.

### 6.1.2 Modèle compact du transistor FinFET pour un simulateur de circuits

Dans le code Verilog-A de notre modèle du transistor FinFET, il faut distinguer deux parties. Une première qui définit le modèle statique et prédit les courants statiques à chacun des terminaux du transistor. Les équations de cette partie sont exactement celles que nous avons publiées dans [3–5], et écrites aux chapitres 2 et 3. En ce qui concerne la partie dynamique du modèle, l'implémentation physique des transcapacités que nous avons développée (Section 3.2.5) conduit à des problèmes de convergences. En effet, nous nous sommes heurtés à des problèmes de convergence systématiques en simulant le modèle du FinFET avec les équations des transcapacités que nous avons développées. La principale difficulté, pour le simulateur, était liée à la discontinuité des quantités et de leur dérivées. En effet, à chaque discontinuité, le simulateur est obligé de recalculer un point de fonctionnement, ce qui n'est pas aisé, surtout autour des points de discontinuité.

Nous proposons dans la suite une explication concernant l'implémentation de notre modèle en repartant de la définition des courants petits signaux simulés dans le modèle Verilog-A :

$$\begin{aligned}
 i_g &= C_{gg} \cdot \frac{dV_g}{dt} - C_{gd} \cdot \frac{dV_d}{dt} - C_{gs} \cdot \frac{dV_s}{dt} \\
 i_d &= -C_{dg} \cdot \frac{dV_g}{dt} + C_{dd} \cdot \frac{dV_d}{dt} - C_{ds} \cdot \frac{dV_s}{dt} \\
 i_s &= -C_{sg} \cdot \frac{dV_g}{dt} - C_{sd} \cdot \frac{dV_d}{dt} + C_{ss} \cdot \frac{dV_s}{dt}
 \end{aligned} \tag{6.1}$$

Ces courants dépendent de la dérivée des potentiels aux terminaux et des transcapacités qui sont calculées à travers un grand nombre d'expressions dépendant fortement des valeurs des potentiels de source, drain et grille. Ainsi, une complexité dans le système d'équations réside naturellement dans les expressions définissant les courants petits signaux. De plus, je rappelle que les expressions des transcapacités définies à la section 3.2.5 sont déterminées différemment

selon que la tension de source-drain  $V_{DS}$  soit proche de zéro ou non. La présence des divisions par le courant normalisé dans les expressions générales des transcapacités, demande ainsi de coder dans le modèle Verilog-A deux blocs d'équations distincts pour gérer ces deux cas de figure. Cependant, la non-continuité de la dérivée des transcapacités qu'implique cette implémentation est une source importante de difficultés pour le solveur d'équations du simulateur.

Nous avons donc décidé de simplifier l'implémentation du calcul des transcapacités de sorte à gommer ces discontinuités. L'objectif est de supprimer leur dépendance avec les potentiels des terminaux, tout en conservant le comportement sous-jacent.

Etant donné que le modèle du transistor FinFET est dédié à la simulation de circuits numériques, les polarisations des terminaux du transistor pour lesquelles les transcapacités doivent être calculés précisément sont celles correspondant aux états logiques haut et bas (0V et  $V_{dd}$ ). En observant le tracé de chaque transcapacité en fonction de la tension de drain ou de grille, nous pouvons représenter leur comportement par des formes simples. Nous distinguons deux formes de tracés, selon que la transcapacité évolue avec la tension de drain à tension de grille constante ou avec la tension de grille à tension de drain constante (Figure 6.1). Les tensions de transition de ces représentations simplifiées dépendent de la tension de seuil et de la tension de saturation du drain. Ainsi, interpolant les courbes par des fonctions continues et continument dérivables, nous pouvons obtenir de nouvelles expressions des transcapacités qui reproduisent la même tendance que les expressions « physiques », qui possèdent les mêmes valeurs aux limites  $C_\epsilon$  et  $C_{V_{dd}}$  mais surtout qui facilitent grandement la convergence des simulations. Les valeurs aux limites de ce modèle dynamique interpolé sont calculées à partir des expressions « physiques » des transcapacités pour chacune d'entre elles à  $V_d = \epsilon$  avec  $\epsilon$  proche de zéro (nous avons choisi  $\epsilon = 20\text{mV}$ ), et  $V_d = V_{dd}$ . Dans cette recherche d'une nouvelle représentation plus simple des transcapacités, nous combinons maintenant les expressions des interpolations selon  $V_d$  et  $V_g$  pour obtenir  $C_{ij}(V_d, V_g)$ . Comme cela est montré à la Figure 6.2, les transcapacités sont données par l'interpolation selon  $V_g$  dont les valeurs aux limites sont déterminées par l'interpolation selon  $V_d$ . Les expressions d'interpolation selon  $V_d$  et  $V_g$  sont données respectivement par (6.2) et (6.3). Précisément, l'interpolation selon  $V_d$  n'est effectuée qu'à  $V_g = V_{dd}$ . Les transcapacités à  $V_g = \epsilon$  sont constantes et ne dépendent que de la géométrie. Elles sont obtenues en utilisant les expressions « physiques » pour une seule polarisation à  $V_d = \epsilon$ . Le

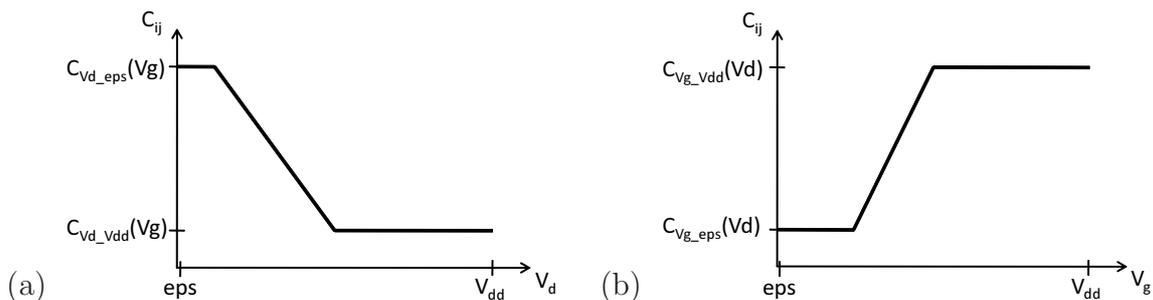


Figure 6.1 – Forme des tracés des transcapacités (a) en fonction de la tension de drain et (b) en fonction de la tension de grille.

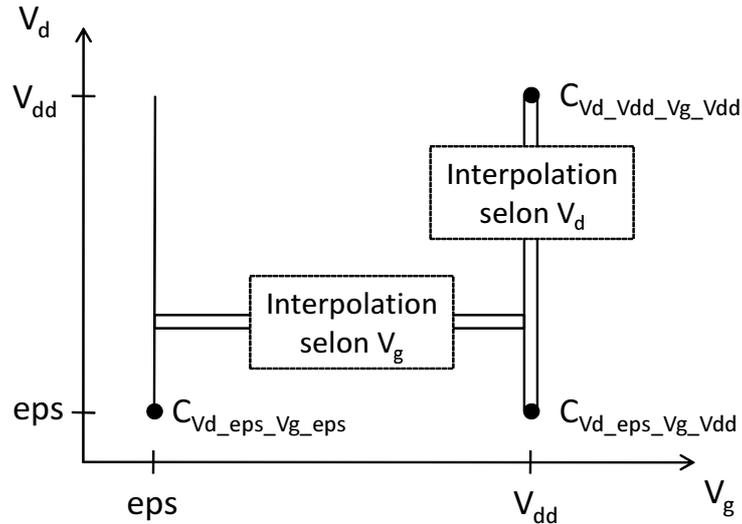


Figure 6.2 – Schéma d'élaboration des transcapacités simplifiées pour la simulation de circuits numériques.

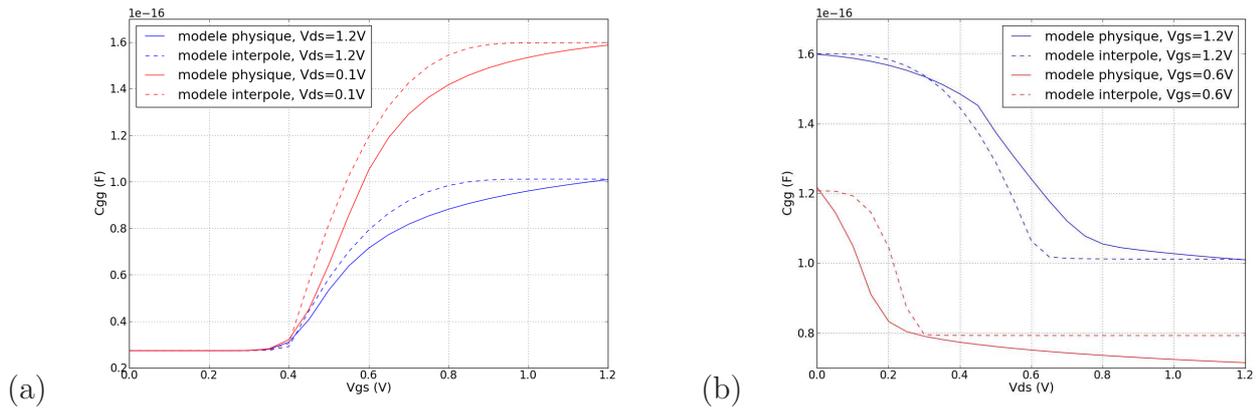


Figure 6.3 – Comparaison entre le modèle physique et le modèle interpolé de  $C_{gg}$  en régime linéaire et saturé (a) en fonction de la tension de grille, (b) en fonction de la tension de drain.  $H_{Si} = 60\text{nm}$ ,  $W_{Si} = 10\text{nm}$ ,  $L = 70\text{nm}$ .

comportement dynamique du modèle compact du transistor FinFET est donc préservé dans le cadre de simulations de circuits numériques pour laquelle il est dédié. Une comparaison entre le modèle « physique » et le modèle interpolé est donnée pour une petite géométrie à la Figure 6.3.

$$\begin{aligned}
 C_{gg\_tmpD} &= [C_{gg}(V_{dd}, \varepsilon) - C_{gg}(V_{dd}, V_{dd})] \\
 &\quad + 0.5 \cdot \left[ a_{D\_Cgg} \cdot (v_d - (\varepsilon + \Delta v_{d\_dyn})) - \sqrt{A^2 + [a_{D\_Cgg} \cdot (v_d - (\varepsilon + \Delta v_{d\_dyn}))]^2} \right] \\
 C_{gg\_vdd} &= C_{gg}(V_{dd}, V_{dd}) + 0.5 \cdot \left[ C_{gg\_tmpD} + \sqrt{A^2 + C_{gg\_tmpD}^2} \right] \quad (6.2)
 \end{aligned}$$

$$\text{avec } \begin{cases} \varepsilon = \frac{0.02 \text{ V}}{U_T} \\ a_{D\_Cgg} = (C_{gg}(V_{dd}, V_{dd}) - C_{gg}(V_{dd}, \varepsilon)) / (v_{dsat} - (\varepsilon + \Delta v_{d\_dyn})) \\ A = 0.1 \cdot (C_{gg}(V_{dd}, V_{dd}) - C_{gg}(V_{dd}, \varepsilon)) \\ v_{d\_dyn} = \frac{0.03 \text{ V}}{U_T} \end{cases}$$

$$\begin{aligned} C_{gg\_tmpG} &= [C_{gg\_vdd} - C_{gg}(V_g = \varepsilon)] \\ &+ 0.5 \cdot \left[ a_{G\_Cgg} \cdot (v_g - (v_{thinter} + \text{delta\_QME})) - \sqrt{A^2 + [a_{G\_Cgg} \cdot (v_d - (\varepsilon + \Delta v_{d\_dyn}))]^2} \right] \\ C_{gg\_vdd} &= C_{gg}(V_{dd}, V_{dd}) + 0.5 \cdot \left[ C_{gg\_tmpG} + \sqrt{A^2 + C_{gg\_tmpG}^2} \right] \end{aligned} \quad (6.3)$$

$$\text{avec } \begin{cases} \varepsilon = \frac{0.02 \text{ V}}{U_T} \\ a_{G\_Cgg} = (C_{gg}(V_{dd}, V_{dd}) - C_{gg}(V_{dd}, \varepsilon)) / (v_{dsat} - (\varepsilon + \Delta v_{d\_dyn})) \\ A = 0.1 \cdot (C_{gg}(V_{dd}, V_{dd}) - C_{gg}(V_{dd}, \varepsilon)) \\ v_{d\_dyn} = \frac{0.03 \text{ V}}{U_T} \end{cases}$$

La structure du code Verilog-A du modèle compact est disponible en annexe (Listing A.1). Maintenant que nous disposons d'un modèle dynamique de FinFET qui est compatible avec de grosses simulations transitoires, nous allons l'utiliser pour prédire le comportement de portes logiques élémentaires.

### 6.1.3 Simulation de portes logiques

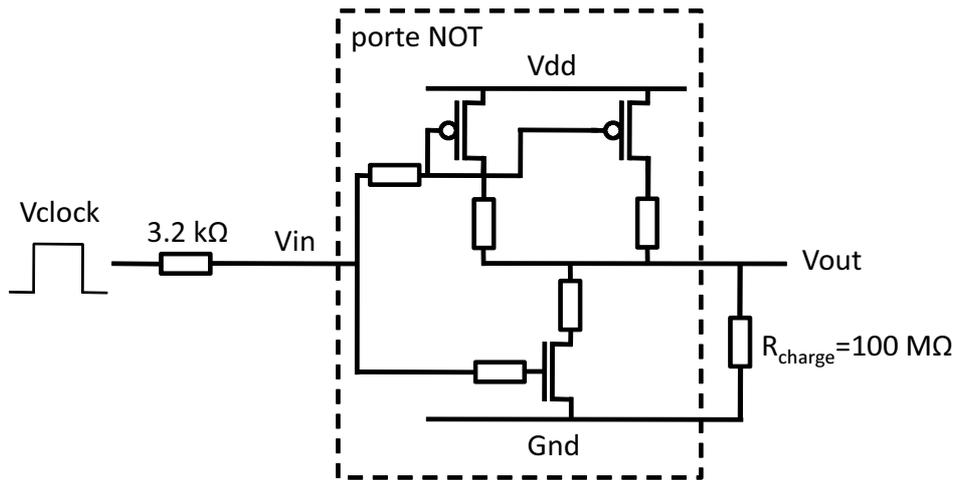


Figure 6.4 – Schéma du circuit de la porte NOT simulée.

Nous avons tout d'abord simulé la porte logique NOT. Le circuit simulé est présenté à la Figure 6.4. Nous considérons une mobilité des charges mobiles deux fois plus faible pour les P-FinFET que pour les N-FinFET. S'il s'agissait de transistors MOSFET bulk planaire, nous compenserions la différence de mobilité en doublant la largeur des P-FET. Cette dimension

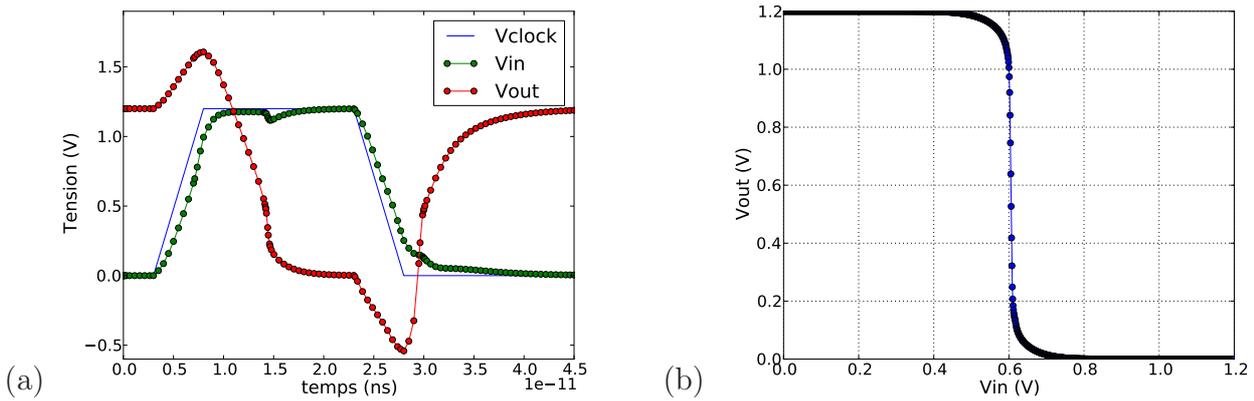


Figure 6.5 – Résultats de simulation de la porte NOT : (a) simulation transitoire, (b) caractéristique de transfert.

de largeur correspond à la hauteur  $H_{Si}$  du film de silicium du transistor FinFET. Cependant, la hauteur est fixée pour une technologie donnée, et n'est donc pas un paramètre que les concepteurs peuvent changer. L'autre paramètre dimensionnel de la section transversale du film, la largeur  $W_{Si}$ , est également fixée pour imposer la performance du transistor en permettant de rendre négligeable les effets canaux courts. Ce second paramètre technologique ne permettrait de toute façon pas la compensation de la mobilité, puisqu'il ne modifie le courant que sous-le-seuil. La seule solution pour augmenter le courant  $I_{on}$  d'un transistor FinFET est de le concevoir avec plusieurs Fin en parallèle. Le courant est donc un multiple du courant du transistor à un seul Fin. Nous simulons donc la porte NOT avec deux transistors P-FinFET en parallèle. Des résistances d'interconnexion de  $1m\Omega$ , modélisant les contacts métalliques, sont placées entre les composants. Elles ajoutent des équations supplémentaires mais en contre-partie, elles facilitent grandement la convergence sans que cela ne change les résultats de simulation. Le circuit est chargé par une résistance de très grande valeur ( $100M\Omega$  en sortie). Une résistance est connectée en série sur l'entrée pour limiter le pic de courant sur la grille lorsqu'un signal d'horloge constitué de fronts assez abrupts et idéaux est généré par le simulateur.

Un module Verilog-A à base de transistors FinFET a été constitué et pourra être appelé pour simuler au niveau-transistor un circuit utilisant des portes NOT.

Le code spectre de la simulation de la porte NOT est en annexe (A.3). La porte NOT est connectée en entrée à un signal d'horloge. L'exemple de résultats de simulation des signaux d'entrées et de sortie sont tracés à la Figure 6.5(a). Nous avons également simulé la caractéristique de transfert de la porte (Figure 6.5(b)) pour nous assurer de la symétrie de son comportement. En effet, la caractéristique passe bien par le point  $(V_{dd}/2, V_{dd}/2)$ .

Nous avons également développé le module Verilog-A de la porte NAND. Le circuit auquel il correspond est représenté à la Figure 6.6. Pour permettre la convergence de sa simulation, nous avons ajouté une équation au modèle en plaçant une capacité de valeur négligeable entre la masse, et le noeud entre les deux transistors N-FinFET. La même approche peut être utilisée pour les autres portes logiques (NOR, XOR, ...).

La simulation au niveau transistors des portes élémentaires est efficace mais conduit ra-

pidement à des jeux d'équations gigantesques, augmentant ainsi drastiquement le temps de calcul et la probabilité de non-convergence. A titre d'exemple, le modèle de la porte NAND est constituée de 221 équations, ce qui conduit à des temps de calculs de 22.62 s pour la simulation d'une seule porte par le simulateur spectre. Pour simplifier les étapes de conception et accéder facilement à des simulations de circuits plus large, nous développons un modèle niveau porte pour chaque sous-bloc logique.

