



HAL
open science

Caractérisation électrique des transistors d'architecture innovante pour les longueurs de grilles décananométriques

Cheikh Diouf

► **To cite this version:**

Cheikh Diouf. Caractérisation électrique des transistors d'architecture innovante pour les longueurs de grilles décananométriques. Energie électrique. Université de Grenoble, 2013. Français. NNT : 2013GRENT082 . tel-01688481

HAL Id: tel-01688481

<https://theses.hal.science/tel-01688481v1>

Submitted on 19 Jan 2018

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Nano électronique et Nano Technologies**

Arrêté ministériel : 7 août 2006

Présentée par

Cheikh DIOUF

Thèse dirigée par **Gérard GHIBAUDO** et codirigée par **Antoine CROS**

Préparée au sein du **Laboratoire IMEP-LAHC** et de l'entreprise **STMicroelectronics** dans l'**École Doctorale EEATS** (**Electronique, Electrotechnique, Automatique et Traitement du Signal**)

Caractérisation électrique des transistors MOS déca-nanométriques d'architecture innovante

Thèse soutenue publiquement le « **31 octobre 2013** », devant le jury composé de :

Mr Georges BREMOND

Pr, INSA Lyon,

Président

Mr Pascal MASSON

Pr, Université de Nice Sophia Antipolis

Rapporteur

Mr Emmanuel DUBOIS

Dr, CNRS Lille

Rapporteur

Mr Stephane MONFRAY

ING, STMicroelectronics

Examineur

Mr Antoine CROS

ING, STMicroelectronics

Co-encadrant

Mr Gérard GHIBAUDO

Dr, CNRS Alpes

Directeur de thèse



REMERCIEMENTS

J'aimerais remercier toutes les personnes que j'ai pu croiser pendant ces trois ans, qu'elles soient de ST ou de l'IMEP. Ils ont été nombreux à m'apporter leur aide précieuse pour mon intégration, mon adaptation et la réussite de mon travail.

J'ai une pensée pour tout le personnel de l'équipe caractérisation électrique et fiabilité (ECR) de ST. J'ai apprécié leur disponibilité à mon arrivée pour faciliter ma formation aux bancs de mesures pour ne citer que cet exemple.

Merci à toutes les personnes avec qui j'ai pu collaborer pour produire ce bon travail: Paolo, Raphael, Imed, Xavier, Denis, Assawer, Sebastien...la liste est longue mais les autres s'y reconnaîtront.

Je tiens spécialement à remercier mon directeur de thèse Gérard Ghibaudo. Tous les thésards qu'il a eu à encadrer ont tous reconnu unanimement ses qualités techniques, ses valeurs humaines, sa réactivité et sa disponibilité. Je ne suis donc pas le premier à louer ses valeurs humaines et je ne serai certainement pas le dernier. J'ai été impressionné par mes mails qui ne restent pas deux heures sans réponses, les rappels immédiats à mes coups de fil manqués, mais aussi par ses corrections rapides et de grande qualité (articles ou chapitres de thèse). Gérard a aussi été très disponible à chaque fois que je l'ai sollicité pour un entrevu malgré son planning chargé de directeur du laboratoire en même temps.

Je remercie aussi mon encadrant industriel Antoine Cros pour sa confiance et son aide précieuse. Il a su faciliter mon intégration et mon adaptation au sein de l'entreprise. Il est aussi bon de rappeler que son apport technique m'a été très utile. Je suis conscient du fait que le temps est très précieux pour un ingénieur à ST mais il a su à chaque fois trouver le moment de faire avancer mon travail.

J'aimerais aussi remercier Yvan Morandini, Daniel Gloria et Michel Buczko pour avoir facilité notre collaboration pour produire le plus important résultat de ma thèse.

Merci aussi à Julien Rosa, pour ses qualités de manager. Je ne peux pas oublier les membres de l'équipe DDCC tels qu'Aurélie, Sebastien, Jean Philippe, Stéphane, Lama et Flore pour les weekly spécialement dédiés aux répétitions de mes présentations. Merci pour leur esprit critique.

Je ne peux pas oublier non plus Mustapha, Salim, Marine, Damien, Wafa, Anas, Franck, Emmanuel, Patrick... (collègues du repas de midi et/ou du foot à 7) pour les moments passés ensemble (au travail et en dehors du travail).

C'est maintenant le moment de remercier tous les membres de ma famille pour leur soutien inconditionnel (mon père Amadou, ma mère Fatou, mes grand-mères Ndeye Mali et Sokhna et mon grand-père Sidi). Sans eux je ne serai pas à ce niveau aujourd'hui. Bravo à mon père pour avoir tenu à me remotiver à chaque fois que j'ai failli dérailler. Même si je trouvais ces méthodes brusques sur le moment, mais je me rends compte maintenant que ce n'était que pour mon propre bien. Merci à lui pour sa vision. J'espère que de la haut il réalise que je suis allé jusqu'au bout et qu'il est fier de son fiston.

Table des matières générale

Table des matières générale

INTRODUCTION GENERALE	1
Partie A	5
Table des matières de la partie A	6
PARTIE A: GENERALITES	8
Chapitre A.I : RAPPEL SUR LE TRANSISTOR MOS	9
A.I.1) Principe de fonctionnement	9
A.I.1.a) Fonctionnement	9
A.I.1.b) Le dopage	10
A.I.2) Architecture de base et régimes de fonctionnement d'un transistor MOS	12
A.I.2.a) Architecture de base	12
A.I.2.b) Les régimes de fonctionnement	13
A.I.3) Les régimes de conduction	14
A.I.4) Les effets de canal court	16
A.I.4.a) Effet de la résistance source-drain	16
A.I.4.b) Influence des zones de déplétion source-drain	17
A.I.4.c) Le DIBL	18
A.I.5) Evolution de la technologie MOS : Du transistor sur substrat massif au FinFET en passant par le FDSOI	19
A.I.5.a) Les solutions apportées sur le matériau canal	19
A.I.5.b) L'empilement de grille	22
A.I.5.c) L'avènement d'autres matériaux pour remplacer le canal Si	23
A.I.5.d) Structures à électrostatique améliorée	26
A.I.5.e) Bilan et conclusion	29
Chapitre A.II : LE TRANSPORT ELECTRONIQUE	30

Table des matières générale

A.II.1) Transport des porteurs dans le canal	30
A.II.1.a) Le transport stationnaire : Du model de Drude au transport dérive diffusion	30
A.II.2) Notion de mobilité et son utilité pour étudier la performance	32
A.II.3) La notion de champ effectif et corrélation avec la mobilité effective.....	33
A.II.3.a) La notion de champ effectif.....	33
A.II.3.b) Corrélation mobilité effective versus champ effectif.....	33
A.II.4) Mécanismes de réduction de mobilité.....	34
A.II.4.a) Interaction avec les vibrations du réseau (phonons).....	34
A.II.4.b) Interactions de coulomb	36
A.II.4.c) La rugosité de surface.....	37
A.II.4.d) Les défauts cristallins non chargés (défauts neutres μ_N).....	37
A.II.5) Transport non stationnaire.....	37
A.II.5.a) Le phénomène de survitesse.....	38
A.II.5.b) Modèle de transport balistique (ou quasi-balistique)	39
A.II.5.c) Balisticité et dégradation de mobilité	42
A.II.5.d) Unification des mécanismes de transport : dérive diffusion et quasi-balistique	43
Table des figures de la partie A	45
Liste des tableaux de la partie A	47
Partie B	48
Table des matières de la partie B	49
PARTIE B: INFLUENCE DU PROCEDE DE FABRICATION ET DE L'ARCHITECTURE SUR LES PERFORMANCES DES TRANSISTORS	51
CHAPITRE B.I : LE RECUIT HAUTE PRESSION SOUS H ₂ OU D ₂	52
B.I.1) Performances électriques à température ambiante.....	53
B.I.1.a) Mobilité à faible charge	54

Table des matières générale

B.I.1.b) Résistance source-drain	55
B.I.1.c) Le pompage de charge pour la caractérisation de l'interface Si/SiO ₂	56
B.I.1.d) La variation de la tension de seuil après HPFA.....	57
B.I.2) Etude de la mobilité à faible charge à basse température	58
B.I.2.a) Description du modèle utilisé	58
B.I.2.b) Effets du HPFA sur les différentes contributions de mobilité à faible charge.....	60
B.I.3) Hydrogène ou Deutérium: Comparaison en mobilité effective, pompage de charge, HCI (hot carrier injection) et NBTI (Negative bias temperature instability)	62
B.I.3.a) Mobilité effective.....	62
B.I.3.b) HCI et NBTI.....	63
B.I.4) Bilan et conclusion	64
CHAPITRE B.II : ARCHITECTURE CANAL A BASE DE SILICIUM-GERMANIUM	65
B.II.1) Pourquoi introduire une proportion de germanium dans le canal?.....	65
B.II.2) Canal silicium-germanium	67
B.II.2.a) Simulations	67
B.II.2.b) Analyse à température ambiante	68
B.II.2.c) Analyse à basse température.....	74
B.II.2.d) Mesures de bruit basse fréquence.....	78
B.II.3) Discussion et conclusion	82
CHAPITRE B.III : L'OXYDE DE GRILLE : LE HIGH- κ ET SON IMPACT	84
B.III.1) Pourquoi l'introduction du high- κ ?.....	84
B.III.2) Les inconvénients du high-k	84
B.III.3) Les transistors de type GO1 et GO2	86
B.III.4) Etude comparative des transistors GO1 et GO2	87
B.III.5) Bilan et conclusion.....	95

Table des matières générale

Partie C	102
Table des matières de la partie C	103
PARTIE C: METHODOLOGIE D'EXTRACTION DES PARAMETRES: ETAT DES LIEUX ET AMELIORATIONS.....	105
CHAPITRE C.I : L'EXTENSION DE LA « FONCTION Y » POUR LE REGIME DE SATURATION	106
C.I.1) La « fonction Y » en linéaire	106
C.I.1.a) Définition de la « fonction Y »	107
C.I.1.b) Extraction de paramètres	107
C.I.2) La « fonction Y » en régime de saturation.....	112
C.I.2.a) La fonction Y du régime linéaire en régime de saturation.....	112
C.I.2.b) Mobilité apparente	116
C.I.2.c) Vitesse de saturation des porteurs en régime de saturation.....	116
C.I.2.c) Vitesse moyenne des porteurs en régime de saturation	118
C.I.3) Application : Comparaison de performance entre silicium-germanium et germanium massif ..	119
C.I.4) Bilan et conclusion	121
CHAPITRE C.II : A LA RECHERCHE DE BALISTICITE DANS LE GERMANIUM : LA METHODE DE LA « VITESSE LIMITE »	122
C.II.1) Les dispositifs étudiés	122
C.II.1.a) Procédé de fabrication [Mitard08].....	123
C.II.1.b) Les courants de fuites dans le germanium.....	125
C.II.2) Vérification des propriétés de transport sur nos dispositifs Ge.....	128
C.II.2.a) Mobilité effective	128
C.II.2.b) Estimation de la balisticité dans le Ge en régime linéaire.....	129
C.II.3) La méthode de la vitesse limitante: le v_{lim}	131
C.II.4) Amélioration de la méthode de « v_{lim} »	132

Table des matières générale

C.II.4.a) Extraction de la tension de seuil en saturation.....	132
C.II.4.b) Confirmation de la méthode d'extraction de V_{THsat} par des simulations TCAD	134
C.II.5) Application de la méthode de « v_{lim} » sur Ge-pMOSFET	138
C.II.5.a) Détermination de la nature du transport	138
C.II.6) Bilan et conclusion	141
CHAPITRE C.III : MESURES HF POUR LE TRANSPORT ELECTRONIQUE.....	142
C.III.1) Contexte	142
C.III.2) Problématique	143
C.III.3) Solutions techniques	143
C.III.3.a) Mesures basses fréquences.....	143
C.III.3.b) Mesures hautes fréquences (HF).....	145
C.III.3.c) Obtention des capacités et de la charge d'inversion.....	148
C.III.4) Extraction de mobilité et de vitesse	151
C.III.5) Application à l'extraction du v_{lim}	155
C.III.6) Bilan et conclusion.....	156
Table des figures de la partie C.....	157
Liste des tableaux de la partie C	161
Conclusion générale et perspectives	162
Première partie:	162
Deuxième partie:	163
Perspectives	165
Références bibliographiques.....	167
Publications de l'auteur.....	177

INTRODUCTION GENERALE

L'électronique qui représente un chiffre d'affaire de plus de deux cents milliards de dollars aujourd'hui, repose principalement sur la microélectronique dont le composant phare est le transistor MOSFET (Metal-Oxide-Semiconducteur field effect transistor).

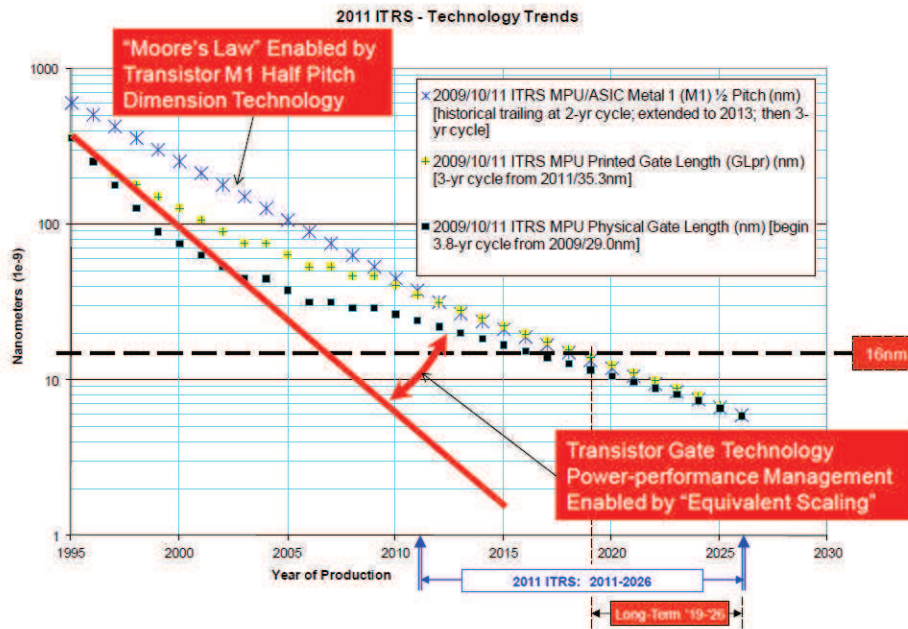
Tout a commencé en Décembre 1947 lorsque W. Shockley, J. Bardeen et W. Brattain découvrirent le transistor. Ces employés d'AT&T *Bell Laboratories* mettaient alors en place un composant qui va particulièrement révolutionner les moyens de communications et qui aujourd'hui s'impose sans appel comme un élément fondamental dans la conception des circuits électroniques.

En 1965, Gordon E. Moore prévoyait déjà une augmentation exponentielle de densité d'intégration de ces circuits électroniques [Moore65]. Aujourd'hui les fabricants de composants microélectroniques sont soumis à une forte concurrence, pour des applications de plus en plus diverses et variées (de l'informatique de bureau, la téléphonie mobile, les consoles de jeux vidéo, la photo numérique jusqu'aux secteurs de l'automobile, militaire mais aussi de l'aérospatiale). Au cours de ces dernières décennies, les fabricants de composants ont dû continuellement augmenter les performances des puces, tout en réduisant le coût de fabrication et leur taille.

La technologie CMOS (Complémentaire-Métal-Oxyde-Semiconducteur) représente près de 90% du marché des semi-conducteurs et poursuit sa route dans la miniaturisation qui amènera l'utilisation de dispositifs MOS de longueur de grille 20-14nm à l'horizon 2015-2020 au plan industriel comme l'illustre la [figure ci-dessous](#).