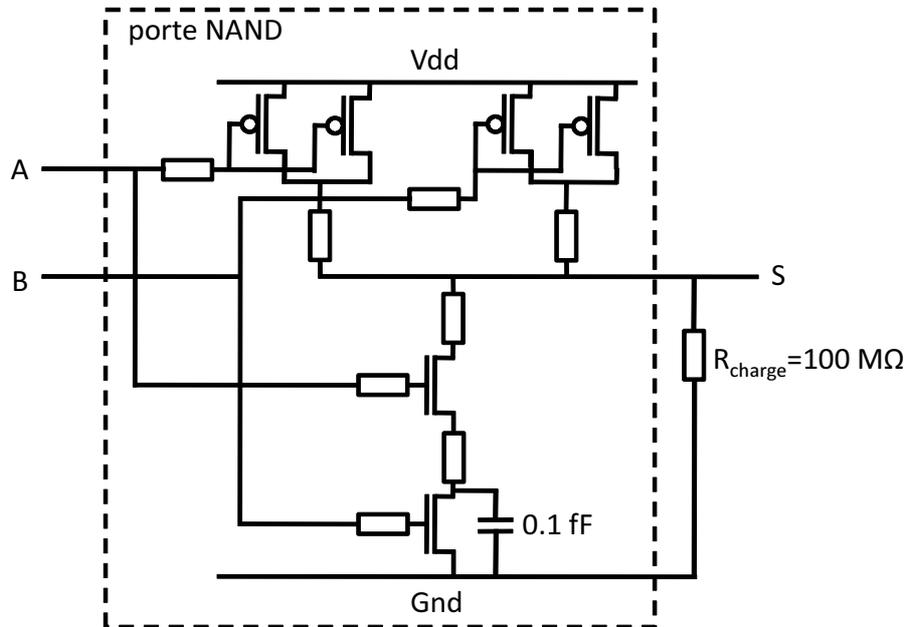


Figure 6.6 – Schéma du circuit de la porte NAND simulée.

## 6.2 Modélisation niveau-porte

### 6.2.1 Principe

La modélisation niveau-porte est un moyen de décrire des circuits numériques complexes, en utilisant pour chaque porte logique impliqué, des macro-modèles au lieu de descriptions au niveau transistor. Il s'agit d'un bon moyen pour obtenir rapidement des résultats de simulation prédictifs s'intéressant au comportement numérique du circuit autant qu'à ces performances temporelles. Par conséquent, le macro-modèle niveau-porte consiste en deux parties : la fonction logique Booléenne accomplie par la porte et des paramètres temporels (temps de propagation de l'entrée vers la sortie).

Les langages de description matériel (HDL) (Verilog, VHDL) sont de bons candidats pour cette approche, dans la mesure où ils mettent en oeuvre par nature des mécanismes pour décrire le comportement numérique en prenant en compte des aspects temporels. En VHDL, cela peut être fait en utilisant le mot-clé AFTER qui peut retarder un évènement apparaissant suite à des transitions sur les signaux numériques. Verilog incorpore également de tels mécanismes. De plus, le langage fournit des macro-modèles pour chaque fonction numérique élémentaire qui demandent trois paramètres temporels : le temps de propagation bas-vers-haut (la sortie passe

d'un état donné à l'état haut), le temps de propagation haut-vers-bas (la sortie passe d'un état donné à l'état bas), et le temps d'extinction (la sortie passe d'un état donné à l'état de haute impédance). Précisons que pour chaque paramètre, une valeur minimale et une valeur maximale peuvent être spécifiées en plus de sa valeur typique.

Au contraire des modèles compacts, qui requiert un simulateur à temps continu et conservatifs, les modèles niveau-porte sont manipulés dans des simulateurs dits « event-driven ». Ce type de simulation se déroule à pas de temps discrets dont les valeurs sont déterminés par les caractéristiques temporelles des circuits. A chaque pas de temps, la simulation se déroule en deux étapes. Tout d'abord, les signaux pour lesquels des transitions ont été prévues sont mis à jour. Ensuite, la mise à jour des signaux entraîne le déclenchement d'un certain nombre de processus (ceux qui sont sensibles aux signaux qui ont évolué). Finalement, l'exécution des processus induit d'autres transitions sur les signaux (au même pas de simulation si la transition est instantanée ou après un certain délai). Ces transitions sont écrites dans un calendrier des transitions. La simulation se termine par le calcul du pas de simulation suivant, qui correspond au premier évènement écrit au calendrier.

De par la nature du simulateur, la prédiction fine de comportements numériques et de performances temporelles de circuits numériques complexes est ainsi possible en de courts temps de simulation. Les analyses des circuits numériques à l'aide de modèles niveau portes peuvent être de trois types :

- l'analyse transitoire standard, qui utilise, pour chaque porte, les paramètres temporels typiques afin d'estimer le parcours moyen critique (pour des circuits combinatoires) et/ou la fréquence d'opération maximale et/ou le temps de propagation horloge-vers-sortie (pour des circuits séquentiels et synchrones).
- l'analyse de la conception du « pire cas » qui consiste à choisir, pour chaque paramètre, la valeur extrême qui mène aux plus mauvaises performances. Cela assure que les performances de chaque système effectivement fabriqué sont toujours meilleures que les performances prédites par le modèle à travers cette méthode.
- l'analyse statistique qui consiste en la simulation du système avec des jeux de paramètres dont les valeurs sont prises au hasard dans des intervalles définies. Cela correspond à des simulations Monte-Carlo et permet d'estimer le comportement statistique des performances temporelles d'un système donné.

Lors de la conception de circuits numériques avec une technologie standard, les fondeurs fournissent un jeu de paramètres temporel pour chaque cellule standard. Dans notre travail, nous nous focalisons sur des circuits numériques à base de transistors FinFET. Pour ces circuits, les bibliothèques standards n'existent pas encore et les concepteur fabriquent leurs circuits en « full-custom ». Ils sont donc amenés à faire varier la géométrie des FinFET. Ainsi, le modèle niveau porte doit pouvoir s'adapter à ces changements de géométrie. De fait, la géométrie deviendra un paramètre d'entrée du FinFET et les caractéristiques temporelles doivent être estimées à partir de la géométrie. Pour cela, nous utiliserons les résultats de simulation obtenus avec un modèle compact précis de FinFET pour extraire les équations liant les performances

temporelles de chaque circuit à sa géométrie.

## 6.2.2 Procédure de paramétrage du modèle niveau-porte

La structure du modèle niveau-porte d'une porte logique à base de FinFETs est donnée par le Listing 6.1. Le modèle comprend quatre parties principales :

- la déclaration des entrées et des sorties de la porte logique ;
- la déclaration des paramètres de la porte numérique (i.e., géométrie des transistors utilisés, paramètres technologiques) ;
- la déclaration de fonctions mathématiques utilisées pour estimer les performances temporelles en fonction des paramètres du modèle compact ;
- le codage de primitives Verilog correspondant au comportement numérique de la porte, et l'appel des fonctions définies précédemment pour déterminer les paramètres temporels de la primitive.

Listing 6.1 – *Structure du code Verilog du modèle niveau-porte.*

```

module FinFET_Not2 (A, Y);

// Input and output
input A;
output Y;

// Geometrical parameters
parameter real L =10e-6 from (0:inf); // FinFET Length
parameter integer N =1 from (1:inf); // Number of fins

// Technological parameters

// Delay calculations

function Trise_typ // Typical rising time
input L, N;
begin
Trise_typ=f(L,N); // To be completed
end
endfunction

function Tfall_typ // Typical falling time
// To be completed
endfunction

// Verilog primitive instantiation

not #(Trise_min(L,N): Trise_typ(L,N): Trise_max(M:N) ,
Tfall_min(L,N): Tfall_typ(L,N): Tfall_max(M:N) ,
Ttoff_min(L,N): Ttoff_typ(L,N): Ttoff_max(M:N)) u1 (A, Y);

endmodule

```

La principale difficulté dans cette approche de modélisation est de déterminer les fonctions mathématiques appropriées pour relier les paramètres géométriques aux paramètres du modèle niveau-porte. Cette analyse suit deux approches. Dans un premier temps, une série de

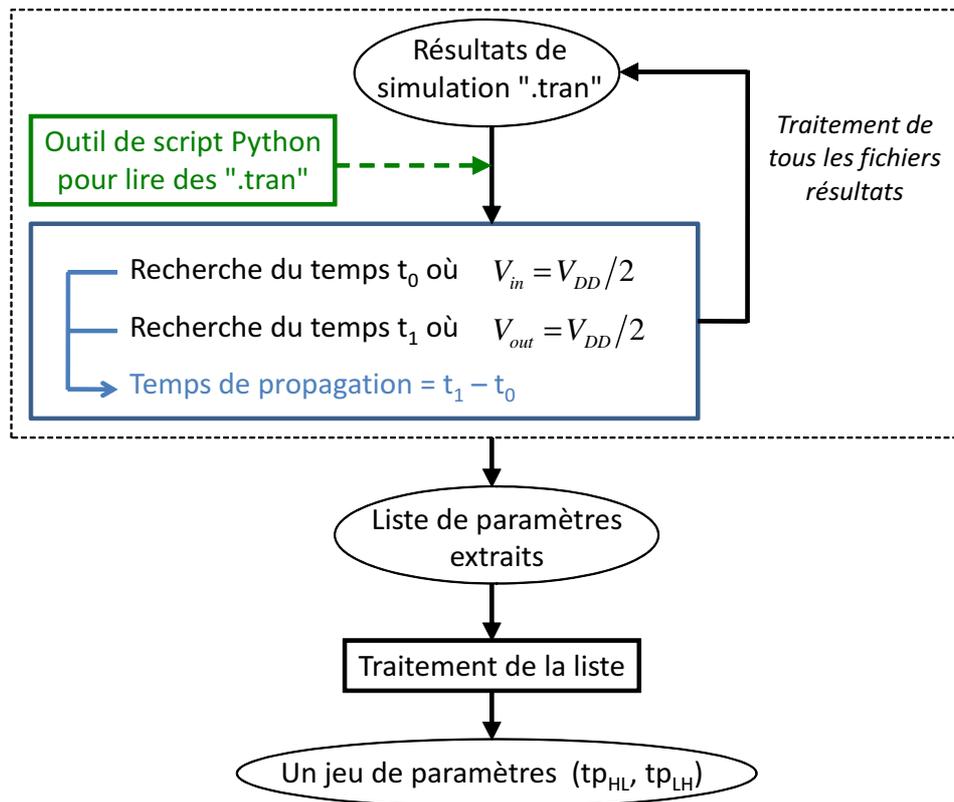


Figure 6.7 – Procédure d'extraction des paramètres temporels.

simulations paramétriques (faisant varier la géométrie) sont utilisées pour obtenir les valeurs nominales des temps de propagation. Ensuite, des simulations stochastiques (Monte Carlo) sont utilisées pour estimer la dispersion de ces caractéristiques autour de la valeur nominale (i.e., leurs valeurs minimum et maximum). Cette dispersion est principalement due aux variations locales des paramètres technologiques et aux problèmes d'appariement.

La première étape du processus d'extraction débute avec un choix de technologie, qui va permettre de fixer les paramètres technologiques du modèle compact. La structure de la porte logique au niveau transistor est alors simulée avec ces paramètres et différentes géométries. Les temps de propagation sont alors extraits, toujours pour différentes géométries à partir des simulations de ce modèle par une routine automatisée. Les résultats sont alors compilés et permettent de définir des fonctions d'interpolations qui seront ensuite intégrés dans le modèle niveau porte.

Dans un second temps, un certain nombre de simulations sont lancés avec des jeux de paramètres qui sont générés aléatoirement autour des valeurs typiques des paramètres technologiques. Pour chaque simulation, les temps de propagation sont automatiquement extraits, conduisant au tracé d'histogrammes représentant la dispersion des paramètres temporels.

Pour l'ensemble de la procédure, nous avons donc développé des scripts Python dont les fonctionnalités sont décrites Figure 6.7.

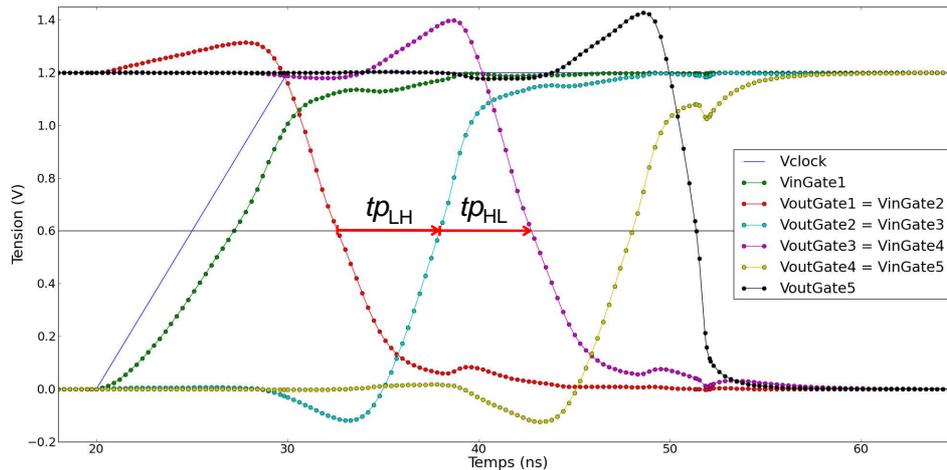


Figure 6.8 – Définition des temps de propagations  $tp_{LH}$  et  $tp_{HL}$  extraits de la simulation de 5 portes inverseuses en cascade.

### 6.2.3 Méthode d'extraction des temps de propagation

Le temps de propagation d'une porte logique est le temps nécessaire pour qu'une transition appliquée à son entrée implique la commutation du dispositif et la transition de sa sortie. Nous nous intéressons ici, comme jusqu'à la fin de ce chapitre, aux temps de propagation bas-vers-haut  $tp_{LH}$  et haut-vers-bas  $tp_{HL}$ . Nous caractérisons l'instant auquel a lieu la transition d'un signal d'un état bas vers un état haut ou d'un état haut vers un état bas, à l'instant où la tension du signal atteint la moitié de la tension d'alimentation ( $V_{dd}/2$ ). Nous pouvons donc simuler une porte câblée en inverseur et lui appliquer un front descendant pour extraire  $tp_{LH}$ , puis un front montant pour extraire  $tp_{HL}$ . Cependant, les temps de propagation, qui trouvent leur origine dans les temps de charge et de décharge des transcapacités des transistors qui composent la porte, sont influencés par la forme du signal d'entrée. Ainsi, pour obtenir un signal d'entrée réaliste (qui correspond à celui que pourrait générer une porte logique) et extraire de manière optimale les temps de propagation, nous simulons le circuit constitué de la mise en cascade de 5 fois la porte à caractériser, chaque porte étant câblée en inverseur (Figure 6.8). La première porte permet d'avoir un signal d'entrée propre et abrupt pour la seconde porte sur laquelle nous extrayons un premier temps de propagation. Puis le second temps est extrait sur la troisième porte. La quatrième porte permet de charger de manière réaliste la troisième porte (elle-même chargeant la seconde). Enfin, pour que la charge qu'est la quatrième porte soit réaliste, elle doit être elle-même chargée. Cela est réalisé par la cinquième porte qui est un simple inverseur à collecteur ouvert utilisant un seul transistor N-FET.

Les résultats d'extraction des temps de propagation de la porte NOT sont présentés au paragraphe suivant.

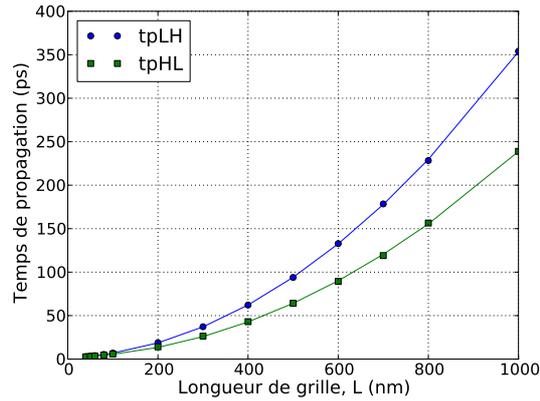


Figure 6.9 – Relation entre les temps de propagation et la longueur de grille pour  $W_{Si}=3$  nm.

## 6.2.4 Résultats d'extraction

La première étape de l'élaboration du modèle niveau-porte est de déterminer la relation donnant les temps de propagation typiques en fonction de la géométrie. Tout d'abord, nous avons simulé une porte NOT avec le modèle niveau-transistor (le modèle compact) sur une large gamme de longueur de grille, de 30 nm à  $1\mu\text{m}$ , et sur une gamme de largeur de silicium de 3 nm à 10 nm. La Figure 6.9 montre le temps de propagation haut-vers-bas  $tp_{HL}$ , et le temps de propagation bas-vers-haut  $tp_{LH}$  en fonction de la longueur de grille  $L$  pour une largeur de silicium  $W_{Si}$  de 3 nm.

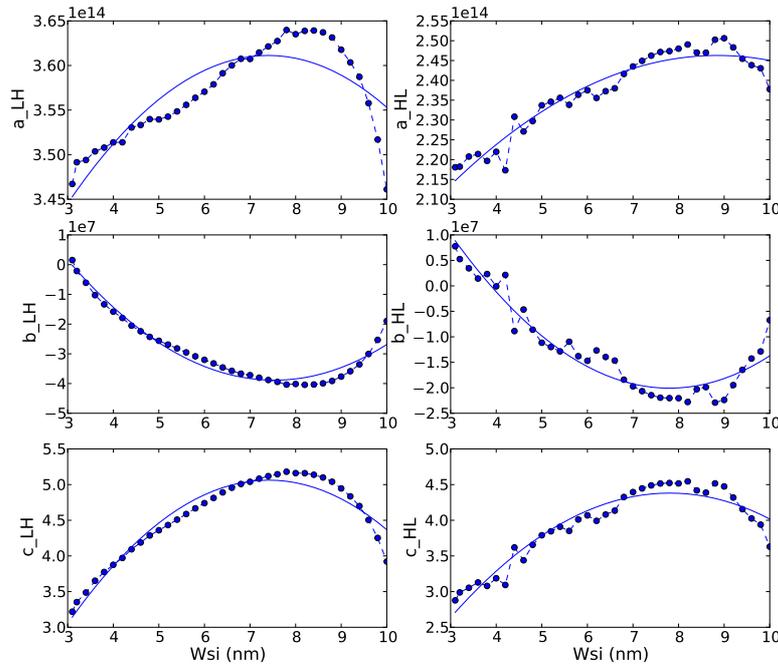


Figure 6.10 – Coefficients ( $a$ ,  $b$ ,  $c$ ) des fonctions paraboliques entre les temps de propagations  $tp_{LH}$  et  $tp_{HL}$  et la longueur de grille  $L$ , en fonction de la largeur de silicium  $W_{Si}$ . (Cercles : simulations de porte NOT, lignes : prédiction par des fonctions paraboliques).

Pour un  $W_{Si}$  donné, les courbes  $tp_{HL}=f(L)$  et  $tp_{LH}=f(L)$  peuvent être très bien prédites par des lois polynomiales du second ordre (Figure 6.9). Ensuite, les coefficients (a,b,c) du polynôme de chaque temps de propagation sont extraits pour différent  $W_{Si}$ . Pour chacune des largeurs de la gamme spécifiée, la prédiction polynomiale reste valide mais les coefficients varient légèrement d'une largeur à l'autre. Par conséquent, pour intégrer dans le modèle niveau-porte l'effet de la largeur, chaque coefficient est lui-même interpolé avec un polynôme du second ordre dépendant de  $W_{Si}$  (Figure 6.10). Ainsi, les fonctions  $tp_{HL}=f(L)$  et  $tp_{LH}=f(L)$  sont :  $tp(L, W_{Si}) = a(W_{Si}) \cdot L^2 + b(W_{Si}) \cdot L + c(W_{Si})$ .

Les relations simples entre la géométrie du transistor et ses temps typiques, permet au concepteur de rapidement trouver une première solution pour son circuit, correspondant à son cahier des charges. L'influence de la dispersion des paramètres technologiques des transistors, inhérente à toute fabrication, doit ensuite être prise en compte.