Cette miniaturisation globale entraîne aussi des effets néfastes sur certains paramètres électriques régissant le fonctionnement des transistors MOS. Pour contrecarrer ces effets, plusieurs solutions technologiques ont été proposées ces dernières années, soit en optimisant les architectures existantes, soit en proposant de nouvelles architectures ou bien en rajoutant des étapes de process lors de la fabrication. Afin de quantifier les avantages et les inconvénients de ses avancées technologiques sur les composants, il est indispensable de pouvoir les caractériser électriquement

de façon efficace ainsi que de modéliser le comportement de leurs paramètres électriques avec la réduction des dimensions.



Évolution de la longueur de grille des transistors (d'après la feuille de route ITRS 2011, actualisé 2012)

Le manuscrit est structuré en trois parties différentes :

- Une première partie traite les généralités du transistor MOS.
 - Dans un premier chapitre, nous parlerons de la théorie et du principe de fonctionnement. Nous y aborderons une problématique actuelle du transistor que sont les effets de canaux courts. Une liste des solutions technologiques sera énumérée.
 - Le deuxième chapitre abordera les différents modèles de transport (déplacement des électrons) existant dans la littérature.
- La deuxième partie de cette thèse examine l'influence du procédé de fabrication et de l'architecture sur les performances électriques du transistor MOS.

INTRODUCTION GENERALE

- Le chapitre I de la deuxième partie porte sur le recuit à haute pression sous enceinte hydrogène ou dihydrogène. L'évaluation du transistor après recuit standard et recuit à haute pression montre que le recuit à haute pression permet d'augmenter les performances.
- Dans le chapitre II, nous évaluons l'effet d'une technologie à base de germanium. L'introduction du germanium comme matériau de canal a pour but d'augmenter les performances sans recourir à la miniaturisation. Cela est possible grâce aux bonnes propriétés de transport du germanium comparé au silicium.
- Le troisième chapitre traite aussi de l'architecture du transistor MOS. Elle s'agit cette fois-ci de l'oxyde de grille. Le high-k a été introduit pour parer aux effets de miniaturisation de l'oxyde de grille. En revanche il s'est avéré qu'il avait des effets néfastes sur la mobilité des porteurs dans le canal.
- La troisième et dernière partie de cette thèse se concentre sur l'extraction de paramètres électriques permettant de caractériser le transport dans le MOS du régime linéaire jusqu'au régime d'utilisation, c'est à dire à forte tension de drain.
 - C'est ainsi que dans le chapitre I de la dernière partie, nous avons étendu la fonction Y jusqu'au régime de saturation. Cela a permis d'y extraire d'une manière fiable un indicateur de performance qui est mobilité apparente à bas champ. La vitesse de saturation a aussi été extraite en se servant de la mobilité apparente comme outil mathématique et en utilisant le concept de dégradation de la mobilité en fonction du champ longitudinale. Les applications qui en sont découlées ont montré toute l'importance d'extraire ces paramètres dans le même régime.
 - Le chapitre II de cette partie porte sur le modèle de transport quasi-balistique. Dans ce chapitre, nous traitons la méthode dite de « v_{lim} » introduite lors d'une thèse précédente. Après des améliorations apportées à cette méthode, elle est

INTRODUCTION GENERALE

appliquée à des technologies germanium et silicium sur substrat massif. Il en a découlé qu'il faudra élaborer des transistors plus courts que 60 nm dans le cadre du germanium sur substrat massif, et 30 nm dans le cadre du silicium-germanium pour espérer rentrer dans la limite quasi-balistique à température ambiante.

- Dans le dernier chapitre de cette dernière partie, nous abordons des mesures de capacité à haute fréquence pour non seulement améliorer la technique du split CV sur des dispositifs courts, mais aussi introduire la possibilité de mesurer des capacités à forte tension de drain. Cela nous a donc permis l'extraction de la vraie charge d'inversion en régime de saturation. Conséquemment, une vitesse moyenne est calculée d'une manière plus exacte.

Partie A

Table des matières de la partie A

Partie A.....	5
Table des matières de la partie A	6
PARTIE A: GENERALITES	8
Chapitre A.I : Rappel sur le transistor MOS.....	9
A.I.1) Principe de fonctionnement	9
A.I.1.a) Fonctionnement	9
A.I.1.b) Le dopage.....	10
A.I.2) Architecture de base et régimes de fonctionnement d'un transistor MOS.....	12
A.I.2.a) Architecture de base.....	12
A.I.2.b) Les régimes de fonctionnement.....	13
A.I.3) Les régimes de conduction.....	14
A.I.4) Les effets de canal court	16
A.I.4.a) Effet de la résistance source-drain	16
A.I.4.b) Influence des zones de déplétion source-drain	17
A.I.4.c) Le DIBL	18
A.I.5) Evolution de la technologie MOS : Du transistor sur substrat massif au FinFET en passant par le FDSOI.....	19
A.I.5.a) Les solutions apportées sur le matériau canal.....	19
A.I.5.b) L'empilement de grille.....	22
A.I.5.c) L'avènement d'autres matériaux pour remplacer le canal Si.....	23
A.I.5.d) Structures à électrostatique améliorée.....	26
A.I.5.e) Bilan et conclusion.....	29
Chapitre A.II : Le transport électronique.....	30

Table des matières de la partie A

A.II.1) Transport des porteurs dans le canal	30
A.II.1.a) Le transport stationnaire : Du model de Drude au transport dérive diffusion	30
A.II.2) Notion de mobilité et son utilité pour étudier la performance	32
A.II.3) La notion de champ effectif et corrélation avec la mobilité effective	33
A.II.3.a) La notion de champ effectif	33
A.II.3.b) Corrélation mobilité effective versus champ effectif	33
A.II.4) Mécanismes de réduction de mobilité	34
A.II.4.a) Interaction avec les vibrations du réseau (phonons)	34
A.II.4.b) Interactions de coulomb	36
A.II.4.c) La rugosité de surface	37
A.II.4.d) Les défauts cristallins non chargés (défauts neutres μ_N)	37
A.II.5) Transport non stationnaire	37
A.II.5.a) Le phénomène de survitesse	38
A.II.5.b) Modèle de transport balistique (ou quasi-balistique)	39
A.II.5.c) Balisticité et dégradation de mobilité	42
A.II.5.d) Unification des mécanismes de transport : dérive diffusion et quasi-balistique	43
Table des figures de la partie A	45
Liste des tableaux de la partie A	47

PARTIE A: GENERALITES

Avant d'entamer la description des résultats, il me semble évident de présenter le principe du transistor métal-oxyde -semi-conducteur dit MOS. C'est ainsi que cette partie dressera les connaissances de base nécessaires pour la compréhension de la physique qui régit le transistor MOS.

Nous rappellerons son principe général qui va de son fonctionnement jusqu'à son architecture de plus en plus complexe.

Nous élucideront le concept de transport électronique qui permet non seulement de comprendre le fonctionnement du transistor, mais aussi d'expliquer les effets du passage d'une architecture à une autre.

Chapitre A.I : RAPPEL SUR LE TRANSISTOR MOS

A.I.1) Principe de fonctionnement

A.I.1.a) Fonctionnement

Le transistor MOS (Metal-Oxide-Semiconductor) comme son nom l'indique, est composé de l'empilement de trois matériaux avec des propriétés différentes. Une première partie est constituée d'un matériau conducteur initialement métallique, mais parfois remplacé par un semiconducteur dopé à dégénérescence. La deuxième partie est un matériau isolant. Ensuite vient s'ajouter un semiconducteur (cf. [figure A.1 & A.2](#)).

Le MOS est un composant microélectronique capable de commuter entre deux états de courant (courant on et courant off). En utilisation logique, on dira qu'il est passant sa résistance est faible (ON) et bloqué (OFF) dans le cas contraire.

Son principe de fonctionnement est régi par un « effet de champ ». Un champ électrique appliqué sur une de ses électrodes qu'est la grille (le matériau conducteur) vient moduler la densité de charge contenue dans le canal (partie du semiconducteur au contact de l'isolant) de conduction lui-même situé entre deux autres électrodes que sont la source et le drain (réservoirs de porteurs). La modulation du canal par la tension appliquée à la grille permet de distinguer les deux états de fonctionnement :

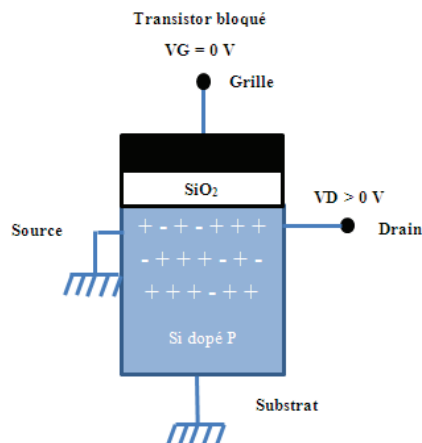


Figure A.1: transistor nMOS à l'état bloqué ($V_g = 0$), valide aussi pour les dispositifs pMOS en inversant les polarisations

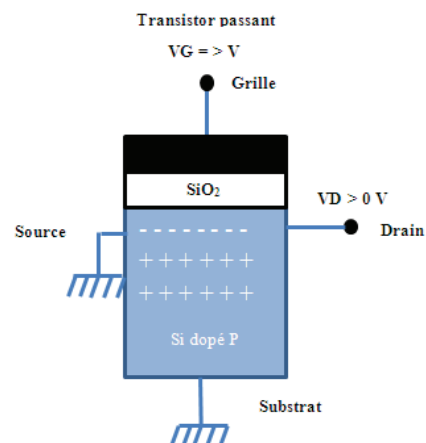


Figure A.2: transistor nMOS à l'état passant ($V_g > 0$), valide aussi pour les dispositifs pMOS en inversant les polarisations

- L'état bloqué : Lorsque le champ de grille est nul, un faible taux de porteurs libres est présent dans le canal et la résistance est forte entre la source et le drain (figure A.1).
- L'état passant : En présence de champ électrique, des porteurs présents dans le canal sont susceptibles de se mouvoir et se déplacent d'un réservoir à l'autre. Si on apporte une différence de potentiel entre la source et le drain, un courant source-drain est alors généré (figure A.2) [Skotnicki03].

Ces deux états de fonctionnement sont possibles grâce à l'effet de champ dans le semiconducteur, qui n'est pas fortement dopé.

A.I.1.b) Le dopage

Le matériau de base principal de la microélectronique est le silicium. Les substrats en sont faits en grande majorité. L'atome de silicium fait partie de la colonne IVA du tableau de Mendeleïev (voir tableau A.1) donc il a 4 électrons sur son niveau de valence. Si on introduit un atome de la colonne VA, ayant donc 5 électrons sur son niveau de valence dans un cristal de silicium cet atome aura tendance à donner au réseau cristallin un électron libre pour se placer dans la bande de conduction du silicium. Donc, si on introduit une dose importante d'atomes de la colonne V dans un cristal de Silicium, on aura un surplus de la densité d'électrons libres d'électrons par rapport à son état initial et l'apparition d'un pseudo niveau d'énergie (E_D) situé juste sous la bande de conduction (E_C). Il devient facile de faire passer un électron d' E_D à E_C (figure A.3a). C'est ce qu'on appelle doper un substrat et dans ce cas on aura un dopage de type N car on aura favorisé les électrons (N=charge négative). De façon symétrique, si on introduit un atome de la colonne IIIA, ayant donc 3 électrons sur son niveau de valence, celui-ci aura tendance à prendre au cristal un électron de la bande de valence du silicium. Cela entraîne l'apparition d'un pseudo niveau d'énergie (E_A) situé juste en dessous de la bande de valence (E_V). Le transfert des électrons d' E_V à E_A est plus facile (figure A.3b). On peut considérer qu'il a donné une charge virtuelle positive au cristal que l'on appelle « trou ». Ceci est un dopage de type P (P=charge positive). Pour des raisons de compatibilité technologique, c'est le Bore qui est utilisé principalement pour doper positivement un substrat Silicium alors que ce sont le Phosphore et l'Arsenic qui sont utilisés habituellement pour le doper négativement.

Dans le cas du nMOS, le semiconducteur est dopé P, les porteurs majoritaires sont donc les trous. Dans le cas de l'application d'une tension de grille négative et d'une tension de drain positive, ces porteurs majoritaires sont attirés en surface et il ne passe aucun courant entre la source et le drain : le transistor est bloqué.

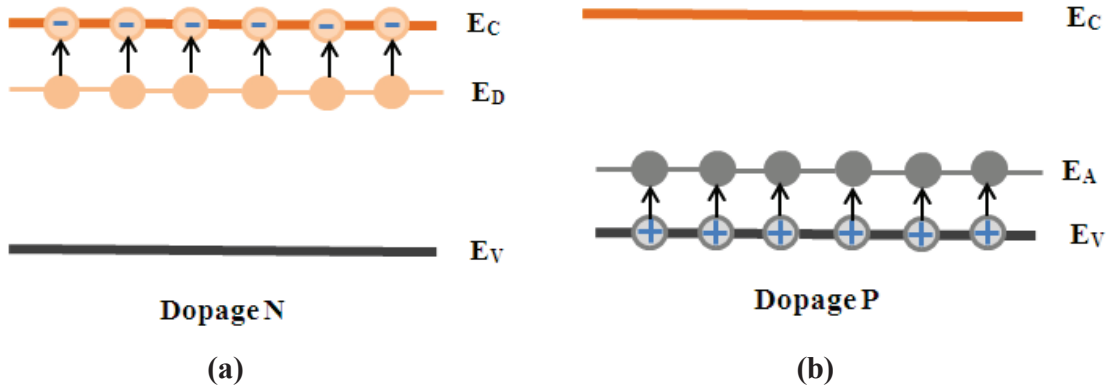


Figure A.3: Schéma de bande d'un semiconducteur dopé N (a) et dopé P (b)

Dans le cas d'une tension positive à la grille, les porteurs minoritaires (ici les électrons) sont attirés en surface (On dit que le canal est inversé). Ces porteurs minoritaires vont pouvoir se mouvoir grâce à la différence de potentiel créée par la tension de drain. Un courant circule dans le canal : le transistor est passant.

Tableau périodique des éléments

IA	Tableau périodique des éléments																VIIIA					
1	H																	He				
2	Li	Be															B	C	N	O	F	Ne
3	Na	Mg											Al	Si	P	S	Cl	Ar				
4	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr				
5	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe				
6	Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn				
7	Fr	Ra	Ac	Rf	Db	Sg	Bh	Hs	Mt	Uu	Uu	Uu	Uu	Uu	Uu	Uu	Uu	Uu				
6	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu								
7	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr								

Tableau A.1: Tableau périodique des éléments: Mendeleïev 1870

A l'état passant du transistor, le canal conduit. Nous allons voir maintenant les différents facteurs qui gouvernent cette conduction.

A.I.2) Architecture de base et régimes de fonctionnement d'un transistor MOS

A.I.2.a) Architecture de base

La [figure A.4](#) présente l'architecture d'un transistor MOS sur silicium massif de type n (le canal en inversion est peuplé d'électrons). L'architecture classique est constituée de quatre terminaux qui permettent d'analyser le comportement électrique du transistor : la grille (V_G), la source (V_S), le drain (V_D) et le contacte du substrat (V_B).