## 6.3 Simulation de l'influence de variations de process

### 6.3.1 Etude de la porte NOT

Dans la seconde étape de l'élaboration du modèle niveau-porte, une estimation de l'intervalle de variation des paramètres temporels autour de leur valeurs typiques est étudiée. Dans ce but, nous avons réalisé des simulations Monte-Carlo pour différentes géométries nominales (Figures 6.11 et 6.12).  $W_{Si}$ ,  $L$  et  $t_{ox}$  varient selon une loi gaussienne sur un intervalle de 5% autour de leur valeurs typiques. Les résultats sont donnés dans les tableaux 6.1 et 6.2. D'après ces simulations, nous pouvons remarquer que le paramètre le plus sensible est la longueur du transistor  $L$  pour lequel les intervalles de variation de  $tp_{HL}$  et  $tp_{LH}$  correspondent à environ 3.5% autour de la valeur typique.

Tableau 6.1 – Synthèse de la dispersion des temps de propagation (en ps) pour la géométrie  $W_{Si}=5\text{ nm}$ ,  $L=30\text{ nm}$ ,  $t_{ox}=1.5\text{ nm}$ .

| Paramètre<br>Variant | $tp_{LH}$ |      |      | $tp_{HL}$ |      |      |
|----------------------|-----------|------|------|-----------|------|------|
|                      | Min       | Typ  | Max  | Min       | Typ  | Max  |
| $L$                  | 2.81      | 2.85 | 2.90 | 2.69      | 2.74 | 2.79 |
| $W_{Si}$             | 2.85      | 2.85 | 2.86 | 2.73      | 2.74 | 2.75 |
| $t_{ox}$             | 2.85      | 2.85 | 2.85 | 2.74      | 2.74 | 2.74 |

Tableau 6.2 – Synthèse de la dispersion des temps de propagation (en ps) pour la géométrie  $L=100\text{ nm}$ ,  $t_{ox}=1.5\text{ nm}$ , pour une variation en  $W_{Si}$ .

| Largeur<br>nominale     | $tp_{LH}$ |      |      | $tp_{HL}$ |      |      |
|-------------------------|-----------|------|------|-----------|------|------|
|                         | Min       | Typ  | Max  | Min       | Typ  | Max  |
| $W_{Si} = 5\text{ nm}$  | 7.83      | 7.85 | 7.87 | 6.79      | 6.82 | 6.85 |
| $W_{Si} = 10\text{ nm}$ | 7.67      | 7.69 | 7.73 | 6.99      | 7.00 | 7.02 |

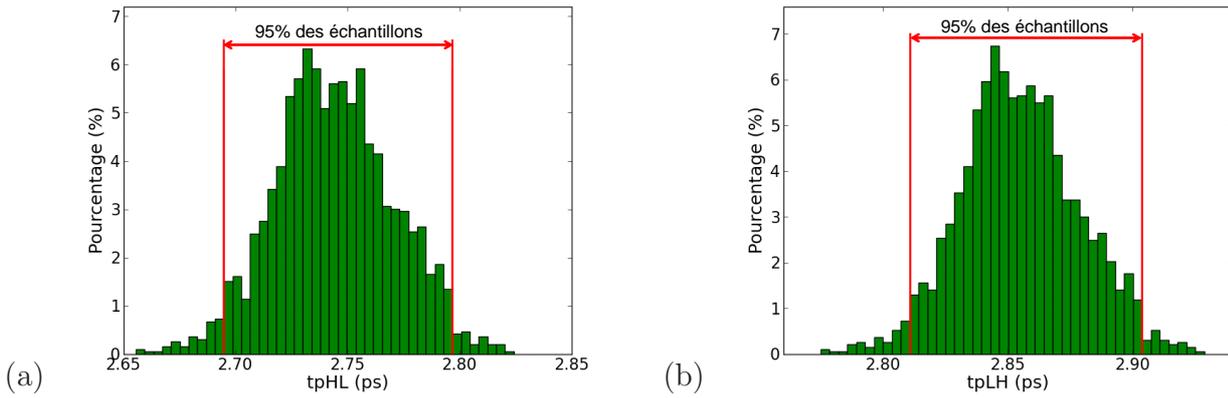


Figure 6.11 – Distribution de 2000 temps de propagation (a)  $tp_{HL}$ , (b)  $tp_{LH}$  extraits à partir d'une distribution gaussienne de la longueur de grille  $L$  pour une valeur moyenne de 30 nm. Trois fois la déviation standard ( $3 \cdot \sigma$ ) est la valeur moyenne fois 5%. La géométrie nominale est  $W_{Si} = 5$  nm,  $L = 30$  nm,  $t_{ox} = 1.5$  nm.

Une approche comparable a été réalisée sur une porte NAND et sera utilisée par la suite dans l'étude de circuits plus complexes.

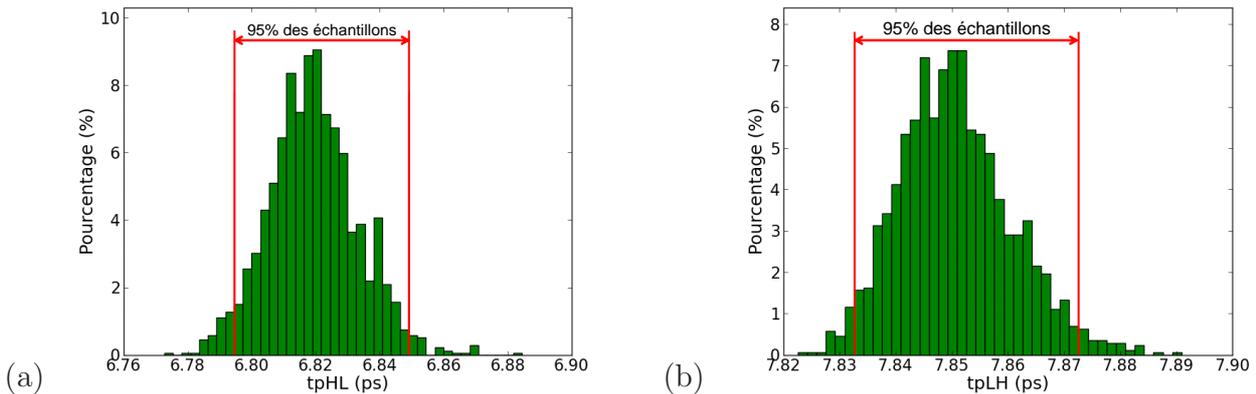


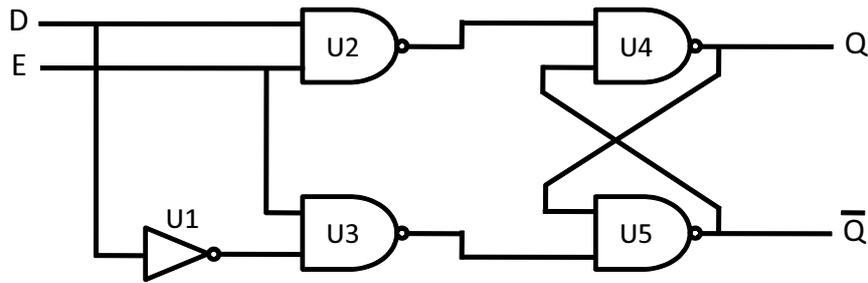
Figure 6.12 – Distribution de 2000 temps de propagation (a)  $tp_{HL}$ , (b)  $tp_{LH}$  extraits à partir d'une distribution gaussienne de la largeur de silicium  $W_{Si}$  pour une valeur moyenne de 5 nm. Trois fois la déviation standard ( $3 \cdot \sigma$ ) est la valeur moyenne fois 5%. La géométrie nominale est  $W_{Si} = 5$  nm,  $L = 100$  nm,  $t_{ox} = 1.5$  nm.

### 6.3.2 Etude de circuits CMOS complexes

L'objectif de cette troisième partie est double. D'une part, il s'agit de valider notre approche en comparant les résultats de simulation obtenus avec les modèles niveau transistor et niveau porte sur des circuits numériques plus complexes que de simple portes logiques. D'autre part, elle permet de démontrer tout l'intérêt du modèle compact.

#### 6.3.2.1 D-latch

Le premier circuit étudié est un verrou D (D-latch) constitué à partir de portes NAND et dont le schéma logique est rappelé sur la Figure 6.13. Le circuit est composé de 5 portes NAND

Figure 6.13 – *D-Latch*.Tableau 6.3 – *Comparaison des temps de propagation (en ps) entre des simulations avec le modèle niveau-porte et le modèle niveau-transistor pour le D-latch.*

| D-latch                            | Modèle niveau-porte | Modèle niveau-transistor | Désaccord |
|------------------------------------|---------------------|--------------------------|-----------|
| Temps de propagation haut-vers-bas | 109.8               | 114                      | -3.7 %    |
| Temps de propagation bas-vers-haut | 54.9                | 49                       | +12 %     |

(l'inverseur est réalisé grâce à une porte NAND dont les deux entrées sont reliées). Chaque porte NAND consiste en 6 transistors de même géométrie : 2 N-FinFET et 4 P-FinFET (6.6). La géométrie de chaque transistor est la suivante :  $W_{Si} = 10$  nm,  $H_{Si} = 50$  nm,  $L = 200$  nm. Le même circuit est modélisé avec un modèle niveau-transistor et un modèle niveau-porte. Les valeurs typiques de  $tp_{LH}$  et  $tp_{HL}$  pour une porte NAND sont respectivement 28.6 ps et 26.4 ps pour cette géométrie. Ces valeurs sont utilisées comme paramètres dans le modèle niveau-porte.

Sur le D-latch, le signal d'inhibition (E) est fixé à '1' et la donnée (D) commute entre des états haut et bas. La comparaison entre les temps de propagation, pour le circuit global, est donnée au tableau 6.3. Le fait que le temps de propagation bas-vers-haut est presque deux fois le temps de propagation haut-vers-bas, peut être expliqué théoriquement. Dans le cas où D passe de l'état bas à l'état haut, la sortie de U2 passe de l'état haut à l'état bas. Comme  $\bar{Q}$  est à l'état haut, U4 commute directement et la sortie passe de l'état bas à l'état haut. Le temps de propagation global est la somme de  $tp_{LH}$  avec  $tp_{HL}$ . Dans un second temps, lorsque D passe de l'état haut à l'état bas, la sortie de U2 passe de l'état bas à l'état haut. Comme  $\bar{Q}$  est à l'état bas, la sortie reste à l'état haut.  $\bar{Q}$  commute seulement après U1, U3 et U5. Par conséquent, pour la transition haut-vers-bas, le chemin critique passe par 4 portes (U1, U3, U5 et U4) au lieu de 2 pour la transition bas-vers-haut.

Les résultats obtenus avec les deux modèles sont très proches. Le léger désaccord peut être expliqué par le fan-out (nombre de portes logiques connectés sur une sortie) qui n'est pas égal à 1 pour chaque porte. Un large fan-out tend à augmenter légèrement le temps de propagation, phénomène qui n'est pas pris en compte dans cette première version du modèle niveau-porte.

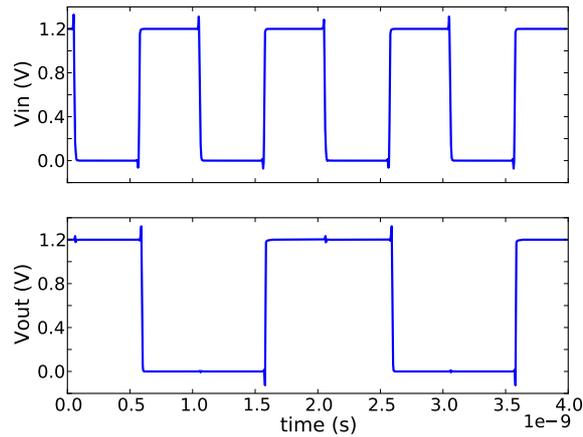


Figure 6.14 – Simulations niveau-transistor du diviseur de fréquence.

Tableau 6.4 – Comparaison du temps Horloge-vers-sortie (en ps) entre des simulations avec le modèle niveau-porte et le modèle niveau-transistor pour le diviseur de fréquence.

| Diviseur de fréquence                                   | Modèle niveau-porte | Modèle niveau-transistor | Désaccord |
|---------------------------------------------------------|---------------------|--------------------------|-----------|
| Temps Horloge-vers-sortie pour une sortie bas-vers-haut | 15.4                | 14.5                     | +6.2 %    |
| Temps Horloge-vers-sortie pour une sortie haut-vers-bas | 21.3                | 20.0                     | +6.5 %    |

### 6.3.2.2 Diviseur de fréquence par 2

Pour aller plus loin, nous simulons maintenant un diviseur de fréquence composé de deux D-latch câblé en maître-esclave. Les résultats de simulation de cette structure sont donnés à la Figure 6.14.

Comme pour le D-latch, les deux approches de modélisation sont comparées. La géométrie du transistor utilisé dans ces circuits est  $W_{Si} = 10$  nm,  $H_{Si} = 50$  nm,  $L = 50$  nm. Pour cette géométrie, les valeurs typiques de  $tp_{LH}$  et  $tp_{HL}$  pour une porte NAND seule sont respectivement 4,8 ps and 5,9 ps. Nous nous focalisons sur l'estimation du « clock-to-output time » (temps entre le front d'horloge et la commutation de la sortie). Les résultats sont donnés au tableau 6.4. Ils sont, eux aussi, en bon accord.

## Conclusion

Nous avons développé dans ce chapitre une nouvelle approche pour développer des modèles niveau-porte pour des circuits numériques à base de transistors FinFET. Comme il n'y a pas de cellules standards dans cette technologie, le concepteur est amené à jouer directement sur les paramètres géométriques des transistors constituant la porte. Ainsi, nous avons défini un modèle niveau-porte dans lequel les paramètres temporels dépendent de la géométrie. Pour ce faire, nous basons notre travail sur le modèle compact du transistor FinFET qui est précis et proche de la physique du dispositif. Une procédure automatisée a été établie afin d'extraire,

pour une porte donnée dans une technologie donnée, les équations mathématiques qui prédisent les macro-modèles niveau-transistor dont les paramètres dépendent de la géométrie. Pour des circuits standards, les résultats de simulation obtenus avec le modèle niveau-porte et le modèle niveau-transistor (utilisant le modèle compact FinFET) sont quantitativement comparables, mais différents en termes de temps de calcul. Pour cette considération, le modèle niveau-porte est clairement plus intéressant et est la seule manière de prédire les performances temporelles d'un circuit complexe dans un temps de calcul raisonnable. Le modèle compact sur lequel le modèle niveau-porte est basé, intègre la plupart des phénomènes physiques qui altèrent les caractéristiques du transistor. De plus, l'approche développée dans ce chapitre, et publiée dans [6], pourra être appliquée aux différents dispositifs multigrilles, notre modélisation des transistors multigrilles étant basée sur celle du transistor FinFET. Pour le modèle niveau-porte lui-même, des améliorations devront être apportées dans une nouvelle version. Une des plus importantes est la prise en compte de l'effet du fan-out qui semble légèrement altérer les temps de propagation d'une porte. De plus, des informations sur la consommation d'énergie ou la surface de silicium nécessaire au circuit peuvent être ajoutées, comme cela est le cas dans les kits de conception. Ces améliorations demandent la manipulation de modèle durant la simulation de process et peuvent être mise en oeuvre seulement par la couche logiciel au-dessus du modèle (par exemple en utilisant des scripts Skill des outils logiciels de Cadence).

## Références

- [1] S. A. BROWN et Z. G. VRANESIC, *Fundamentals of Digital Logic With Verilog Design*, MCGRAW HILL HIGHER EDUCATION, éd. 2007.
- [2] P. ASHENDEN, *The Designer's Guide to VHDL*, MORGAN KAUFMANN PUBLISHERS IN, éd. 2008.
- [3] J.-M. SALLESE, F. KRUMMENACHER, F. PRÉGALDINY, C. LALLEMENT, A. ROY et C. ENZ, "A design oriented charge-based current model for symmetric DG MOSFET and its correlation with the EKV formalism", *Solid-State Electron.*, vol. 49, n° 3, pp. 485–489, 2005.
- [4] F. PRÉGALDINY, F. KRUMMENACHER, B. DIAGNE, F. PÊCHEUX, J.-M. SALLESE et C. LALLEMENT, "Explicit modelling of the double-gate MOSFET with VHDL-AMS", *Int. J. Numer. Model.*, vol. 19, n° 3, pp. 239–256, 2006.
- [5] A. YESAYAN, F. PRÉGALDINY, N. CHEVILLON, C. LALLEMENT et J.-M. SALLESE, "Physics-based compact model for ultra-scaled FinFETs", *Solid-State Electron.*, vol. 62, n° 1, pp. 165–173, 2011.
- [6] N. CHEVILLON, M. MADEC et C. LALLEMENT, "Gate-level modeling for CMOS circuit simulation with ultimate FinFETs", dans *IEEE International Symposium on Nanoscale Architectures (NANOARCH'12)*, 2012, pp. 22–29.

# Chapitre 7

## Modélisation du transistor MOSFET double grille sans jonction

### Sommaire

---

|            |                                                                                                                |            |
|------------|----------------------------------------------------------------------------------------------------------------|------------|
| <b>7.1</b> | <b>Etat de l'art . . . . .</b>                                                                                 | <b>138</b> |
| <b>7.2</b> | <b>Étude électrostatique dans les DG MOSFET sans jonction . . . . .</b>                                        | <b>139</b> |
| 7.2.1      | Limitations du modèle de DG MOSFET à base de jonctions pour<br>simuler les dispositifs sans jonction . . . . . | 139        |
| 7.2.2      | Solution approchée de l'équation de Poisson . . . . .                                                          | 141        |
| 7.2.3      | Approche générale du calcul de la charge mobile . . . . .                                                      | 143        |
| 7.2.4      | Expressions approchées pour les charges . . . . .                                                              | 144        |
| 7.2.5      | Tension de seuil dans les transistors sans jonction . . . . .                                                  | 149        |
| <b>7.3</b> | <b>Développement de l'expression de courant . . . . .</b>                                                      | <b>151</b> |
| 7.3.1      | Courant en accumulation . . . . .                                                                              | 152        |
| 7.3.2      | Courant en déplétion . . . . .                                                                                 | 152        |
| 7.3.3      | Traitement général du courant dans le DG MOSFET sans jonction . . . . .                                        | 152        |
| 7.3.4      | Comparaison avec des simulations numériques . . . . .                                                          | 154        |
|            | <b>Conclusion . . . . .</b>                                                                                    | <b>157</b> |
|            | <b>Références . . . . .</b>                                                                                    | <b>158</b> |

---

Nous avons développé un modèle analytique pour le transistor MOS à effet de champ double grille (DG MOSFET) sans jonction. Le principe de ce dispositif a été démontré récemment [1]. Malgré des similarités avec le DG MOSFET classique à jonctions, les relations charges-potentiels sont assez différentes. Un modèle du MOSFET sans jonction ne peut pas simplement être basé sur un formalisme existant de transistors multigrilles. Cela est particulièrement vrai si l'on considère certains paramètres technologiques importants que sont le dopage du film de silicium et sa largeur  $W_{Si}$ , pour des niveaux de dopage excédant  $10^{19} \text{ cm}^{-3}$  et des largeurs autour de 10 à 20 nm. Des simulations numériques 2D TCAD ont permis d'évaluer le modèle et de confirmer sa validité pour toutes les régions d'opération, c'est-à-dire de la déplétion profonde à l'accumulation, et du régime linéaire à la saturation.

## 7.1 Etat de l'art

Un nouveau concept de transistor à effet de champ (FET) sans jonction disposant d'une architecture triple grille a été récemment simulé [1] et validé expérimentalement [2]. Selon le même concept, d'autres topologies ont été proposées [3, 4], connues sous le nom de dispositifs « vertical-slit FET ». Dans une tentative de développer un corps de modèle générique pour cette famille de dispositifs émergeant, nous avons focalisé notre recherche sur une structure double grille symétrique (Figure 7.1). Une représentation de sa distribution de potentiel se trouve à la Figure 7.2. La structure consiste en une couche de silicium uniformément dopée de type N, de quelques dizaines de nanomètre de largeur et situé entre deux grilles de métal. A la différence d'un DG MOSFET classique, les régions de source et de drain ont un dopage du même type que le silicium situé entre ces deux régions. Ainsi, le courant est uniquement issu des porteurs majoritaires, contrairement au canal d'inversion du DG MOSFET. La technologie sans jonction présente plusieurs avantages [1, 3] tels que la disparition de jonctions abruptes dont la fabrication peut être difficilement contrôlable à l'échelle nanométrique. Le processus de fabrication est simplifié, la variabilité du niveau de dopage est plus faible et l'existence d'une conduction volumique implique une minimisation de la dégradation de la mobilité due à la rugosité de surface et une minimisation du bruit rose. Il a été montré que les dispositifs DG à fine couche utilisés en accumulation, ont une meilleure immunité contre les effets canaux courts et l'effet DIBL, que les transistors multigrilles dont les canaux fonctionnent en inversion [4]. De grands ratios  $I_{on}/I_{off}$  et une pente sous-le-seuil idéale ont été simulés [1] et mesurés. Bien que la modélisation ait été faite pour le MOSFET SOI fonctionnant en accumulation [5, 6], il n'y a toujours pas de modèle analytique pour les dispositifs sans jonction DG à couche ultra-mince, qui fonctionnent en régimes d'accumulation et de déplétion.

Dans ce chapitre, nous proposons tout d'abord d'estimer si le modèle basé sur la charge des dispositifs DG MOSFET avec jonction à canal dopé et non dopé, utilisant le concept de largeur équivalente [7, 8], peut être exploité pour obtenir un modèle des DG sans jonction. En effet, la seule différence technologique entre ces deux types de transistor MOS est le signe du dopage du film de silicium. Nous pouvons simplement penser qu'il suffit d'inverser ce signe dans le modèle du DG dopé avec jonction pour obtenir un modèle du DG sans jonction. Nous estimerons donc la limite de validité du modèle des DG classiques lorsqu'il est appliqué aux DG MOSFET sans

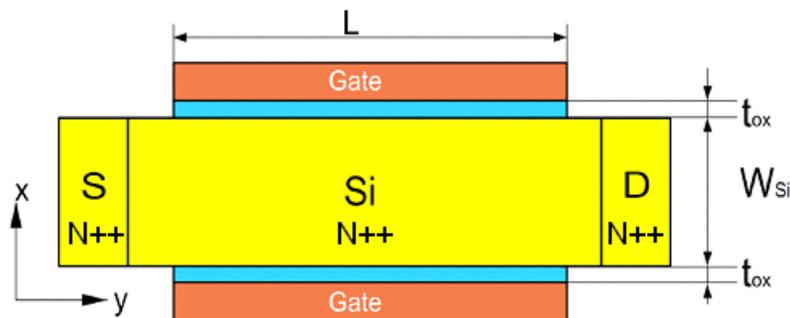


Figure 7.1 – Coupe schématique du transistor DG N-MOS sans jonction.

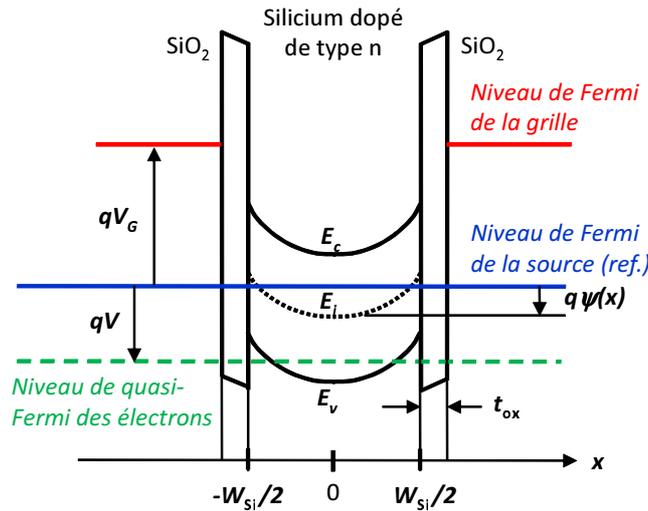


Figure 7.2 – Diagramme de bandes d'énergie du transistor DG N-MOS sans jonction.

jonction. Il s'avèrera que cette méthode atteint rapidement ses limites. Nous proposons donc une approche totalement différente afin de déterminer la relation entre charges et potentiels, ainsi que l'expression du courant de drain. Ce nouveau modèle est l'objet de ce chapitre.

## 7.2 Étude électrostatique dans les DG MOSFET sans jonction

Nous considérons un DG MOSFET sans jonction ayant une largeur de silicium  $W_{Si}$  et une concentration de donneur  $N_D$ . Une coupe transversale de sa structure en termes de potentiels est montrée à la Figure 7.2.

### 7.2.1 Limitations du modèle de DG MOSFET à base de jonctions pour simuler les dispositifs sans jonction

Bien que les architectures à base de jonctions et celles sans jonction soient très similaires, le principe de fonctionnement des dispositifs sans jonction est assez différent de celui des DG MOSFET standards. Dans un DG MOSFET sans jonction, l'effet de la grille est de dépléter ou d'accumuler des porteurs majoritaires de la couche de silicium dopée, et non de créer une couche d'inversion à l'interface du semi-conducteur et de l'isolant. Ainsi, même à fortes tensions de grille, le flot de porteurs majoritaires se fait essentiellement à travers le volume du semi-conducteur.

Rappelons les relations de base du DG MOSFET à canal dopé, qui sont indépendantes de la nature du dopage des régions source et drain : comme pour le DG MOSFET à base de jonctions [8], lier l'équation de Poisson aux statistiques de Boltzmann non dégénérées mène à une relation entre le champ électrique de surface et le potentiel de surface. La source est utilisée comme référence pour les potentiels (seul le signe est changé par rapport à [8] lorsque

nous considérons des atomes ionisés de type donneur). En négligeant les trous, nous avons donc :

$$\frac{d^2\Psi}{dx^2} = \frac{e}{\varepsilon_{Si}} \cdot \left( n_i \cdot e^{\frac{\Psi - V_{ch}}{U_T}} - N_D \right) \quad (7.1)$$

Cette équation différentielle n'a pas de solution analytique. Nous pouvons néanmoins l'intégrer pour obtenir une expression analytique du champ électrique, comme dans [8]. En exprimant la dérivée seconde en fonction de la dérivée première à partir de l'expression  $\frac{d}{dx} \left[ \left( \frac{d\Psi}{dx} \right)^2 \right] = 2 \cdot \frac{d\Psi}{dx} \cdot \frac{d^2\Psi}{dx^2}$ , nous pouvons transformer (7.1) :

$$\frac{d}{dx} \left[ \left( \frac{d\Psi}{dx} \right)^2 \right] = 2 \cdot \frac{d\Psi}{dx} \cdot \frac{e}{\varepsilon_{Si}} \cdot \left( n_i \cdot e^{\frac{\Psi - V_{ch}}{U_T}} - N_D \right) \quad (7.2)$$

En intégrant (7.2) du centre à l'interface canal/oxyde, nous pouvons finalement exprimer le champ électrique de surface :

$$E_S^2 = \int_{\Psi_0}^{\Psi_s} d \left( \frac{d\Psi}{dx} \right)^2 = \int_{\Psi_0}^{\Psi_s} \frac{2 \cdot e}{\varepsilon_{Si}} \cdot \left( n_i \cdot e^{\frac{\Psi - V_{ch}}{U_T}} - N_D \right) \cdot d\Psi$$

$$\left( \frac{Q_{Si}}{2 \cdot \varepsilon_{Si}} \right)^2 = \frac{2 \cdot e \cdot n_i \cdot U_T}{\varepsilon_{Si}} \cdot \left[ \left( e^{\frac{\Psi_s - V_{ch}}{U_T}} - e^{\frac{\Psi_0 - V_{ch}}{U_T}} \right) - \frac{N_D}{n_i} \cdot \frac{\Psi_s - \Psi_0}{U_T} \right] \quad (7.3)$$

où  $Q_{Si}$  est la densité de charge dans le canal de silicium,  $\Psi_s = \Psi(W_{Si}/2)$  est le potentiel de surface et  $\Psi_0 = \Psi(0)$  est le potentiel au centre du canal. Ainsi, sans tenir compte du signe de la charge ionisée, l'équation différentielle gouvernant la distribution de la charge (voir relation (7.1)) est apparemment la même que celle du DG MOSFET dopé classique [8]. Dans ce dernier, la concentration du dopage est prise en compte dans une définition d'une largeur équivalente de silicium (voir section 2.3). Puisque de bons résultats ont été obtenus par cette méthode, nous pouvons nous demander si cette définition est toujours un bon concept pour les dispositifs sans jonction, même en sachant que le principe de fonctionnement est totalement différent. En reprenant la définition de la largeur équivalente de silicium  $W_{eq}$  [8] pour des couches de silicium dopées de type N, cela donne :

$$\frac{2}{W_{eq}} = \left[ \int_{-W_{Si}/2}^0 e^{\frac{-e \cdot N_D}{2 \cdot \varepsilon_{Si} \cdot U_T} \cdot x^2} dx \right]^{-1} - \frac{e \cdot N_D \cdot W_{Si}}{2 \cdot \varepsilon_{Si} \cdot U_T} \quad (7.4)$$

Selon (7.4), nous trouvons que dans les architectures sans jonction,  $W_{eq}$  peut prendre une valeur négative pour de relativement grandes concentrations de dopage, donnant lieu à des densités de charges imaginaires et donc non physiques (Figure 7.3). Comme cela sera confirmé au long du chapitre, bien que cette définition soit liée à un modèle en particulier, ce non sens physique révèle que ces dispositifs sans jonction demandent une approche de modélisation différente. Celle-ci est développée dans les sections suivantes.

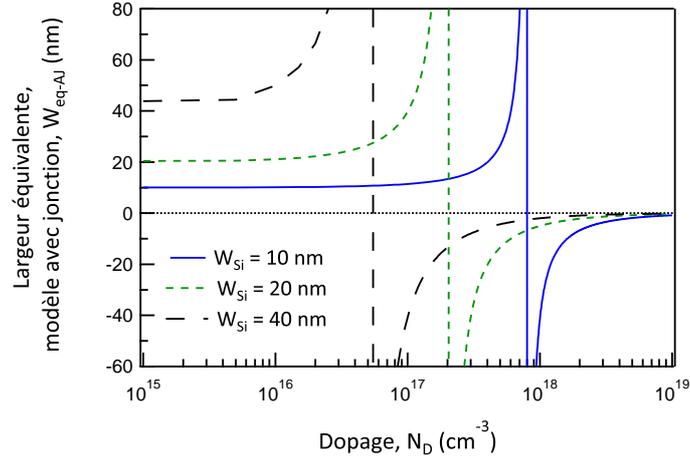


Figure 7.3 – Largeur équivalente de silicium calculée selon (7.4) pour différentes largeurs physiques de silicium.

### 7.2.2 Solution approchée de l'équation de Poisson

Nous considérons un semi-conducteur dopé de type N, et notre analyse porte sur les modes de déplétion et d'accumulation qui sont les modes pertinents du DG MOSFET sans jonction.

Comme cela a déjà été mentionné, en négligeant la densité des trous et en supposant le semi-conducteur non-dégénéré, l'équation de Poisson mène à l'équation différentielle (7.1), qui à notre connaissance n'a pas de solution exacte. Avant d'aller plus loin, il est nécessaire d'évaluer la limite de validité des statistiques de Boltzmann adoptées dans (7.1), puisque de grandes concentrations de dopage seront à l'étude. Idéalement, les statistiques de Boltzmann ne devraient pas être utilisées lorsque la densité de porteurs libres excède la densité effective d'état. Du point de vue des électrons dans le silicium, cela doit correspondre à environ  $3 \times 10^{19} \text{ cm}^{-3}$  à température ambiante. Cependant, comme nous le verrons, nous avons constaté que l'utilisation des statistiques de Boltzmann avec des niveaux de dopage tels que  $5 \times 10^{19} \text{ cm}^{-3}$  n'introduit pas d'erreurs substantielles. Il est à préciser que cela est également constaté dans les canaux de MOSFET bulk conventionnel pour lesquels la densité de feuille de charge peut aller au-delà de cette limite.

Maintenant, nous proposons de rechercher une solution approchée de (7.1) en s'appuyant sur une méthode de « différence-finie » en employant un pas de discrétisation grossier. Selon les valeurs discrètes représentées à la Figure 7.2, la dérivée seconde du potentiel au centre du canal peut être exprimée selon :

$$\left. \frac{d^2\Psi}{dx^2} \right|_{x=0} = \left( \frac{\Psi\left(-\frac{W_{Si}}{2}\right) - \psi(0)}{\frac{W_{Si}}{2}} - \frac{\psi(0) - \Psi\left(\frac{W_{Si}}{2}\right)}{\frac{W_{Si}}{2}} \right) \cdot \frac{1}{\frac{W_{Si}}{2}} \quad (7.5)$$

De plus, en fonctionnement symétrique, le champ électrique au centre vaut zéro, ce qui donne :

$$\left. \frac{d\Psi}{dx} \right|_{x=0} = \frac{\Psi\left(\frac{W_{Si}}{2}\right) - \psi\left(-\frac{W_{Si}}{2}\right)}{W_{Si}} = 0 \quad (7.6)$$

Ainsi selon (7.6),  $\Psi\left(\frac{W_{Si}}{2}\right) = \psi\left(-\frac{W_{Si}}{2}\right) = \Psi_S$ . Et en associant finalement les relations (7.1) et

(7.5), nous aboutissons à l'expression de la différence de potentiel entre la surface et le centre du canal :

$$(\Psi_s - \Psi_0) = \frac{e \cdot W_{Si}^2}{8 \cdot \varepsilon_{Si}} \cdot \left( n_i \cdot e^{\frac{\Psi_0 - V_{ch}}{U_T}} - N_D \right) \quad (7.7)$$

Puisque (7.7) lie le potentiel de surface au potentiel au centre du canal, une fois que  $\Psi_0$  est évalué,  $\Psi_s$  peut être calculé.

Nous poursuivons maintenant avec quelques remarques qui peuvent déjà être faites sur cette nouvelle approche. Lorsque le pas de discrétisation de valeur  $W_{Si}/2$  est remplacé par la coordonnée spatiale  $x$ , (7.7) peut être considéré comme une généralisation de la bien connue approximation de déplétion qui a été utilisée dans [8], et qui est rappelée ici :

$$(\Psi(x) - \Psi_0) = \frac{e \cdot x^2}{2 \cdot \varepsilon_{Si}} \cdot \left( n_i \cdot e^{\frac{\Psi_0 - V_{ch}}{U_T}} - N_D \right) \quad (7.8)$$

En effet, lorsque la densité de charges mobiles est négligeable devant la densité de charges fixes, cette relation revient à la solution exacte de l'équation de Poisson dans les semi-conducteurs totalement déplétés. De même, lorsque les charges mobiles ne peuvent pas être négligées, nous nous retrouvons avec un profil de potentiel parabolique selon  $x$ . Néanmoins, il est intéressant de remarquer que cette forme parabolique n'est pas imposée a priori, mais résulte de la discrétisation de l'équation de Poisson (une équation du second ordre) sur trois points seulement. Nous verrons que la densité de charges totale ne sera pas obtenue par l'intégration de la densité des porteurs sur la parabole, mais qu'elle s'appuiera sur une approche plus élaborée qui, d'une certaine façon relâchera l'approximation. En outre, plus la couche de semi-conducteur sera mince et plus l'approximation sera valide, étant donné que le pas de discrétisation deviendra plus petit. Lorsque le dispositif opère en accumulation plutôt qu'en déplétion, le profil de potentiel a une forme non parabolique davantage prononcée. Ce comportement n'est pas cohérent avec la forme de 7.8, cependant il convient de relativiser ce problème. En effet, les dispositifs qui seront réellement utilisés seront fortement dopés, dans le but d'atteindre de grandes performances, et il est peu probable d'avoir besoin que les transistors opèrent en forte accumulation. Ils fonctionneront surtout en régime de déplétion, et à sa limite qu'est la condition de bandes plates, le courant est déjà important et cohérent avec un courant  $I_{ON}$  que peut demander un concepteur. De plus, la forme parabolique de la relation (7.7) n'est utilisée que pour définir le potentiel au centre du canal en fonction du potentiel de surface, qui lui est imposé par la tension de grille à travers la capacité d'oxyde de grille. En effet, nous ne passons pas par l'établissement d'une expression des potentiels du canal entre son centre et sa surface, tel que peut le permettre la relation (7.8). Enfin, selon l'hypothèse de la non parabolicité du profil de potentiel, cela n'engendre qu'une erreur sur le potentiel au centre du canal, qui a très peu d'influence en forte accumulation sur la densité de charge, et donc le courant. Ainsi, nous nous attendons à ce que cette approche soit toujours assez juste pour prédire les caractéristiques de dispositifs fortement dopés, et de dopages cohérents avec des applications réelles.

Selon (7.3), la densité de charges dans le silicium peut être directement exprimée en fonction

du potentiel au centre du canal et du potentiel de Fermi, c'est-à-dire :

$$Q_{Si} = 2 \cdot \text{signe}(\Psi_0 - \Psi_s) \cdot \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{\left( e^{\frac{\Psi_s - V_{ch}}{U_T}} - e^{\frac{\Psi_0 - V_{ch}}{U_T}} \right) - \frac{N_D}{n_i} \cdot \frac{\Psi_s - \Psi_0}{U_T}} \quad (7.9)$$

où la fonction *signe* est cohérente avec la densité de charge du silicium. En effet, contrairement au DG MOSFET classique,  $Q_{Si}$  peut être positif ou négatif, impliquant respectivement un potentiel de surface plus petit ou plus grand que celui du centre étant donnée l'équation de Poisson.

Ainsi, un cas limite existe à la condition de bandes plates. Selon (7.1), la neutralité du silicium est atteinte lorsque  $d^2\Psi/dx^2 = 0$ , c'est-à-dire lorsque :

$$\Psi_0 = \Psi_{0FB} = V_{ch} + U_T \cdot \ln\left(\frac{N_D}{n_i}\right) \quad (7.10)$$

Ici,  $\Psi_{0FB}$  se réfère au potentiel de surface à la condition de bandes plates (FB). Lorsque  $\Psi_0 = \Psi_{0FB}$ , la relation (7.1) impose  $\Psi(x) = \Psi_{0FB}$  pour toutes valeurs de  $x$ , c'est-à-dire que localement une neutralité prend place. Le potentiel est ainsi constant à travers toute la couche de silicium ( $\Psi_s = \Psi_0$ ).