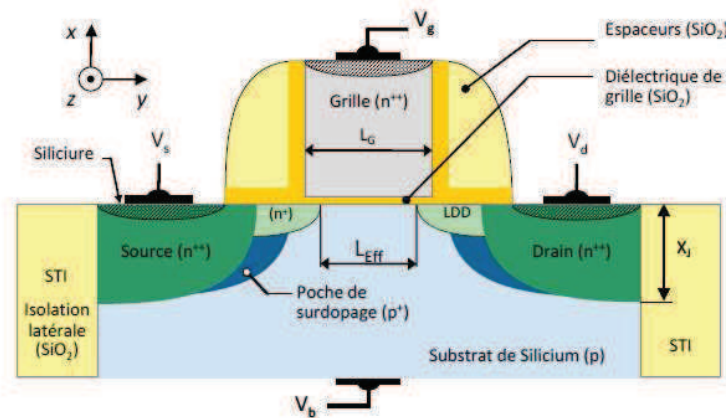


Figure A.4: Architecture d'un transistor *n*MOS sur silicium massif, d'après [Skotnicki03]

La structure du transistor MOS est constituée des éléments suivants :

- Une électrode de grille qui permet de contrôler l'état du canal et le passage du courant,
- un canal qui permet le passage des porteurs entre la source et le drain,
- un diélectrique permettant d'isoler la grille du canal pour permettre un couplage grille-canal capacitif,
- Les réservoirs de porteurs source et drain qui fournissent les porteurs minoritaires et permettent la formation d'une couche d'inversion.

Le principe du transistor MOS consiste à moduler la densité de porteurs du canal en jouant sur la polarisation de grille (V_{GS}). Celle-ci permet, par le biais du champ électrique ainsi créé, d'attirer une quantité de porteurs depuis les réservoirs à la surface du semiconducteur. Une couche

conductrice se forme alors à ce niveau et permet la libre circulation d'un courant (I_D), sous réserve de l'application d'un champ électrique longitudinal (i.e. d'une tension V_{DS}). Cette couche conductrice est appelée "couche d'inversion" ou parfois "canal d'inversion". Cf. [Skotnicki03]. Selon les tensions V_G et V_D appliqué, on peut distinguer le fonctionnement en trois principaux régimes.

A.I.2.b) Les régimes de fonctionnement

La tension V_{GS} (égale à V_G quand $V_S = 0V$) impose le régime de fonctionnement via la modulation d'une barrière de potentiel V_{bi} . Cette barrière de potentiel (built-in potential) qui vient du type de dopage opposé entre la jonction source-substrat d'une part et drain-substrat d'autre part, empêche la libre circulation des porteurs [Sze81]. On définit aussi une tension de bande plate V_{FB} (lorsque concentrations de porteur majoritaires et minoritaires sont constantes dans tout le substrat), qui est la tension qu'il est nécessaire d'appliquer sur la grille pour obtenir un régime de bandes plates dans la structure :

$$V_{FB} = (\Phi_m - \Phi_s) - \frac{Q_0}{C_{ox}} \quad (A.1).$$

Φ_m et Φ_s sont les travaux de sortie du métal (grille) et du semiconducteur (substrat) respectivement. Q_0 est la charge d'interface et C_{ox} la capacité de l'isolant. Cf. [Sze81].

On définit ainsi quatre régimes de fonctionnement selon la tension V_G appliquée :

- **Le régime d'accumulation** : La tension $V_G < V_{FB}$, les porteurs majoritaires du substrat sont attirés en surface. Le transistor est bloqué et il ne passe aucun courant significatif (diode en inverse).
- **Le régime de bandes plates** : $V_G = V_{FB}$, les concentrations de porteur majoritaires et minoritaires sont constantes dans tout le substrat.
- **Le régime de désertion et d'inversion** : $V_G > V_{FB}$, les porteurs majoritaires sont repoussés de la surface du substrat et laissent la place aux porteurs minoritaires. Lorsque V_G devient supérieure à une tension de seuil V_{TH} , la concentration des minoritaires

devient supérieure à celle des majoritaires à la surface du substrat, c'est l'inversion. La tension de seuil est donnée par :

$$V_{TH} = V_{FB} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}} \quad (\text{A.2}) \text{ où } \Phi_F \text{ est le niveau de fermi et } Q_{dep} \text{ est la charge de déplétion.}$$

Si les régimes de fonctionnement sont définis selon V_G , les régimes de conduction sont définis selon V_D .

A.I.3) Les régimes de conduction

Le courant délivré par le transistor est gouverné par trois facteurs :

- La charge d'inversion dans le canal qui dépend de la tension de grille et qui est donnée en régime d'inversion forte en un point x du canal par :

$$Q_{inv} = C_{ox} \cdot (V_G - V_{TH} - V(x)) \quad (\text{A.3})$$

C_{ox} est la capacité de l'oxyde, V_G est la tension de grille, V_{TH} est la tension de grille minimale pour créer l'inversion dans le canal et $V(x)$ est le potentiel au point x .

- La tension de drain V_D qui permet le déplacement des porteurs de la source au drain.
- La mobilité des porteurs μ_{eff} qui quantifie la vitesse acquise par les porteurs sous l'action d'un champ électrique (cette notion sera détaillée plus tard).

Deux régimes de conduction distincts peuvent avoir lieu en fonction des valeurs de V_{DS} (égale à V_D quand $V_S = 0$):

- Le régime linéaire qui est caractérisé par une dépendance linéaire du courant en fonction de V_{DS} (faible V_{DS}). Dans ce régime, la vitesse des porteurs est aussi linéairement dépendante du champ électrique engendré par la tension de drain. Le facteur de linéarité est défini comme étant la mobilité (μ).
- Quant au régime de saturation (fort V_{DS}), il se traduit par la perte de cette linéarité du courant avec V_{DS} : le courant varie peu avec V_{DS} . En même temps, la vitesse des porteurs perd sa linéarité avec le champ électrique dû à V_{DS} . Le terme mobilité perd tout son sens et la charge d'inversion diminue au drain.

Dans ce régime, le canal est dit pincé et V_{DS} ne contrôle plus ce qui s'y passe. La tension de pincement est définie par : $V_P = V_{GS} - V_{TH}$ (A.4).

Entre les deux régimes il existe une zone de raccordement dans laquelle le courant délivré dépend à la fois de V_{GS} et de V_{DS} (figure A.5).