### 7.2.3 Approche générale du calcul de la charge mobile

La charge totale  $Q_{Si}$  est composée des densités de charges mobiles  $Q_m$  et de charges fixes des atomes donneurs ionisés :

$$Q_{Si} = Q_m + e \cdot N_D \cdot W_{Si} \quad (7.11)$$

D'autre part, la densité de charge du silicium est aussi reliée à la chute de potentiel à travers l'isolant de grille, c'est-à-dire :

$$(V_G - \Delta\Phi - \Psi_s) \cdot 2 \cdot C_{ox} = -Q_{Si} \quad (7.12)$$

En associant les relations (7.12) et (7.9) selon une égalité sur la charge totale, nous pouvons écrire la relation suivante :

$$(V_G - \Delta\Phi - \Psi_s) \cdot 2 \cdot C_{ox} = -2 \cdot \text{signe}(\Psi_0 - \Psi_s) \cdot \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{\left( e^{\frac{\Psi_s - V_{ch}}{U_T}} - e^{\frac{\Psi_0 - V_{ch}}{U_T}} \right) - \frac{N_D}{n_i} \cdot \frac{\Psi_s - \Psi_0}{U_T}} \quad (7.13)$$

Ensuite, en remplaçant dans (7.13) les termes  $\Psi_s$  du potentiel de surface par son expression issue de la relation (7.7), et en définissant  $C = e^{\frac{\Psi_0 - V_{ch}}{U_T}}$  et  $K = (e \cdot n_i \cdot W_{Si}^2) / (8 \cdot U_T \cdot \varepsilon_{Si})$ , nous aboutissons à une relation implicite entre les tensions appliquées au dispositif, et le potentiel

au centre du canal exprimé par l'intermédiaire de la variable  $C$  :

$$\begin{aligned}
 V_G - V_{ch} - \Delta\Phi = & U_T \cdot \ln(C) \\
 & + K \cdot U_T \cdot \left( C - \frac{N_D}{n_i} \right) - \text{signe}(\Psi_0 - \Psi_s) \cdot \frac{\sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T}}{C_{ox}} \\
 & \cdot \sqrt{C \cdot \left[ \exp \left( K \cdot \left( C - \frac{N_D}{n_i} \right) \right) - \frac{N_D}{n_i} \cdot K - 1 \right] + \left( \frac{N_D}{n_i} \right)^2 \cdot K} \quad (7.14)
 \end{aligned}$$

Pour des potentiels donnés de grille et de canal, la variable  $C$  peut être numériquement résolue avec des routines standards tel que Newton-Raphson. Enfin, la densité de charge est directement obtenue à partir de l'expression (7.12) qui intègre l'expression du potentiel de surface toujours issue de (7.7). La densité de charge est ainsi calculée en fonction de  $C$ , et son expression est la suivante :

$$Q_{Si} = -2 \cdot C_{ox} \cdot \left( V_G - V_{ch} - \Delta\Phi + K \cdot U_T \cdot \frac{N_D}{n_i} - K \cdot U_T \cdot C - U_T \cdot \ln(C) \right) \quad (7.15)$$

Ici, il est intéressant de remarquer que la forme parabolique du potentiel a simplement été utilisée pour relier le potentiel de surface à celui du centre, permettant de calculer le potentiel au centre du canal par (7.14) à partir de (7.13). En effet, la relation (7.13) a été obtenue à partir d'une intégration exacte de l'équation de Poisson-Boltzmann (7.1). Comme discuté précédemment, ceci est une différence majeure en comparaison d'une approche entièrement faite par différences finies, demandant d'imposer un profil de potentiel pour calculer la densité de charges à travers le silicium.

Nous ne sommes pas encore au stade où nous disposons d'une expression explicite de la charge. Toutefois, notre approche représente une formulation intéressante qui évite la résolution de l'équation différentielle du second ordre non linéaire, tout en incluant la contrainte résultant de la tension de grille [voir (7.1) et (7.12)]. Nous avons vérifié que (7.14) et (7.15) s'applique aussi à des dispositifs presque non dopés (Figure 7.6(d)), c'est-à-dire pour des concentrations du dopage du silicium aussi faible que  $10^{15} \text{ cm}^{-3}$ . L'application à des DG MOSFET à base de jonctions en prenant  $-N_A$  au lieu de  $N_D$  est tout à fait envisageable, mais n'est pas étudiée ici.

Cependant, nous allons voir que la formulation obtenue jusqu'à maintenant demande des simplifications si l'on souhaite exprimer le courant de manière analytique en fonction de la charge. Ainsi, partant de (7.14), nous proposons de développer de nouvelles expressions pour la relation charge-potential, menant à des formulations plus explicites.

## 7.2.4 Expressions approchées pour les charges

Des expressions simplifiées reliant charges et potentiels sont maintenant développées en fonction du ratio entre la densité de charges mobiles et celle de charges fixes. Par la suite, nous allons distinguer les modes de fonctionnement du transistor selon qu'il est en accumulation ou en déplétion.

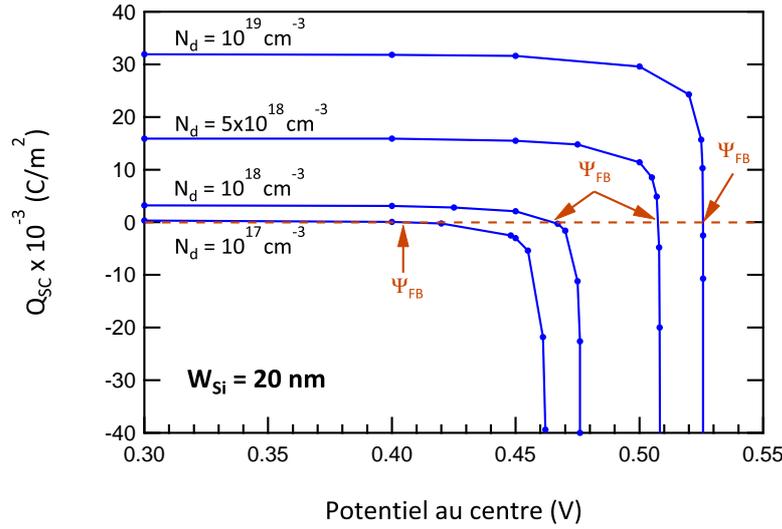


Figure 7.4 – Densité de charge du semi-conducteur en fonction du potentiel au centre d’une couche de silicium de 20 nm de largeur, d’un DG MOSFET sans jonction selon différentes concentrations de dopage. Les flèches indiquent le potentiel de surface correspondant à la condition de bandes plates.

**Le DG MOSFET sans jonction en accumulation :** En résolvant numériquement l’équation de Poisson-Boltzmann (7.1), il ressort que le potentiel au centre du canal atteint une asymptote lorsque le transistor entre en mode d’accumulation, soit lorsque  $\Psi_0$  évolue au-dessus de  $\Psi_{0FB}$ . Dans les DG MOSFET à jonctions et non dopés, une limite supérieure est en effet prédite à partir de la solution exacte [9] (solution exacte qui n’existe pas pour le cas dopé). D’un point de vue physique, cette limite peut être interprétée comme un écran du champ électrique dû à la forte densité des porteurs mobiles dans le silicium. Ainsi nous trouvons que le potentiel au centre reste très proche de sa valeur à la condition de bandes plates, soit  $\Psi_0 \simeq \Psi_{0FB}$ . Cette hypothèse est d’autant plus valide que le dopage est important. La Figure 7.4 montre le potentiel au centre qui varie en fonction de la densité de charges mobiles pour différentes concentrations de dopage. Par exemple, en prenant une couche de silicium de 20 nm dopée jusqu’à  $5 \times 10^{18} \text{ cm}^{-3}$ , la borne supérieure de  $\Psi_0$  est trouvée à 0.508 V, soit presque  $\Psi_{0FB}$ . De plus, le graphe suggère une accentuation de l’asymptote verticale, lorsque la concentration du dopage augmente. Ceci est un résultat assez intéressant qui va nous aider à développer une expression approchée pour la condition où les potentiels du canal sont supérieurs au potentiel des bandes plates.

L’hypothèse développée précédemment est appliquée à l’expression (7.9), dans laquelle  $\Psi_0$  est donc remplacée par  $\Psi_{0FB}$ , et nous mène à une expression de la densité de charge valide en accumulation (rappel : en accumulation  $\Psi_0 - \Psi_s \leq 0$ ) :

$$Q_{Si}^{\text{acc}} \simeq -2 \cdot \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{\left( e^{\frac{\Psi_s - V_{ch}}{U_T}} - \frac{N_D}{n_i} \right) - \left[ \frac{N_D}{n_i} \cdot \left( \frac{\Psi_s - V_{ch}}{U_T} - \ln \left( \frac{N_D}{n_i} \right) \right) \right]} \quad (7.16)$$

L’expression (7.16) peut être simplifiée en étudiant le ratio entre le terme du côté gauche de la

seconde racine carrée et celui du côté droit, soit :

$$\left( e^{\frac{\Psi_s - V_{ch}}{U_T}} - \frac{N_D}{n_i} \right) / \left[ \frac{N_D}{n_i} \cdot \left( \frac{\Psi_s - V_{ch}}{U_T} - \ln \left( \frac{N_D}{n_i} \right) \right) \right] \quad (7.17)$$

Après quelques transformations, ce ratio peut être exprimé sous la forme suivante :

$$\frac{\left( e^{\frac{\Psi_s - \Psi_{0FB}}{U_T}} - 1 \right)}{\frac{\Psi_s - \Psi_{0FB}}{U_T}} \quad (7.18)$$

En accumulation ( $\Psi_S > \Psi_{0FB}$ ), nous remarquons que ce ratio est toujours plus grand que un (ce ratio devient un à la condition de bandes plates), et qu'il augmente rapidement avec l'augmentation du potentiel de surface. Ainsi, nous proposons d'omettre le terme du côté droit dans l'expression de la densité de charge, sans que cela n'ait de conséquence sur le calcul de la charge au cas limite de la condition de bandes plates, à laquelle la valeur du terme omis est nulle. En mode d'accumulation, (7.16) se simplifie donc en :

$$Q_{Si}^{\text{acc}} \simeq -2 \cdot \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{\left( e^{\frac{\Psi_s - V_{ch}}{U_T}} - \frac{N_D}{n_i} \right)} \quad (7.19)$$

Selon (7.19), nous pouvons vérifier que la condition de neutralité est toujours satisfaite lorsque  $\Psi_s = \psi_0$ .

En exprimant  $\Psi_s$  en fonction de  $Q_{Si}$  à partir de (7.19), puis en le remplaçant dans (7.12), nous obtenons après quelques manipulations une relation charge-potentiel implicite :

$$V_G - V_{ch} - \Delta\Phi - U_T \cdot \ln \left( \frac{N_D}{n_i} \right) \stackrel{\text{acc}}{\simeq} \frac{-Q_{Si}}{2 \cdot C_{ox}} + U_T \cdot \ln \left( 1 + \frac{Q_{Si}^2}{8 \cdot e \cdot N_D \cdot \varepsilon_{Si} \cdot U_T} \right) \quad (7.20)$$

Selon (7.20), la condition de bandes plates ( $Q_{Si} = 0$ ) correspond à  $V_G - V_{ch} - \Delta\Phi = U_T \cdot \ln(N_D/n_i)$ , ce qui est le résultat attendu.

**Le DG MOSFET sans jonction en déplétion :** Dans le cas de la déplétion, le potentiel au centre du canal devient plus grand que le potentiel de surface. Le silicium devient donc chargé positivement ( $Q_{Si} \geq 0$ ). Nous exprimons dans un premier temps la densité de charges en fonction du potentiel au centre du canal, en remplaçant l'expression de  $\Psi_S$  issue de (7.7) dans l'expression de la charge (7.9) :

$$Q_{Si}^{\text{dep}} \stackrel{\text{dep}}{=} 2 \cdot \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{C \cdot \left[ \exp \left( K \cdot \left( C - \frac{N_D}{n_i} \right) \right) - \frac{N_D}{n_i} \cdot K - 1 \right] + \left( \frac{N_D}{n_i} \right)^2 \cdot K} \quad (7.21)$$

Le terme exponentiel de (7.21) peut s'écrire sous la forme suivante :

$$\exp \left( K \cdot \left( e^{\frac{\Psi_0 - V_{ch}}{U_T}} - e^{\frac{\Psi_{0FB} - V_{ch}}{U_T}} \right) \right) \quad (7.22)$$

et nous remarquons qu'en déplétion il est toujours plus petit que un. Cela permet une nouvelle simplification. En effet, en considérant nulle la valeur de  $\left(\exp\left(K \cdot \left(C - \frac{N_D}{n_i}\right)\right) - 1\right)$  dans la seconde racine carré, de bons résultats ont été obtenus. Cette simplification suppose  $C = N_D/n_i$  dans le terme exponentiel de (7.21), soit une sorte de condition de bandes plates aboutissant à l'expression suivante :

$$Q_{Si}^{\text{dep}} \simeq 2 \cdot \sqrt{2 \cdot e \cdot n_i \cdot \varepsilon_{Si} \cdot U_T} \cdot \sqrt{K \cdot \frac{N_D}{n_i} \cdot \left(\frac{N_D}{n_i} - C\right)} \quad (7.23)$$

En exprimant la variable  $C$  en fonction de la charge totale dans le semi-conducteur définie par (7.23), nous avons :

$$C \stackrel{\text{dep}}{\simeq} \frac{N_D}{n_i} \cdot \left(1 - \left(\frac{Q_{Si}}{e \cdot N_D \cdot W_{Si}}\right)^2\right) \quad (7.24)$$

En remplaçant (7.24) dans (7.15) et en réarrangeant les termes, nous obtenons une relation charge-potentiel implicite valide en mode de déplétion :

$$V_G - V_{ch} - \Delta\Phi - U_T \cdot \ln\left(\frac{N_D}{n_i}\right) \stackrel{\text{dep}}{\simeq} -\frac{Q_{Si}^2}{8 \cdot e \cdot N_D \cdot \varepsilon_{Si}} - \frac{Q_{Si}}{2 \cdot C_{ox}} + U_T \cdot \ln\left(1 - \left(\frac{Q_{Si}}{e \cdot N_D \cdot W_{Si}}\right)^2\right) \quad (7.25)$$

Tout comme dans le cas de l'accumulation, (7.25) prédit que la tension de grille satisfait  $V_G - V_{ch} - \Delta\Phi = U_T \cdot \ln(N_D/n_i)$  à la condition de bandes plates ( $Q_{Si} = 0$ ). Ce point est très important puisqu'il assure que la continuité est satisfaite à la transition entre la déplétion et l'accumulation. Les équations (7.20) et (7.25) sont ainsi compatibles.

Ces approximations « régionales » ont été évaluées par rapport à la solution « exacte » donnée par (7.15) et (7.14). Les résultats de ces calculs sont tracés à la Figure 7.5. Les cercles présents sur la figure indiquent la transition entre les modes d'accumulation et de déplétion, dont les tracés sont respectivement obtenus par les relations (7.20) et (7.25). Il sort que les approximations proposées jusqu'ici sont suffisamment précises pour être appliquées.

**Evaluation de la densité de charges mobiles** Afin de vérifier la validité de notre approche, nous avons exécuté des simulations numériques TCAD 2D. La mobilité des électrons a été fixée constante à  $0,1 \text{ m}^2/V.s$ . La longueur et la largeur du canal du dispositif sont de  $1 \mu\text{m}$  pour éviter des effets de canaux courts et étroits, et l'épaisseur d'oxyde est de  $1,5 \text{ nm}$ . Nous avons ignoré la quantification du potentiel dans le canal. Finalement, le même jeu de paramètres a été utilisé pour les simulations TCAD, et celles du modèle sans l'introduction d'aucun paramètre empirique.

La densité de charges mobiles évaluée à partir de (7.20) et (7.25) a été comparée avec des simulations TCAD de DG MOSFET sans jonction pour différentes largeurs de silicium et différentes concentrations du dopage. Le potentiel de quasi-Fermi du canal est imposé à  $0 \text{ V}$ .

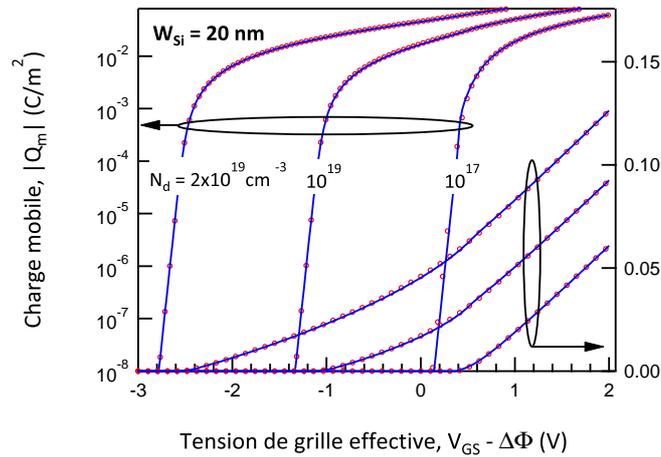


Figure 7.5 – Comparaison de la densité de charge mobile dans un DG MOSFET de 20 nm de largeur obtenue (cercles) des relations (7.14) et (7.15), et (lignes pleines) des approximations « régionales » données par les relations (7.20) et (7.25).

Les Figures 7.6(a, b, c) correspondent respectivement aux largeurs de silicium de 10, 20 et 40 nm.

Tant sur l'échelle linéaire que logarithmique, l'accord entre le modèle et les simulations TCAD est bon. En particulier, le coude, visible sur l'échelle linéaire et qui apparaît juste sous la condition de bandes plates, est bien pris en compte par le modèle. La pente de la charge en fonction de la tension de grille ne correspond plus à la capacité d'oxyde de grille  $C_{ox}$ , lorsque la tension est bien au-dessus du seuil (nous définirons les conditions du seuil par la suite). Cette différence de comportement avec le DG MOSFET à base de jonctions est une des particularités majeures du DG MOSFET sans jonction. Ce nouveau type de dispositif ne peut pas être modélisé par le modèle dans [8] ou tout autres approches pour dispositifs à base de jonctions.

Néanmoins, nous pouvons mentionner que l'existence du coude est moins prononcée pour de fines largeurs de silicium et/ou de faibles concentrations du dopage. Dans ce cas, il se peut qu'un modèle de DG MOSFET à base de jonctions puisse prédire les caractéristiques électriques de dispositifs sans jonction faiblement dopés. Cependant avec une telle solution il n'est plus possible de prédire les caractéristiques directement avec les paramètres technologiques.

En forte déplétion, la densité de charges mobiles varie selon une pente de 60 mV/dec à température ambiante, tout comme le DG MOSFET à base de jonctions en régime de faible inversion.