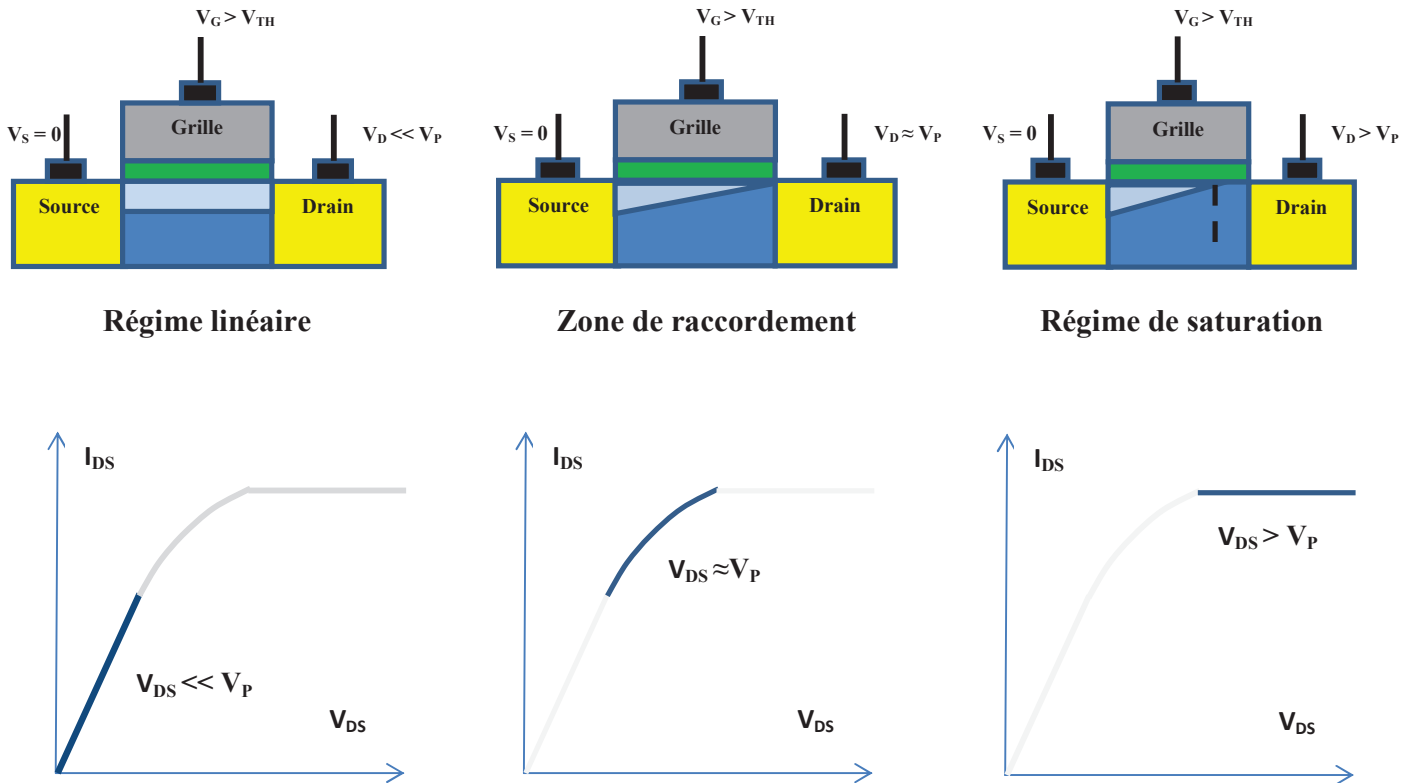


Figure A.5 : Illustration des trois régimes de conduction qui caractérisent le fonctionnement d'un transistor nMOS [Skotnicki03]

Comme indiqué en introduction générale, la taille des transistors n'a pas cessé de réduire au fil du temps (loi de Moore). Cette réduction des dimensions permet de gagner en vitesse et en fréquence de fonctionnement. Ce gain en performance s'accompagne d'effets indésirables que sont les effets de canaux courts.

A.I.4) Les effets de canal court

A.I.4.a) Effet de la résistance source-drain

La résistance totale dans le canal (R_{TOT}) du transistor est donnée en régime linéaire par la résistance du canal (R_{CH}) et celle de la source (R_S) et drain (R_D) (figure A.6) ($R_{SD}=R_S+R_D$) :

$$R_{TOT} = R_{CH} + R_{SD} \quad (A.5).$$

Dans les transistors longs, $R_{CH} \gg R_{SD}$, donc la résistance totale s'apparente à la résistance du canal. Au fur et à mesure que les dimensions diminuent (notamment la longueur du canal), R_{SD} ($R_S + R_D$) n'est plus négligeable devant R_{CH} [Hu87- Chang10-Fleury09c]. Dans ces conditions, ces résistances série engendrent une diminution du courant. En régime ohmique, le courant de drain corrigé des R_{SD} est donc donné par :

$$I_{D,corr} = \frac{I_D}{1 - \frac{I_D \cdot R_{SD}}{V_D}} \quad (A.6).$$

En régime saturé, cette correction se fait souvent sur la tension de grille pour compenser la chute de potentiel à la source. En effet on redéfinit V_G comme suit :

$$V_{G,corr} = V_G - R_S \cdot I_D \quad (A.7a).$$

Similairement, on compense la chute de potentiel au drain par :

$$V_{DS,corr} = V_{DS} - R_{SD} \cdot I_D \quad (A.7b)$$

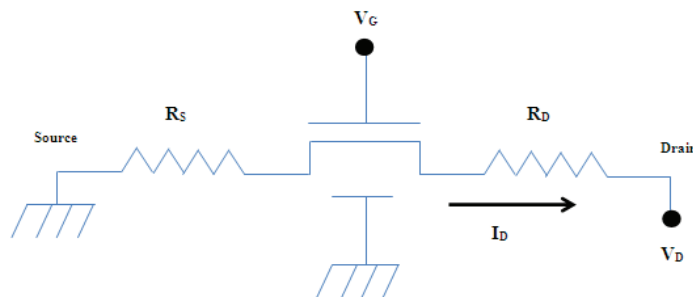


Figure A.6: Schéma électrique équivalent du MOSFET - influence des résistances séries de source et de drain

A.I.4.b) Influence des zones de déplétion source-drain

Le contact des sources et drain avec le canal crée des zones de déplétions comme dans tout jonction PN ou NP (le dopage des sources et drain sont opposés à celui du canal). Tant qu'on avait des grandes longueurs de canal, ces zones de déplétion avaient peu d'effet. Avec la diminution des dimensions, ces zones de déplétion ne sont plus négligeables et doivent être prises en compte.

➤ Partage de charge

Une conséquence majeure du rapprochement des jonctions source et drain est la perte par la grille du contrôle d'une partie des charges situées à sa verticale. C'est ce que l'on nomme « partage de charge ou charge sharing », ses répercussions sont importantes sur certains paramètres électriques, notamment la tension de seuil V_{TH} .

Lorsque les extensions des zones de désertion de jonction sont de l'ordre de grandeur de la longueur de grille, la part de la surface du semiconducteur contrôlée par la grille diminue. Avec la réduction de la longueur de canal, la charge Q du canal n'est donc plus efficacement contrôlée par la tension de grille et la barrière de potentiel entre la source et le drain s'abaisse (SCE) (figure A.7).

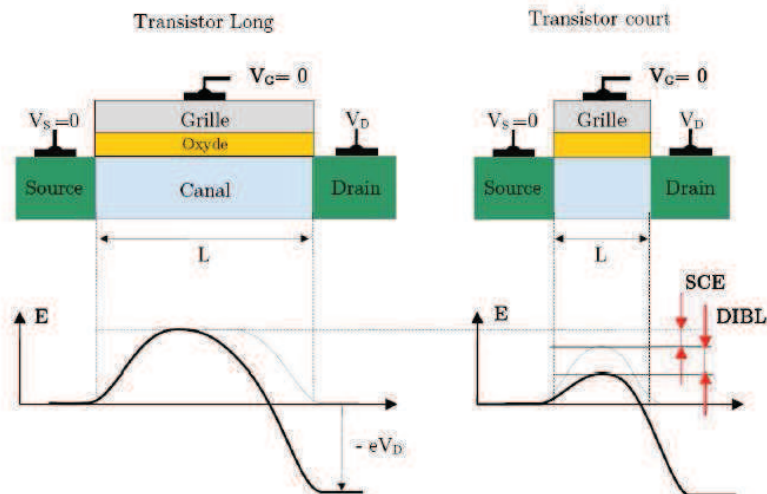


Figure A.7: Illustration de l'abaissement de la barrière de potentiel source-drain induit par les effets de canaux courts (CSE et DIBL)

A.I.4.c) Le DIBL

Sur les transistors courts, l'application d'une forte tension V_D sur le drain, provoque une chute supplémentaire de la barrière de potentiel source-drain. Le contrôle de la charge devient dépendant de la grille et du **drain** : c'est le DIBL (Drain Induced barrier lowering) [Fikri94-Ghibaudo97-Deen92-Ayhan00]. Cet effet plus présent au fur et à mesure que la longueur de grille diminue, est illustré sur les figures A.7 & A.8.

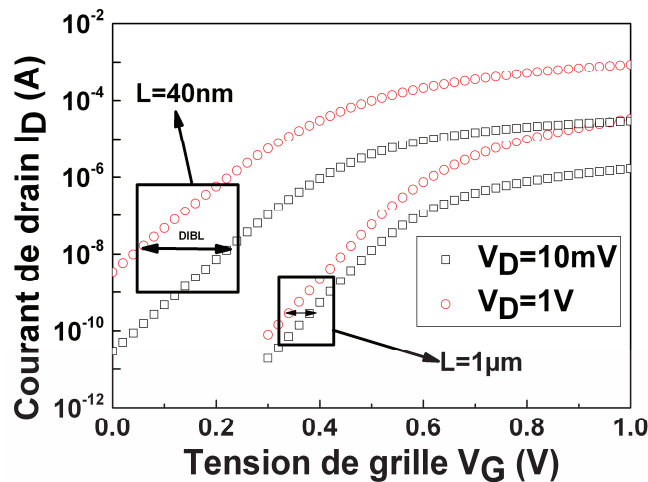


Figure A.8: Courant de drain en fonction de la tension de grille pour un transistor long ($L=1\mu\text{m}$) et un transistor ultracourt ($L=40\text{nm}$) à faible (10mV) et forte (1V) polarisation de drain.