Pour les couches de 20 et 40 nm, nous observons qu'une limite supérieure de la concentration du dopage conditionne la déplétion totale dans le barreau de silicium. Au-delà de cette limite (Figure 7.7), le dispositif n'est plus capable d'atteindre un état off. Sous conditions d'équilibre, cela vient de la couche d'inversion (de trous dans notre cas) qui se forme aux interfaces Si/SiO<sub>2</sub>. L'écran de la couche d'inversion sur le champ électrique de la grille a par exemple lieu autour de  $V_G - \Delta\Phi \simeq -2V$  dans des largeurs de 20 nm de silicium dopé à  $2 \times 10^{19} \text{ cm}^{-3}$ , et autour de  $V_G - \Delta\Phi \simeq -1V$  dans des largeurs de 40 nm de silicium dopé à  $5 \times 10^{18} \text{ cm}^{-3}$ . Ce comportement n'est pas prédit par le modèle qui suggère toujours que la déplétion totale peut être atteinte

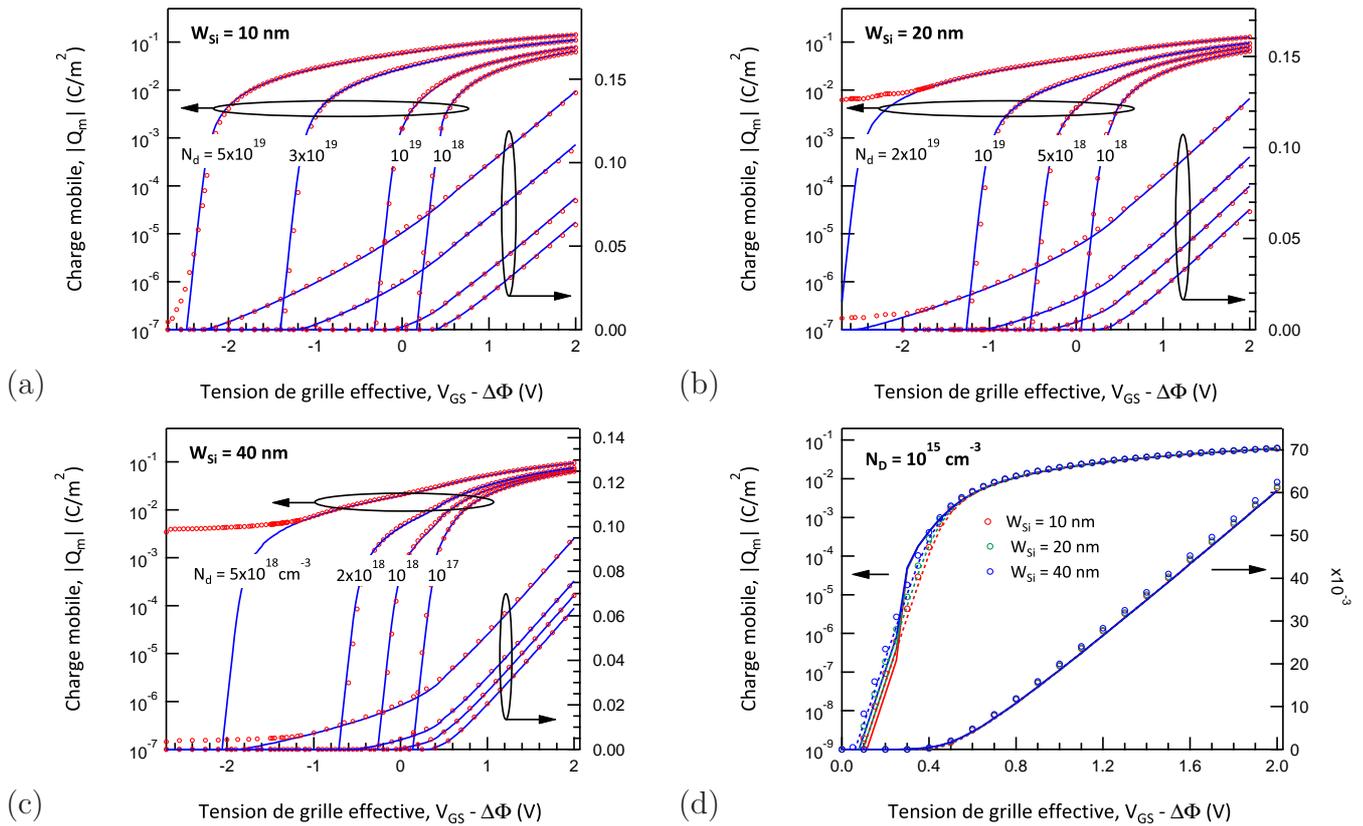


Figure 7.6 – Densité de charge mobile en fonction de la tension de grille pour un DG MOSFET sans jonction, de largeur de silicium de (a) 10 nm, (b) 20 nm, (c) 40 nm pour différentes concentrations de dopages, et (d) de faible de concentration de dopage pour différentes largeurs de silicium. Lignes pleines : modèle avec approximations régionales, lignes pointillées : modèle exact, cercles : simulations TCAD.

même pour de faibles tensions de grille. Cela représente une limite intrinsèque au modèle dont on doit être conscient lorsqu'on traite des dispositifs sans jonction larges et fortement dopés.

### 7.2.5 Tension de seuil dans les transistors sans jonction

Contrairement au DG MOSFET classique, la notion de tension de seuil est beaucoup moins évidente dans le cas des dispositifs sans jonction. En effet, extrapoler la densité de charges mobiles de l'asymptote presque linéaire de la région de forte inversion, n'est plus valide lorsque nous considérons de forts niveaux de dopage et/ou d'épaisseurs couches de silicium. Par exemple, pour un DG MOSFET sans jonction de 20 nm, les Figures 7.5 et 7.6(b) mettent en évidence deux pentes dans la dépendance Q-V, toujours pour les tensions inférieures à la tension de bandes plates. Cette caractéristique inhabituelle des transistors double grille vient du terme au carré dans (7.25) lorsque le dispositif opère en déplétion, régime à atteindre pour commuter à l'état off le canal.

Cependant, nous pouvons toujours définir une tension de seuil comme étant la tension de grille qui annule la densité de charges majoritaires mobiles lorsque le terme logarithmique est négligé. Dans ce cas, nous ignorons simplement la déplétion totale, région de fonctionnement dans laquelle la densité de charges mobiles varie exponentiellement avec la tension de grille. Il

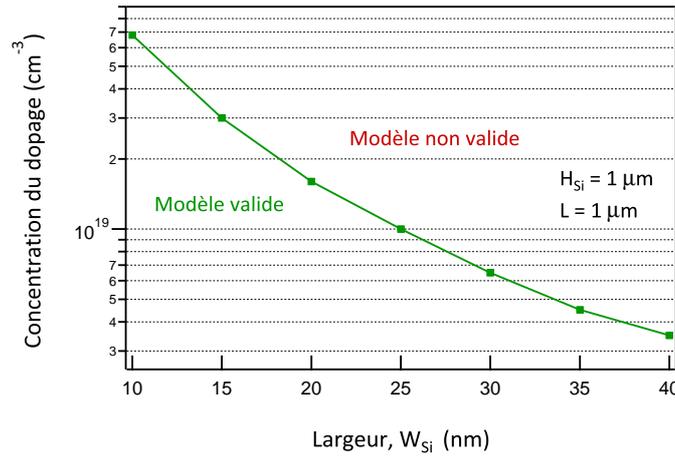


Figure 7.7 – *Domaine de validité du modèle : limite supérieure de la concentration de dopage en fonction de la largeur de silicium pour laquelle le canal peut atteindre le régime de déplétion totale. Résultats obtenus à partir de simulations TCAD 2D.*

est important de remarquer que cette définition revient à celle utilisée pour les DG MOSFET à base de jonctions dans [8].

En adoptant cette définition, et en forçant donc  $Q_m = 0$  dans (7.11), la relation (7.25) devient :

$$V_T - \Delta\Phi - U_T \cdot \ln\left(\frac{N_D}{n_i}\right) = -\frac{(e \cdot N_D \cdot W_{Si})^2}{8 \cdot e \cdot N_D \cdot \epsilon_{Si}} - \frac{e \cdot N_D \cdot W_{Si}}{2 \cdot C_{ox}}$$

nous permettant d'aboutir à l'expression de tension de seuil suivante :

$$V_T = \Delta\Phi + U_T \cdot \ln\left(\frac{N_D}{n_i}\right) - e \cdot N_D \cdot W_{Si} \cdot \left(\frac{1}{2 \cdot C_{ox}} + \frac{1}{8 \cdot C_{Si}}\right) \quad (7.26)$$

Selon (7.26), la tension de seuil conjugue une dépendance linéaire et une dépendance logarithmique, vis-à-vis de la concentration du dopage. La Figure 7.8 révèle que ces contributions varient dans des sens opposés. Parmi les deux termes de l'expression, nous distinguons le terme "électrostatique" qui seul varie avec la largeur de silicium, et le terme « semi-conducteur » indépendant des paramètres technologiques.

Tant que la densité de dopage reste inférieure à  $10^{18} \text{ cm}^{-3}$ , la tension de seuil est en légère augmentation avec l'augmentation du niveau de dopage, puis une fois la valeur  $10^{18} \text{ cm}^{-3}$  dépassée, elle chute rapidement. Ce comportement inhabituel d'une variation non monotone de la tension de seuil avec le dopage est assez intéressant. Il peut être exploité pour trouver la meilleure combinaison largeur - dopage étant donnée la tension de seuil. Cette liberté est sans équivalent par rapport aux MOSFET à base de jonctions. Pour une largeur de silicium donnée, il est également possible de trouver un niveau de dopage qui engendre la même tension de seuil que dans un DG MOSFET non dopé à base de jonction et de mêmes dimensions.

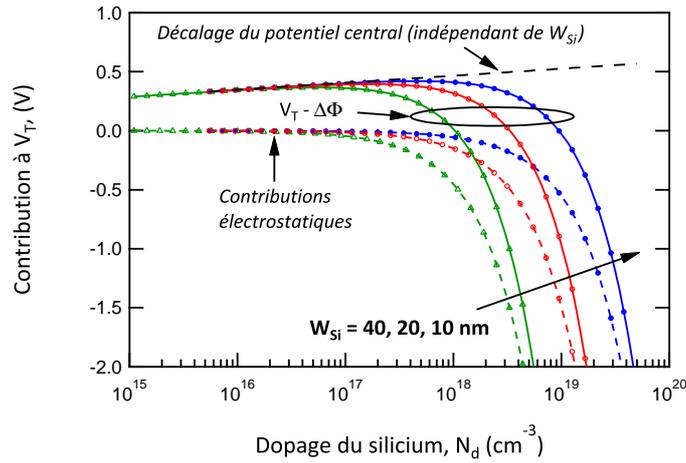


Figure 7.8 – Contributions de la tension de seuil exprimée par (7.26) en fonction de la concentration du dopage dans le DG MOSFET sans jonction pour les largeurs de silicium 10 nm, 20 nm et 40 nm. Lignes pleines :  $V_T - \Delta\Phi$ . Lignes pointillées avec symboles : contribution électrostatique [dernier terme de (7.26)]. Lignes pointillées sans symboles : potentiel central [avant-dernier terme de (7.24), indépendant de la largeur de silicium  $W_{Si}$ ]

### 7.3 Développement de l'expression de courant

Comme pour le DG MOSFET à base de jonctions [8], nous supposons que la densité de courant peut être calculée selon le modèle de « drift-diffusion ». Dans ce cas, le courant est donné par :

$$I_D = -H_{Si} \cdot \mu \cdot Q_m \cdot \frac{dV_{ch}}{dy} \quad (7.27)$$

où  $\mu$  est la mobilité des porteurs supposée constante le long du canal,  $Q_m$  est la densité de charges mobiles locale. En introduisant l'expression de la densité de charges mobiles issue de (7.11) dans (7.27), puis en l'intégrant de la source au drain, nous obtenons la relation générale suivante :

$$\begin{aligned} I_D &= \frac{H_{Si}}{L} \cdot \mu \cdot \int_S^D (e \cdot N_D \cdot W_{Si} - Q_{Si}) \cdot dV_{ch} \\ &= \frac{H_{Si}}{L} \cdot \mu \cdot e \cdot N_D \cdot W_{Si} \cdot V_{DS} - \frac{H_{Si}}{L} \cdot \mu \cdot \int_S^D Q_{Si} \cdot dV_{ch} \end{aligned} \quad (7.28)$$

Malheureusement, la relation assez complexe entre la densité de charge  $Q_{Si}$  et le potentiel  $V_{ch}$  telle qu'obtenue de la dépendance charge-potentiel générale (7.14), ne mène à aucune expression analytique du courant. Par contre, à partir des expressions simplifiées développées pour les modes d'accumulation et de déplétion, nous proposons de développer une expression analytique complète pour le courant valide dans toutes les régions de fonctionnement.

### 7.3.1 Courant en accumulation

En suivant l'approche qui vient d'être présentée, nous pouvons maintenant utiliser (7.20) pour calculer le courant en accumulation. L'expression de la dérivée de (7.20) par rapport à  $Q_{Si}$  nous donne le terme différentiel de l'intégrale de (7.28), soit :

$$Q_{Si} \cdot dV_{ch} \stackrel{\text{acc}}{\simeq} \frac{Q_{Si} \cdot dQ_{Si}}{2 \cdot C_{ox}} - \frac{\frac{Q_{Si}^2 \cdot dQ_{Si}}{4 \cdot e \cdot N_D \cdot \varepsilon_{Si}}}{1 + \frac{Q_{Si}^2}{8 \cdot e \cdot N_D \cdot \varepsilon_{Si} \cdot U_T}} \quad (7.29)$$

Après intégration, nous obtenons :

$$\int_S^D Q_{Si} \cdot dV_{ch} \Big|_{\text{acc}} = \frac{1}{4 \cdot C_{ox}} \cdot Q_{Si}^2 \Big|_S^D - 2 \cdot U_T \cdot Q_{Si} \Big|_S^D + 2 \cdot U_T \cdot \sqrt{8 \cdot e \cdot N_D \cdot \varepsilon_{Si} \cdot U_T} \cdot \arctan \left( \frac{Q_{Si}}{\sqrt{8 \cdot e \cdot N_D \cdot \varepsilon_{Si} \cdot U_T}} \right) \Big|_S^D \quad (7.30)$$

Ensuite, en insérant cette expression dans (7.28), le courant en accumulation peut être calculé à partir des valeurs des densités de charges (mobiles et fixes) évaluées par (7.20) à la source et au drain.

### 7.3.2 Courant en déplétion

De même, nous exploitons (7.25) pour évaluer le terme différentiel de (7.28), soit :

$$Q_{Si} \cdot dV_{ch} \stackrel{\text{dep}}{\simeq} \frac{Q_{Si}^2 \cdot dQ_{Si}}{4 \cdot e \cdot N_D \cdot \varepsilon_{Si}} + \frac{Q_{Si} \cdot dQ_{Si}}{2 \cdot C_{ox}} + U_T \cdot \frac{\frac{2 \cdot Q_{Si}^2 \cdot dQ_{Si}}{(e \cdot N_D \cdot W_{Si})^2}}{1 - \left( \frac{Q_{Si}}{e \cdot N_D \cdot W_{Si}} \right)^2} \quad (7.31)$$

Comme pour l'accumulation, l'expression analytique suivante est obtenue après intégration :

$$\int_S^D Q_{Si} \cdot dV_{ch} \Big|_{\text{dep}} = \frac{1}{12 \cdot e \cdot N_D \cdot \varepsilon_{Si}} \cdot Q_{Si}^3 \Big|_S^D + \frac{1}{4 \cdot C_{ox}} \cdot Q_{Si}^2 \Big|_S^D - 2 \cdot U_T \cdot Q_{Si} \Big|_S^D + U_T \cdot (e \cdot N_D \cdot W_{Si}) \cdot \left( \ln \left( 1 + \frac{Q_{Si}}{e \cdot N_D \cdot W_{Si}} \right) - \ln \left( 1 - \frac{Q_{Si}}{e \cdot N_D \cdot W_{Si}} \right) \right) \Big|_S^D \quad (7.32)$$

En insérant cette expression dans (7.28) et en évaluant les charges à la source et au drain à partir de (7.25), le courant peut être calculé en déplétion.

### 7.3.3 Traitement général du courant dans le DG MOSFET sans jonction

Dans les dernières sous-parties, nous avons été capable d'exprimer le courant dans un canal de silicium de la source au drain en accumulation ou en déplétion. Cependant, selon les

potentiels appliqués au dispositif, une portion du canal peut être en accumulation (du côté de la source) tandis que l'autre est en déplétion (du côté du drain). Cette sorte d'état de canal hybride requiert un traitement spécial. Sachant que la transition a lieu à la tension de bandes plates, nous pouvons identifier trois situations différentes (en supposant toujours  $V_{DS} > 0$ ).

**Canal entièrement en accumulation** Pour que le canal entier soit en accumulation, nous devons satisfaire l'inégalité suivante  $V_G - V_D - \Delta\Phi \geq U_T \cdot \ln(N_D/n_i)$ , impliquant aussi que  $V_G - V_S - \Delta\Phi \geq U_T \cdot \ln(N_D/n_i)$  puisque  $V_D > V_S$ . Ensuite nous avons :

$$I_D = \frac{H_{Si}}{L} \cdot \mu \cdot e \cdot N_D \cdot W_{Si} \cdot V_{DS} - \frac{H_{Si}}{L} \cdot \mu \cdot \left( \int_S^D Q_{Si} \cdot dV_{ch} \Big|_{acc} \right) \quad (7.33)$$

**Canal entièrement en déplétion** De même, en déplétion, nous devons satisfaire  $V_G - V_S - \Delta\Phi \leq U_T \cdot \ln(N_D/n_i)$ , impliquant aussi que  $V_G - V_D - \Delta\Phi \leq U_T \cdot \ln(N_D/n_i)$ . L'expression du courant est dans ce cas :

$$I_D = \frac{H_{Si}}{L} \cdot \mu \cdot e \cdot N_D \cdot W_{Si} \cdot V_{DS} - \frac{H_{Si}}{L} \cdot \mu \cdot \left( \int_S^D Q_{Si} \cdot dV_{ch} \Big|_{dep} \right) \quad (7.34)$$

**Canal « hybride »** Le cas intermédiaire entre ces deux comportements asymptotiques est le cas le plus important. Une portion du canal est alors en accumulation, c'est-à-dire de la source à une position correspondant à l'établissement des bandes plates dans le canal, et le reste du canal est en déplétion, c'est-à-dire de la position des bandes plates au drain.

Ce cas spécial se produit lorsque  $V_G - V_S - \Delta\Phi \geq U_T \cdot \ln(N_D/n_i)$  et  $V_G - V_D - \Delta\Phi \leq U_T \cdot \ln(N_D/n_i)$ . La première inégalité exprime le fait que le semi-conducteur est en accumulation au niveau de la source, alors que la seconde indique que le semi-conducteur est déplété au niveau du drain.

Ainsi, le terme intégrale dans (7.28) doit être décomposé en deux parties distinctes, pour écrire le courant total comme la somme de deux courants distincts qui sont chacun évalués soit par (7.30), soit par (7.32). Nous avons donc le courant total suivant quelle que soit la polarisation du transistor :

$$I_D = \frac{H_{Si}}{L} \cdot \mu \cdot e \cdot N_D \cdot W_{Si} \cdot V_{DS} - \frac{H_{Si}}{L} \cdot \mu \cdot \left( \int_S^{FB} Q_{Si} \cdot dV_{ch} \Big|_{acc} + \int_{FB}^D Q_{Si} \cdot dV_{ch} \Big|_{dep} \right) \quad (7.35)$$

Remarquons que la densité de charge évaluée à  $V_{FB}$  donne par définition,  $Q_{Si} = 0$ . Ainsi le calcul des deux termes intégrales ne demande pas plus de calculs comparé au calcul du courant d'un canal entièrement en accumulation ou en déplétion.

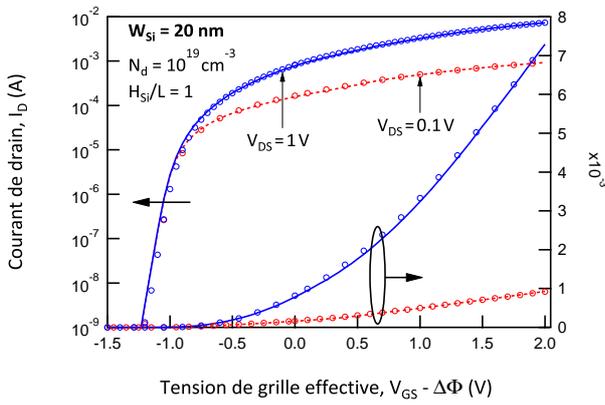


Figure 7.9 – Courant de drain en fonction de la tension de grille en régimes linéaire et saturé dans un DG MOSFET sans jonction de 20 nm de largeur de silicium et dopé à  $10^{19} \text{ cm}^{-3}$ . (Lignes pleines) modèle, (lignes pointillées) simulations TCAD.

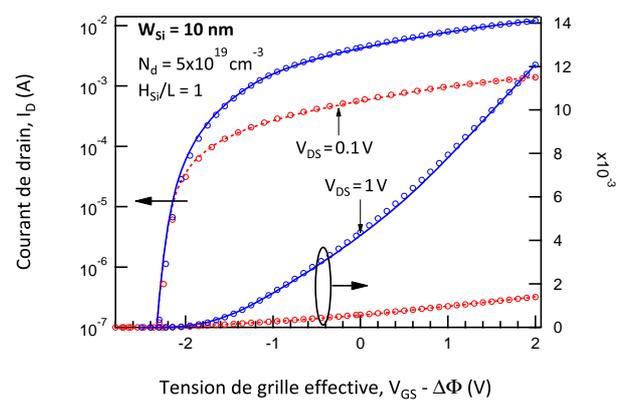


Figure 7.10 – Courant de drain en fonction de la tension de grille en régimes linéaire et saturé dans un DG MOSFET sans jonction de 10 nm de largeur de silicium et dopé à  $5 \times 10^{19} \text{ cm}^{-3}$ . (Lignes pleines) modèle, (lignes pointillées) simulations TCAD.

### 7.3.4 Comparaison avec des simulations numériques

La Figure 7.9 montre le courant en fonction de la tension de grille à tension de drain faible ( $V_{DS} = 0,1 \text{ V}$ ) et forte ( $V_{DS} = 1 \text{ V}$ ) et pour un dispositif dont la largeur de silicium est de 20 nm et le dopage de  $10^{19} \text{ cm}^{-3}$ . Selon la valeur de la tension de grille  $V_G$ , le canal passe par l'état « hybride », en partie déplété, en partie en accumulation. La concordance entre modèle et simulations numériques est suffisamment bonne de la déplétion totale à l'accumulation. Comme la densité de charges a été correctement prédite à la condition de bandes plates pour les modes de déplétion et d'accumulation, il n'y a pas de discontinuité de courant lors du passage d'un mode à l'autre.

De même, la Figure 7.10 confirme que nous obtenons de bon résultats pour une largeur de silicium de 10 nm malgré le très haut niveau de concentration du dopage de  $5 \times 10^{19} \text{ cm}^{-3}$ , valeur proche de celle rapportée dans [2].

Le léger écart de tension de seuil visible entre modèle et simulations, sur les graphes de courant précédemment évoqués, reste très faible comparé à la variation de la tension seuil en fonction du niveau de dopage représentée à la Figure 7.11. Selon cette figure, les tensions de seuil du modèle et des DG MOSFET sans jonction simulés par TCAD sont très similaires et leur variation en fonction du dopage de leur canal concorde. La Figure 7.12 rend directement compte de l'erreur de tension de seuil des caractéristiques de courant du modèle. Quelle que soit la largeur de silicium, l'erreur qui est contenue dans l'intervalle  $[-35 \text{ mV}; +18 \text{ mV}]$  est négligeable pour un modèle compact sans aucun paramètre empirique. Dans une volonté cependant, d'identifier précisément la source de la sous-estimation de la tension de seuil de plus en plus importante avec le niveau de dopage, nous avons étudié la qualité de la modélisation du potentiel central  $\Psi_0$  du canal en fonction du niveau de dopage. En effet, le potentiel central est le potentiel le plus élevé dans la largeur du canal dopé en déplétion, et à ce titre est une image du courant de drain en régime de déplétion. Nous nous attendons donc à trouver

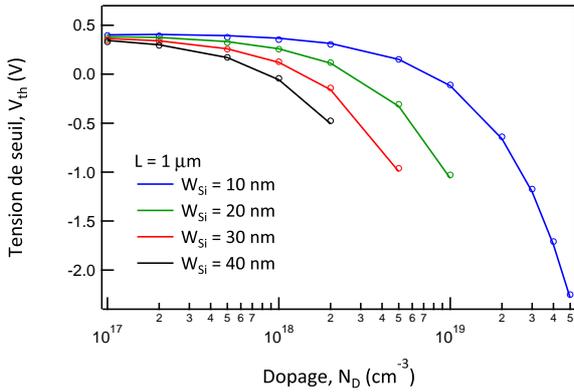


Figure 7.11 – Tension de seuil du DG MOS-FET sans jonction en fonction de la concentration du dopage pour différentes largeurs de silicium. (Lignes pleines) modèle, (cercles) simulations TCAD.

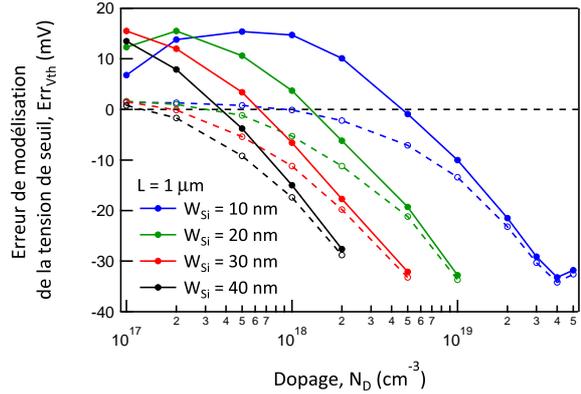


Figure 7.12 – Différence de tension de seuil du modèle par rapport aux simulations, du DG MOS-FET sans jonction en fonction de la concentration du dopage pour différentes largeurs de silicium. (Lignes pleines) modèle avec approximation « régionale », (lignes pointillées) modèle exact.

une sur-estimation de ce potentiel, qui implique une sur-estimation du courant soit une sous-estimation de la tension de seuil. Ce comportement du potentiel central est bien celui observé à la Figure 7.13 qui représente l'erreur de modélisation de  $\Psi_0$ . Parmi les relations à la base du modèle compact du transistor sans jonction, seule la relation 7.7, exprimant la différence de potentiel entre la surface et le centre du canal, est une relation non exacte. L'erreur, faible, induite dans cette relation par la méthode de résolution de l'équation de Poisson 7.1, est également représentée à la Figure 7.13, et représente en effet toute l'erreur de la modélisation de  $\Psi_0$  du modèle compact. Une différence cependant existe aux très forts niveaux de dopage en limite de fonctionnement du transistor, à laquelle nous n'avons pas d'explication.

La variation de l'erreur de modélisation du potentiel central explique exactement la variation de l'erreur de tension de seuil du modèle exact (lignes pointillées de la Figure 7.12). Cela n'est pas le cas de la tension de seuil du modèle avec approximations « régionales » aux relativement faibles niveaux de dopage. Cependant, la sous-estimation du courant de drain induite par l'approximation de (7.9) en (7.23) disparaît aux forts niveaux de dopage, les dopages auxquels serait exploité le transistor sans jonction dans des applications réelles. Le modèle avec approximations « régionales » se justifie ici pleinement pour permettre l'obtention d'une expression explicite du courant de drain en fonction de la charge.

Le courant en fonction de la tension de drain a aussi été traité, dans les Figures 7.14 et 7.15 pour respectivement des largeurs de silicium de 20 et 10 nm. Nous observons des différences entre le modèle et les simulations TCAD qui sont attribuées à la non prise en compte de la longueur d'overlap que nous ne modélisons pas. Ce désaccord ne provient pas des relations de l'approximation « régionale », puisqu'elle est toujours présente lorsque le courant est évalué par intégration numérique.

Des caractéristiques « petits signaux » ont été également simulées. La Figure 7.16 représente la transconductance de grille  $g_m$  pour une largeur de silicium de 10 nm avec différentes densités

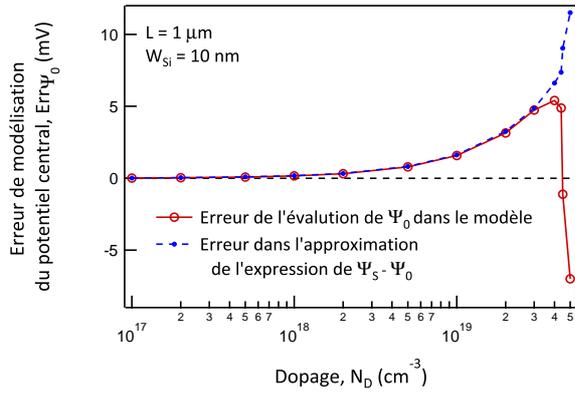


Figure 7.13 – Erreur de modélisation du potentiel central  $\Psi_0$  en régime de déplétion en fonction du niveau de dopage pour une largeur de silicium de 10 nm.

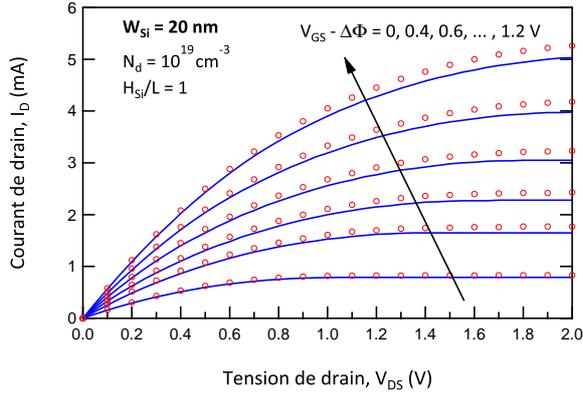


Figure 7.14 – Courant de drain en fonction de la tension de drain pour différentes tensions de grille dans un DG MOSFET sans jonction de 20 nm de largeur de silicium et dopé à  $10^{19} \text{ cm}^{-3}$ . (Lignes pleines) modèle, (lignes pointillées) simulations TCAD.

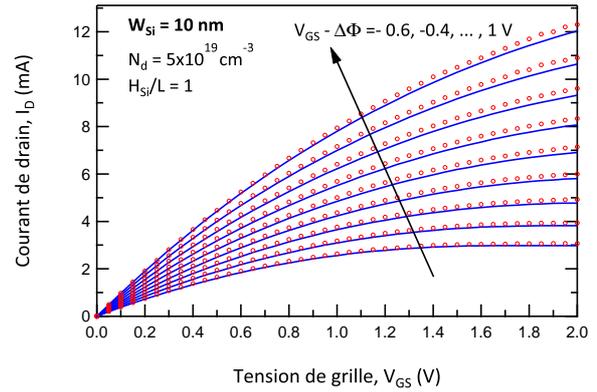


Figure 7.15 – Courant de drain en fonction de la tension de drain pour différentes tensions de grille dans un DG MOSFET sans jonction de 10 nm de largeur de silicium et dopé à  $5 \times 10^{19} \text{ cm}^{-3}$ . (Lignes pleines) modèle, (lignes pointillées) simulations TCAD.

de dopage, soient  $10^{19} \text{ cm}^{-3}$ ,  $3 \times 10^{19} \text{ cm}^{-3}$  et  $5 \times 10^{19} \text{ cm}^{-3}$ . Le potentiel de drain a été fixé à 1 V pour rendre possible l'observation de la transition entre les modes saturé et linéaire. Les courbes représentent les caractéristiques de transconductance de grille dérivées des simulations TCAD (points), du courant obtenu de (7.27) (pointillés), et du courant du modèle « hybride » évalué par (7.35) (lignes pleines). Des écarts apparaissent entre les simulations TCAD et le modèle original, c'est-à-dire sans utiliser les approximations régionales. Cependant, étant donné que les dérivées amplifient les petites variations et que nous n'avons introduit aucun paramètre empirique, ces résultats confirment que l'approche par différences finies selon un pas de discrétisation grossier est satisfaisante. Concernant le courant calculé à partir de l'approximation régionale, une bonne concordance est encore constatée avec les simulations TCAD. De plus, aucune discontinuité n'est observé sur  $g_m$ . De même que pour les densités de charges mobiles, les dérivées premières convergent vers les mêmes valeurs en bandes plates.

Ces transconductances de grille sont également instructives pour mettre en évidence les différences dans les caractéristiques de transfert entre le DG MOSFET sans jonction et celui à base de jonction. Les différentes pentes des caractéristiques  $I_D$ - $V_G$  sont clairement visibles à travers la transconductance de grille. Cette observation est particulière aux dispositifs sans jonction et ne peut pas être simulée avec un modèle de DG MOSFET standard.

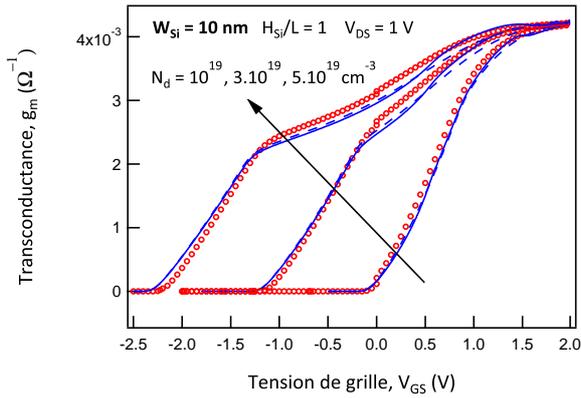


Figure 7.16 – Transconductance de grille en fonction de la tension de grille à  $V_{DS} = 1V$  pour un DG MOSFET sans jonction de largeur de 10nm et de différentes concentrations de dopage ( $10^{19} \text{ cm}^{-3}$ ,  $3 \times 10^{19} \text{ cm}^{-3}$  et  $5 \times 10^{19} \text{ cm}^{-3}$ ). Symboles : simulations TCAD. Lignes continues : modèle analytique (7.35) combiné avec (7.30) et (7.32). Lignes pointillées : modèle basé sur l'intégration numérique de (7.27).

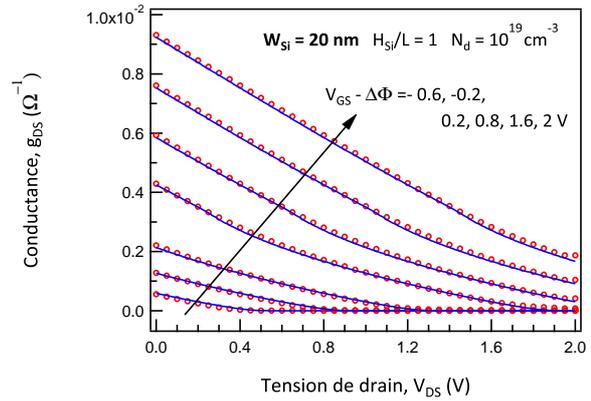


Figure 7.17 – Transconductance de drain en fonction de la tension de de drain pour différentes tensions de grille, pour un DG MOSFET sans jonction de largeur de 20nm dopé à  $10^{19} \text{ cm}^{-3}$ . Symboles : simulations TCAD. Lignes continues : modèle analytique (7.35) combiné avec (7.30) et (7.32).

Dans l'étude de la conductance de sortie  $g_{ds}$ , nous voyons que l'accord entre le modèle « hybride » et les simulations TCAD était bon, tel que le montre la Figure 7.17. Étant donnée cette constatation, nous avons seulement inclut dans la figure des simulations du modèle avec approximations régionales (7.35), puisqu'il n'y a presque pas de différence entre ce modèle et l'évaluation de  $g_{ds}$  par l'intégration numérique de (7.27).

## Conclusion

Nous avons proposé un modèle analytique pour calculer la densité de charge et le courant dans les transistors à effet de champ double grille symétriques sans jonction. Le modèle, publié dans [10], est valide dans toutes les régions d'opération, de la déplétion totale à l'accumulation, et du régime linéaire au saturé. Cela est confirmé par des comparaisons détaillées avec des simulations numériques TCAD. Aucun paramètre empirique n'a été utilisé, confirmant les racines physiques saines du corps du modèle. En particulier, l'existence de deux pentes distinctes dans la dépendance charge-potentiel a été très bien prédite, soit une différence majeure par rapport au DG MOSFET à base de jonctions. Le modèle présenté dans ce chapitre peut ainsi servir de base à de nouveaux développements qui prendrait en compte les effets canaux courts, la quantification de l'énergie des porteurs dans le canal et la dégradation de la mobilité, entre autres phénomènes pertinents à devoir être modélisés.

## Références

- [1] C. LEE, A. AFZALIAN, N. D. AKHAVAN, R. YAN, I. FERAIN et J. P. COLINGE, “Junctionless multigate field-effect transistor”, *Appl. Phys. Lett.*, vol. 94, n° 5, pp. 053 511–12–, 2009.
- [2] J. P. COLINGE, C.-W. LEE, A. AFZALIAN, N. D. AKHAVAN, R. YAN, I. FERAIN, P. RAZAVI, B. O’NEILL, A. BLAKE, M. WHITE, A.-M. KELLEHER, B. MCCARTHY et R. MURPHY, “Nanowire transistors without junctions”, *Nature Nanotechnology*, vol. 5, n° 3, pp. 225–229, 2010.
- [3] A. PFITZNER, M. STANIEWSKI et M. STRZYGA, “DC characteristics of junction vertical slit field-effect transistor (JVeSFET)”, dans *16th Int. Conf. MIXDES*, Lodz, Poland, 2009, pp. 420–423.
- [4] M. WEIS, A. PFITZNER, K. KASPROWICZ, Y.-W. LIN, T. FISCHER, R. EMLING, M. MAREK-SADOWSKA, D. SCHMITT-LANDSIEDEL et W. MALY, “Low power SRAM cell using vertical slit field effect transistor (VeSFET)”, dans *ESSCIRC Fringe P6*, 2008.
- [5] E. RAULY, B. IÑIGUEZ et D. FLANDRE, “Investigation of deep submicron single and double gate SOI MOSFETs in accumulation mode for enhanced performance”, *Electrochem. Solid-State Lett.*, vol. 4, n° 3, G28–G30, 2001.
- [6] B. IÑIGUEZ, B. GENTINNE, V. DESSARD et D. FLANDRE, “A physically-based continuous model for accumulation-mode SOI pMOSFETs”, *IEEE Trans. Electron Devices*, vol. 46, n° 12, pp. 2295–2303, 1999.
- [7] J.-M. SALLESE, F. KRUMMENACHER, F. PRÉGALDINY, C. LALLEMENT, A. ROY et C. ENZ, “A design oriented charge-based current model for symmetric DG MOSFET and its correlation with the EKV formalism”, *Solid-State Electron.*, vol. 49, n° 3, pp. 485–489, 2005.
- [8] J.-M. SALLESE, N. CHEVILLON, F. PRÉGALDINY, C. LALLEMENT et B. IÑIGUEZ, “The equivalent-thickness concept for doped symmetric DG MOSFETs”, *IEEE Trans. Electron Devices*, vol. 57, n° 11, pp. 2917–2924, 2010.
- [9] Y. TAUR, “Analytic solutions of charge and capacitance in symmetric and asymmetric double-gate MOSFETs”, *IEEE Trans. Electron Devices*, vol. 48, n° 12, pp. 2861–2869, 2001.
- [10] J.-M. SALLESE, N. CHEVILLON, C. LALLEMENT, B. IÑIGUEZ et F. PRÉGALDINY, “Charge-based modeling of junctionless double-gate field-effect transistors”, *IEEE Trans. Electron Devices*, vol. 58, n° 8, pp. 2628–2637, 2011.

# Conclusion et perspectives

Depuis le début de l'évolution de la microélectronique, la technologie MOS Bulk joue un rôle essentiel. Depuis une trentaine d'années, les effets néfastes liés à la réduction de la taille de transistor sont devenus de plus en plus importants et vont aboutir à la fin de la technologie MOS Bulk d'ici la fin de la décennie 2010 (cf. ITRS'11). La technologie multigrille devrait permettre de poursuivre cette évolution en attendant que la microélectronique passe à un autre matériau que le silicium. Grâce à son excellent contrôle du canal, la technologie SOI multigrille permet de réduire les effets néfastes liés à la réduction de la taille de transistor.

Dans le Chapitre 1 de ce manuscrit, nous avons présenté l'évolution de la technologie du transistor MOSFET qui est imposée par les effets néfastes de la réduction de sa géométrie. Du transistor MOSFET bulk planaire à l'origine de la microélectronique, nous sommes maintenant passés au transistor MOSFET bulk triple grille dans le cas de la dernière génération des processeurs Intel.

Les chapitres 2, 3 et 4 ont traité de la modélisation compacte des transistors MOSFET multigrilles. Leur relation est représentée par le schéma de la Figure 8.1. Nous sommes partis de la modélisation compacte du transistor MOSFET double grille canal long faiblement dopé qui est au cœur de nos travaux, pour développer des modèles compacts dont le domaine de validité s'applique à un grand nombre de structure de transistors multigrilles.