Pendant longtemps, la réduction des dimensions (autrement appelé le downscaling) s'est poursuivie sans se soucier des effets néfastes qui y sont liés. Cependant, les effets de canaux courts sont devenus tellement importants et prédominants que l'on ne peut plus les négliger. Plusieurs solutions technologiques ont été proposées pour parer ces effets. Ces solutions englobent le dopage du canal mais aussi une mutation de l'architecture du transistor en lui-même.

Pour éviter les effets néfastes du downscaling, une alternative est de proposer des solutions technologiques qui augmentent les performances sans avoir besoin de réduire les dimensions (on peut citer la contrainte mécanique, l'avènement des matériaux à haute permittivité, ...etc.). Une autre solution consiste à changer l'architecture pour mieux supporter les effets canaux courts.

A.I.5) Evolution de la technologie MOS : Du transistor sur substrat massif au FinFET en passant par le FDSOI

La technologie du transistor a beaucoup évolué afin de contrecarrer les effets de canaux courts tout en augmentant la performance. C'est ce que l'on peut voir sur la roadmap ITRS 2011 à la **figure A.9**. Comme on peut le voir sur cette figure, les solutions technologiques les plus utilisées ces dernières années peuvent se diviser en trois familles :

- Des alternatives sont proposées pour l'empilement métal-oxyde, cet empilement est communément appelé empilement de grille (ou gate-stack en anglais).

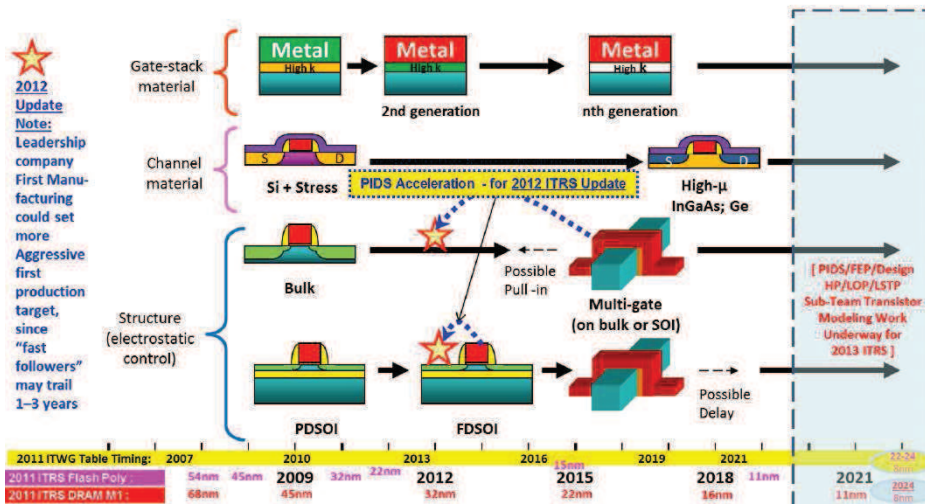


Figure A.9: L'évolution des technologies selon la roadmap ITRS 2011, actualisé pour 2012

- Une deuxième solution consiste à travailler sur le matériau de canal en lui-même.
- Une troisième solution est de changer l'architecture pour un meilleur contrôle de l'électrostatique.

A.I.5.a) Les solutions apportées sur le matériau canal

Le but ultime des sociétés de microélectronique est d'augmenter la performance des transistors en augmentant le courant débité. Le temps de commutation τ des circuits (constitués

de portes logiques à base de transistors) est inversement proportionnel au courant débité. Le circuit sera d'autant plus rapide que le courant sera important :

$$\tau = \frac{C_{eff} \cdot V_{DD}}{I_{eff}} \quad (\text{A.8})$$

- ✓ C_{eff} est la capacité effective, caractéristique du couplage entre la grille et le canal.
- ✓ I_{eff} est le courant moyen en régime de fonctionnement et V_{DD} est la tension d'alimentation du circuit.

Une solution pour augmenter les performances est de jouer sur le matériau de canal en lui-même pour augmenter la mobilité des porteurs et donc le courant sans autant réduire la longueur (on évite les effets de canaux courts). Une étape est ajoutée dans le procédé de fabrication du dispositif afin de produire une contrainte dans le canal. Cette contrainte permet de modifier le peuplement des bandes dans le matériau. La mobilité μ des porteurs est ainsi améliorée, et conséquemment le courant débité aussi ($I_{eff} \sim \mu$). Il existe plusieurs façons de produire cette contrainte.

- Une méthode consiste à faire les sources/drain avec un matériau dont le paramètre de maille est différent de celui du matériau de canal. C'est donc ce désaccord de paramètre de maille qui va induire une compression ou une tension du matériau de canal. Grâce à sa différence de paramètre de maille avec le silicium, le silicium-germanium (SiGe) peut être utilisé pour induire une contrainte compressive (favorable à la mobilité des électrons) ($a=5.43\text{Å}$ pour le Silicium et 5.64Å pour le germanium) [Ghani03-Thompson04], cf. figure A.10. Le carbure de silicium ($\text{Si}_{1-x}\text{C}_x$) peut aussi être utilisé comme matériau source/drain pour induire une contrainte extensive dans le canal (favorable à la mobilité des trous). Là aussi, c'est le désaccord de paramètre de maille qui rentre en jeu comme dans [Ang04]. Il s'agit d'une contrainte dite uniaxiale (dans la direction du transport)

- Une autre méthode pour apporter une contrainte dans le canal consiste à insérer dans le dispositif une couche de matériau de paramètre de maille différent à celui du silicium, voir sur la [figure A.11 \[Welsey92\]](#). La recette reste la même que dans le cas des source/drain SiGe, c'est-à-dire le désaccord de maille entre le silicium et le SiGe par exemple. Dans ce cas, on parle de contrainte biaxiale.

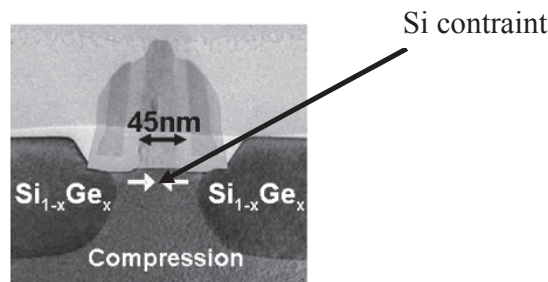


Figure A.10: Illustration de la contrainte compressive du canal par les sources/drain via un mismatch de paramètre de maille entre le matériau source/drain et le matériau canal [\[Thompson04\]](#)

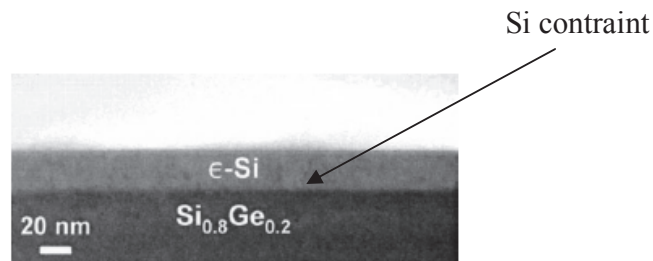


Figure A.11: Illustration d'un canal silicium contraint dans une heterostructure [\[Lee05\]](#)

- La troisième méthode de contrainte que nous évoquerons est celle appelé CESL (contact-etching stop-layer). Elle consiste à déposer une couche de diélectrique (qui peut être du nitrure de silicium par exemple) au-dessus des zones de grille et des source/drain comme illustré sur la [figure A.12 \[Wang10\]](#). L'effet de cette contrainte, de type uniaxiale, sera d'autant plus important que la longueur du canal diminue.