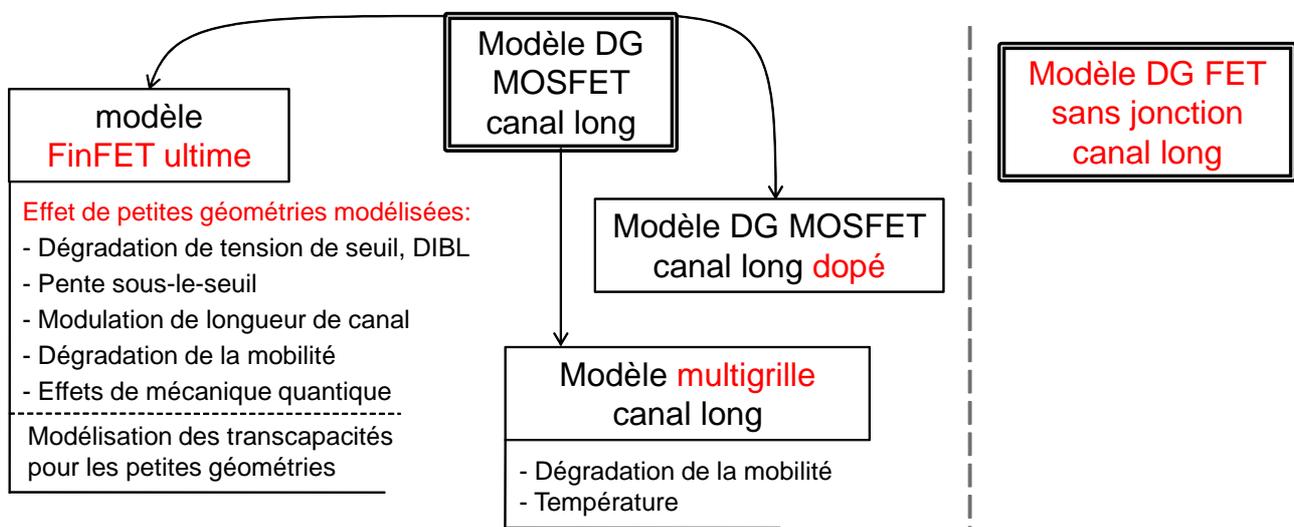


Figure 8.1 – Schéma de synthèse des travaux de modélisation développés dans cette thèse.

L'intérêt croissant porté ces deux dernières années sur un nouveau type de transistor, le transistor MOS double grille sans jonction, et sa similitude apparente avec le transistor MOSFET double grille, nous a conduit à développer un premier modèle de courant de drain de transistor multigrille sans jonction. Le chapitre 7 a proposé une approche de modélisation « régionale » pour prendre en compte le comportement spécifique du transistor sans jonction.

Le chapitre 5 a présenté les différents outils logiciels qui ont été nécessaires au développement des travaux de cette thèse. Cet environnement logiciel complet n'est pas indispensable au développement de modèles mais il permet de fortement l'accélérer.

Enfin le chapitre 6 a exploité le modèle compact du transistor FinFET implémenté en langage de description matériel (Verilog-A) pour réaliser des simulations de circuit logique. Nous avons pu tester notre modèle pour ce à quoi il est destiné. Nous avons poussé la modélisation jusqu'à la modélisation niveau-porte qui est indispensable à la simulation de circuits numériques complexes. Le développement d'une méthode d'étude statistique basée sur le modèle compact a permis de prendre en compte dans le modèle niveau-porte, la variabilité du processus de fabrication du transistor.

Les perspectives de ces travaux sont multiples. La première d'entre elles sera le développement d'un modèle compact de MOSFET multigrille incluant les effets canaux courts quelle que soit l'architecture multigrille. Des travaux préliminaires sur le MOSFET triple grille canal court ont déjà été effectués dans ce sens. Le modèle devra tenir compte du niveau de dopage du film de silicium, et ainsi permettre un ajustement de la tension de seuil des dispositifs multigrilles. Le modèle devra également être valide de la basse à la haute température. Les transistors multigrilles peuvent être réalisés sur des substrat bulk ou SOI. Le choix de la technologie du substrat devra être pris en compte dans le futur modèle compact.

La solution alternative que représente le transistor MOSFET sans jonction pourra être étudiée plus avant. Deux extensions de notre modèle sont à envisager. Une première permettrait l'extension du modèle double grille à des dispositifs multigrilles, et notamment triple grille. L'étude et la modélisation des effets canaux courts pourra constituer une part importante de futurs travaux.

Enfin, l'implémentation de tels modèles dans un langage de description matériel donnera les outils pour concevoir des circuits d'architectures innovantes. Le développement d'un modèle compact générique de transistors multigrilles ultimes permettra très aisément de comparer les performances de circuits entre les différentes architectures de transistors.

# Annexe A

## Simulation de circuits numériques

### A.1 Structure du code Verilog-A du modèle compact du transistor FinFET

Listing A.1 – Code Verilog-A du modèle compact du FinFET

```
//*****//
// Title: FinFET/DG-MOSFET Model Verilog-A source code
// Authors: Nicolas Chevillon, Fabien Pregaldiny, Morgan Madec,
// Christophe Lallement, Jean-Michel Sallese.
//
// Version Date
//-----
// 1.0 – Beta version February 16, 2012
// (Transcapacitances modeling including)
//
//*****//
// The Verilog-A code for the FinFET/DG-MOSFET model has been developed in
// cooperation among Institut d'Electronique du solide et des systemes
// (InESS) Universite de Strasbourg (UdS) / CNRS, Strasbourg, France and
// Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.
//
// Contact: Pr. Christophe Lallement, InESS
// christophe.lallement@ensps.u-strasbg.fr
//
// Main contributions:
//
// Coordination Pr. Christophe Lallement, InESS.
// Coding, implementation and testing Ph.D Nicolas Chevillon, Dr. Morgan
// Madec, InESS.
// Contributions to model formulation Ph.D Nicolas Chevillon, Dr. Fabien
// Pregaldiny, Pr. Christophe Lallement, InESS, Dr. Jean-Michel Sallese,
// EPFL.
//
// This model and its code are confidential property of the FinFET/DG-MOSFET
// model development team, and can be shared out only with the people of the
// COMON project.
// The model and its code are made available 'as is', without any warranty
// regarding its functionality or fitness of performance.
//*****//
```

```

‘include ”disciplines.vams”

‘define M_PI 3.141592653589793
‘define EPS0 8.854e-12
‘define EPSSI_REL 11.7
‘define EPSOX_REL 3.9
‘define H_Planck 6.626068e-34
‘define M_0 0.910956e-30

module finfetdg(d,g,s);

// External nodes declaration
inout d,g,s;
electrical d,g,s;
electrical di,si;
electrical igdyn,iddyn;

// Model parameter declaration
parameter real W = 20e-9 from [0:inf]; // Silicon width [m]
parameter real L = 0.1e-6 from [0:inf]; // Gate length [m]
parameter real H = 1e-6 from [0:inf]; // Silicon film thickness [m]
parameter integer type = 1 from [-1:1] exclude 0; // MOS type (N=1, P=-1)
parameter real tox = 1.5e-9 from [0:inf]; // Gate oxide thickness [m]
parameter real WFdiff = 0.0 from [0:inf]; // Work function difference [V]
parameter real mu0 = 0.1 from [0:1]; // Low-field mobility at 300K [m^2/(V*sec)]
parameter real E0 = 1.4 from [0:inf]; // Mobility parameter 1 for the degradation
// in strong [] inversion
parameter real E1 = 400.0 from [0:inf]; // Mobility parameter 2 for the degradation
// in strong [] inversion
parameter real n = 1.3 from [0:inf]; // Parameter of the temperature
// for the mobility model []
parameter real gamma = 2.0 from [0:inf]; // Parameter of the phonon scattering
// modeling term []
parameter real theta = 2.5 from [0:inf]; // CLM parameter []
parameter real Lov = 1e-9 from [0:inf]; // Overlap length [m]
parameter real Na = 5e20 from (0:inf); // Acceptor concentration [m^-3]
parameter real Nd = 5e27 from (0:inf); // Donor concentration [m^-3]
parameter real Rserial = 1e-9 from [0:inf]; // Serial resistance [Ohms]

//*** Definitions of functions ***
...

//*** Constant declaration ***
real kb, q, Tk, UT;
...

//*** Analog description of ultimate DG-FinFET ***
analog begin : DGFinFET_model

@(initial_step) begin
// Initialisation of physical constants and model constants
...
kb = 1.380658e-23; // Boltzmann's constant
q = 1.602e-19; // absolute charge of electron
Tk = $temperature; // ambient temperature in Kelvin
UT = kb*Tk/q; // thermodynamic voltage
...
end

```

```

// Normalized voltages on gate, source and drain
vgn = type * (V(g)-V(si) + WFdiff) / UT;
vdn = type * (V(di)-V(si)) / UT;
vsn = 0.0;

// Derivatives of voltages
V(tvg) <+ type * (ddt(V(g)) /1e9);
V(tvd) <+ type * (ddt(V(di)) /1e9);
V(tvs) <+ type * (ddt(V(si)) /1e9);

//*****
//*** Static model ***
//*****
...
// Determination of static drain current
ids = ...

//*****
//*** Dynamic model ***
//*****
...
// Determination of small signal currents
V(igdyn) <+ (Cgg * V(tvg) - Cgd * V(tvd) - Cgs * V(tvs)) * 1e9;
V(iddyn) <+ (-Cdg * V(tvg) + Cdd * V(tvd) - Cds * V(tvs)) * 1e9;
V(isdyn) <+ (-Csg * V(tvg) - Csd * V(tvd) + Css * V(tvs)) * 1e9;

I(g) <+ type * V(igdyn);
I(di) <+ type * ids + type * V(iddyn);
I(si) <+ -I(g) - I(di);

// Source serial resistance
I(s,si) <+ V(s,si) / (Rserial/2.0);

// Drain serial resistance
I(d,di) <+ V(d,di) / (Rserial/2.0);

end
endmodule

```

Listing A.2 – Bloc d'initialisation du modèle du transistor.

```

@(initial_step) begin
 // bloc d'équation définissant les constantes physiques
 // et les constantes propres au modèle
end

```

Pour un gain du temps de calcul lors de chaque itération de la simulation de circuit, il faut remarquer l'utilisation de la commande *initial\_step* (Listing A.2) sur tout le contenu du bloc d'initialisation des grandeurs électriques et physiques du modèle. Cette commande indique au simulateur de circuit que ce bloc de code ne devra être exécuté qu'une seule fois lors de toute la simulation du circuit.

## A.2 Code spectre de la simulation d'une porte NOT

Listing A.3 – Code spectre de la simulation d'une porte NOT.

```

Simulation Input File
.options
+ TEMP = 26.85 // Temperature=300K

simulator lang=spectre
saveOptions options save=lv1pub

parameters LG_W=10E-9
parameters LG_L=200E-9
parameters LG_H=60E-9
parameters LG_type=1
parameters LG_Rserial=1E1
parameters LG_Lov=3E-9;

simulator lang=spectre

ahdl_include "/home/chevillo/Verilog-A/DGFinFET/comon/notgate.va"
ahdl_include "/home/chevillo/SourcesModeleVA/capacitor.va"

subckt nfet (D G S)
parameters W=LG_W L=LG_L H=LG_H type=1 Lov=LG_Lov Rserial=LG_Rserial Mu0=0.1
Q1 D G S finfetdg W=W L=L H=H type=type Lov=Lov Rserial=Rserial mu0=Mu0
ends nfet

subckt cmosnotgate (A B GND VDD)
parameters W=LG_W L=LG_L H=LG_H Lov=LG_Lov Rserial=LG_Rserial Mu0=0.1
CMOS1 A B GND VDD notgate W=W L=L H=H Lov=Lov Rserial=Rserial mu0=Mu0
ends cmosnotgate

subckt resistance (A B)
parameters R=2e3
R12 A B res resist=R
ends resistance

subckt charge (A B)
parameters R=100e6
R34 A B res resist=R
ends charge

subckt Cout (A B)
parameters C=1e-15
C23 A B capa C=C
ends Cout

subckt Rinv (A B)
parameters R=2e3
R45 A B res resist=R
ends Rinv

RIN 3 6 resistance
GATE1 6 4 1 2 cmosnotgate
ROUT1 4 1 charge
DCKINOUT 5 4 1 nfet
RINVERTOR 5 2 Rinv
CINVOUT 5 1 Cout
RINVOUT 5 1 charge

```

```
simulator lang=spice
* START SOURCES
VSNGROUND 1 0 DC 0
VSPGROUND 2 0 DC 1.2
VINGROUND 3 0 PULSE(0 1.2 5p 5p 5p 20n 40n)
* END SOURCES

simulator lang=spectre
simulator lang=spice

* Paramètres de la simulation transitoire ".tran" :
* 1) "pas_de_temps_minimum"
* 2) "durée_de_simulation"
* 3) "instant_de_début_d'enregistrement_des_résultats"
* 4) "pas_de_temps_maximum"
.tran 10f 35p 0 0.5p
```



# Bibliographie personnelle

## Revue internationale

Jean-Michel Sallese, Nicolas Chevillon, Fabien Prégaldiny, Christophe Lallement and Benjamin Iniguez, “The equivalent thickness concept for doped symmetric DG MOSFETs,” *IEEE Transaction on Electron Devices*, vol. 57, no. 11, pp. 2917-2924, Nov. 2010.

Ashkhen Yesayan, Fabien Prégaldiny, Nicolas Chevillon, Christophe Lallement and Jean-Michel Sallese, “Physics-based compact model for ultra-scaled FinFETs”, *Solid-State Electronics*, vol. 62, no. 1, pp. 165-173, Apr. 2011.

Jean-Michel Sallese, Nicolas Chevillon, Christophe Lallement, Benjamin Iniguez and Fabien Prégaldiny, “Charge-Based Modeling of Junctionless Double-Gate Field-Effect Transistors”, *IEEE Transaction on Electron Devices*, vol. 58, no. 8, pp. 2628-2637, Aug. 2011.

Nicolas Chevillon, Jean-Michel Sallese, Christophe Lallement, Fabien Prégaldiny, Morgan Madec, Josef Sedlmeir and Jasmin Aghassi, “Generalization of the Concept of Equivalent Thickness and Capacitance to Multigate MOSFETs Modeling”, *IEEE Transaction on Electron Devices*, vol. 59, no. 1, pp. 60-71, Jan. 2012.

## Communications et workshop internationaux

Nicolas Chevillon, Mingchun Tang, Fabien Prégaldiny, Christophe Lallement and Morgan Madec, “FinFET compact modeling and parameter extraction,” *Proceeding IEEE Mixed Design of Integrated Circuits and Systems (MIXDES’09)*, pp. 55-60, Jun. 2009.

Ashkhen Yesayan, Nicolas Chevillon, Fabien Prégaldiny and Christophe Lallement, “Compact physics-based model for ultrashort FinFETs,” *Proceeding IEEE Mixed Design of Integrated Circuits and Systems (MIXDES’10)*, pp. 75-80, Jun. 2010.

Ashkhen Yesayan, Nicolas Chevillon, Fabien Prégaldiny, Morgan Madec, Christophe Lallement and Jean-Michel Sallese, “Physics-based compact model for ultimate FinFETs”, *Workshop MOS-AK*, Paris, 8 Apr. 2011.

Nicolas Chevillon, Fabien Prégaldiny, Christophe Lallement and Jean-Michel Sallese, “Standardization of Multigate MOSFET Modeling,” *Proceeding IEEE Mixed Design of Integrated Circuits and Systems (MIXDES’12)*, pp. 78-83, May 2012.

Nicolas Chevillon, Morgan Madec and Christophe Lallement, “Gate-Level Modeling for CMOS Circuit Simulation with Ultimate FinFETs,” *Proceeding IEEE International Symposium on Nanoscale Architectures (NANOARCH’12)*, pp. 22-29, Jul. 2012.

## Colloques nationaux

Nicolas Chevillon, Mingchun Tang, Fabien Prégaldiny, Christophe Lallement and Morgan Madec, “Modèle compact de FinFET et extraction de paramètres” Poster, *Journées-atelier des Groupements de Recherche SiP Soc et Nanoélectronique (GDR SiP-Soc/Nano)*, Bordeaux, 9-11 décembre 2009.

Nicolas Chevillon, Fabien Prégaldiny, Morgan Madec et Christophe Lallement, “Modélisation compacte du transistor FinFET pour la conception des circuits intégrés en technologie CMOS ultime”, Poster, *Journées Nationales du Réseau Doctoral en Micro-nanoélectronique (JNRDM'11)*, ENS Cachan, 23-25 mai 2011.



# Étude et modélisation compacte du transistor FinFET ultime

## Résumé :

Une des principales solutions technologiques liées à la réduction d'échelle de la technologie CMOS est aujourd'hui clairement orientée vers les transistors MOSFET faiblement dopés à multiples grilles. Ceux-ci proposent une meilleure immunité contre les effets canaux courts comparés aux transistors MOSFET bulk planaires (cf. ITRS 2011). Parmi les MOSFETs à multiples grilles, le transistor FinFET SOI est un candidat intéressant de par la similarité de son processus de fabrication avec la technologie des transistors planaires. En parallèle, il existe une réelle attente de la part des concepteurs et des fonderies à disposer de modèles compacts efficaces numériquement, précis et proches de la physique, insérés dans les « design tools » permettant alors d'étudier et d'élaborer des circuits ambitieux en technologie FinFET. Cette thèse porte sur l'élaboration d'un modèle compact orienté conception du transistor FinFET valide aux dimensions nanométriques. Ce modèle prend en compte les effets canaux courts, la modulation de longueur de canal, la dégradation de la mobilité, les effets de mécanique quantique et les transcapacités. Une validation de ce modèle est réalisée par des comparaisons avec des simulations TCAD 3D. Le modèle compact est implémenté en langage Verilog-A afin de simuler des circuits innovants à base de transistors FinFET. Une modélisation niveau-porte est développée pour la simulation de circuits numériques complexes. Cette thèse présente également un modèle compact générique de transistors MOSFET SOI canaux long faiblement dopés à multiple grilles. La dépendance à la température est prise en compte. Selon un concept de transformation géométrique, notre modèle compact du transistor MOSFET double grille planaire est étendu pour s'appliquer à tout autre type de transistor MOSFET à multiple grille (MuGFET). Une validation expérimentale du modèle MuGFET sur un transistor triple grille est proposée. Cette thèse apporte enfin des solutions pour la modélisation des transistors MOSFET double grille sans jonction.

**Mots-clés :** FinFET, MOSFET à multiple grilles, sans jonction, modèle compact, température, TCAD, Verilog-A, extraction de paramètres, simulation de circuits, modélisation niveau-porte.

---

## Study and compact modeling of ultimate FinFET transistor

### Abstract :

One of the main technological solutions related to downscaling of CMOS technology is now clearly oriented to lightly doped multigate MOSFETs. They offer better immunity against short channel effects compared to planar bulk MOSFETs (see ITRS 2011). Among the multigate MOSFETs, the SOI FinFET transistor is an interesting candidate because of the similarity of its manufacturing process with the planar transistor technology. In parallel, there is a real expectation on the part of designers and foundries to have compact models numerically efficient, accurate and close to the physics, and then inserted into the design tools in order to study and develop ambitious circuits in FinFET technology. This thesis focuses on the development of a design-oriented compact model of FinFET transistor valid to nanoscale dimensions. This model takes into account the short channel effects, the channel length modulation, the mobility degradation, the quantum mechanic effects and the transcapacitances. A validation of this model is carried out by comparisons with 3D TCAD simulations. The compact model is implemented in Verilog-A to simulate innovative FinFET-based circuits. A gate-level modeling is developed for the simulation of complex digital circuits. This thesis also presents a generic compact modeling of multigate SOI MOSFETs with lightly doped channels and temperature dependent. According to a concept of geometric transformation, our compact model of the planar double-gate MOSFET is extended to be applied to any other type of multigate MOSFETs (MuGFET). An experimental validation of the MuGFET compact model with a triple gate transistor is proposed. This thesis finally brings solutions for the modeling of junctionless double-gate MOSFET.

**Keywords :** FinFET, multigate MOSFET, junctionless FET, compact model, temperature, TCAD, Verilog-A, parameter extraction, circuit simulation, gate-level modeling.