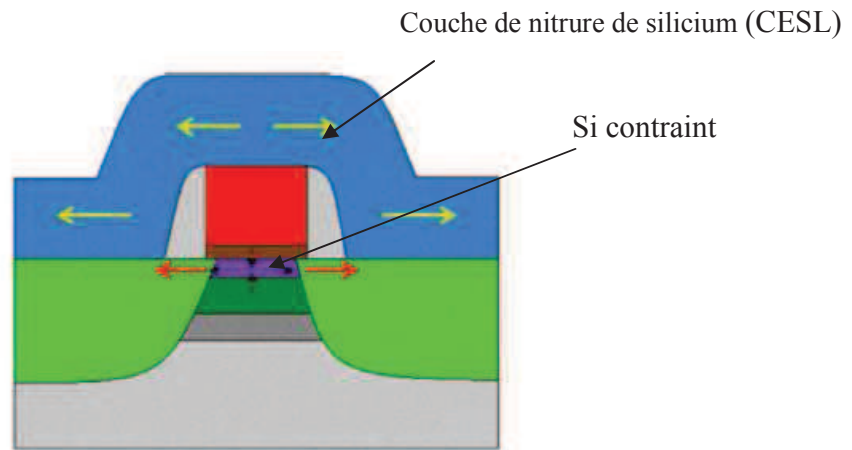


Figure A.12: Illustration d'une contrainte extensive via CESL sur le canal [Wang10]

A.I.5.b) L'empilement de grille

L'empilement de grille est historiquement constitué de poly-silicium surdopé comme matériau de grille et de dioxyde de silicium comme oxyde de grille. Il a été difficile de maintenir cette architecture pour contrecarrer les effets de la miniaturisation. Aujourd'hui l'empilement de grille est constitué de multiples couches constituées de métaux et d'oxydes.

L'oxyde de grille initialement utilisé pour la fabrication des transistors MOS est le dioxyde de silicium SiO_2 . L'utilisation du SiO_2 atteint sa limite avec la réduction continue des dimensions. Les épaisseurs d'oxyde atteignent les alentours de 1.5 nm et cela engendre des problèmes tels que l'effet tunnel des électrons à travers ce même oxyde. Il s'en suit donc une augmentation inévitable et non souhaitable du courant de fuite (I_{off}) dans les dispositifs, et donc de la consommation électrique à l'état off (lorsque le dispositif n'est pas en phase d'utilisation) :

$$P_{\text{off}} = V_{DD} \cdot I_{\text{off}} \quad (\text{A.9})$$

P_{off} est la puissance consommée passivement et I_{off} est le courant débité par le transistor au repos (par l'effet tunnel évoqué plus haut par exemple).

C'est ainsi qu'une solution technologique est proposée pour parer à ces difficultés. Cela consiste à utiliser un autre type d'oxyde de grille avec une constante diélectrique plus grande que celle du SiO_2 . Ces oxydes appelés high-k permettent l'utilisation d'épaisseurs plus grandes pour un

même champ électrique. On évite ainsi les fuites de grille [Taur97-Liu01]. Nous reviendrons plus largement sur les effets de ce type d'oxyde de grille dans la partie « influence du procédé de fabrication sur les performances du transistor ».

La mauvaise qualité de l'interface high-k/poly-silicium [Misra05] a été un facteur déterminant pour réfléchir à un autre type de grille [Hobbs01]. Le remplacement du poly-silicium par un métal a permis de réduire ces effets néfastes dus à la formation de dipôles à l'interface high-k/poly-silicium. Il faut signaler que le métal le plus couramment utilisé aujourd'hui est le nitrure de titane (TiN) [Maiti98-Liu09-Andrieu06]. De plus, l'utilisation d'un métal de grille, dont la densité de porteur est plus grande que celle du polysilicium dopé, permet de réduire très fortement le phénomène de désertion des porteurs à l'interface grille/diélectrique rencontré dans le polysilicium, et donc d'augmenter la capacité de couplage.

Une solution transversale pour augmenter les performances en s'affranchissant des effets canaux court consiste à changer le matériau de canal, qui jusqu'ici est constitué de silicium.

A.I.5.c) L'avènement d'autres matériaux pour remplacer le canal Si

La réduction sans limite des dimensions fait que les solutions que nous venons d'évoquer ne suffisent plus pour faire face aux difficultés liés à la miniaturisation.

L'idée est d'utiliser d'autres matériaux de mobilité meilleure que celle du silicium pour faire le canal. Parmi les candidats pour remplacer le canal silicium, on trouve le silicium-germanium, le germanium et les matériaux III/V comme l'indium-arséniure de gallium (InGaAs).

- **Canal à base de silicium-germanium (SiGe) : Couche de SiGe épitaxiée**

La mobilité des trous dans le silicium est faible comparée à celle de ses électrons. Cela pose des problèmes lorsqu'il faut associer des transistors de types nMOSFETs (conduction par les électrons) et de types pMOSFETs (conduction par les trous) pour des applications logiques. D'où la nécessité de trouver un moyen d'augmenter la performance des pMOSFETs. Des transistors de ce type se sont donc développés avec une couche de SiGe épitaxié sur du substrat silicium (figure A.13). L'idée est de former un canal SiGe dans laquelle la mobilité des trous est meilleure que celle dans le silicium [Alatise09-Sarah04-Ke91].

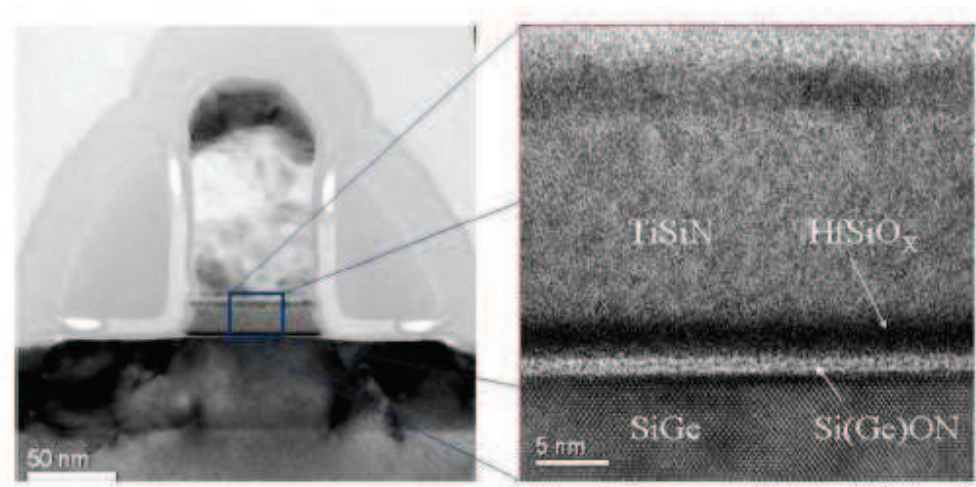


Figure A.13: Coupe TEM de pMOSFETs SiGe contraint avec le stack de grille correspondant [Alatise09]

- Technologie germanium (Ge) bulk : substrat Ge

Pour encore augmenter les performances des pMOSFETs, des technologies germanium sont développés en laboratoire. Notons que la piste menant au germanium a déjà constitué un thème de recherche mais les problèmes liés à sa petite bande interdite (de l'ordre de 0.66 eV) et à sa grande constante diélectrique avaient poussé les chercheurs à son abandon. A partir des années 2000, la démonstration d'oxyde stable (HfO_2 , ZrO_2 ,...etc.) et compatible sur germanium a à nouveau ressuscité les recherches [Chui04]. Si le pMOSFET germanium a suscité de l'engouement, le nMOSFETs présente de pauvres résultats en termes de performance. Des transistors à base de germanium fonctionnels (cf. figure A.14) ont donc vu le jour avec des performances acceptables [Hellings08-Mitard08].

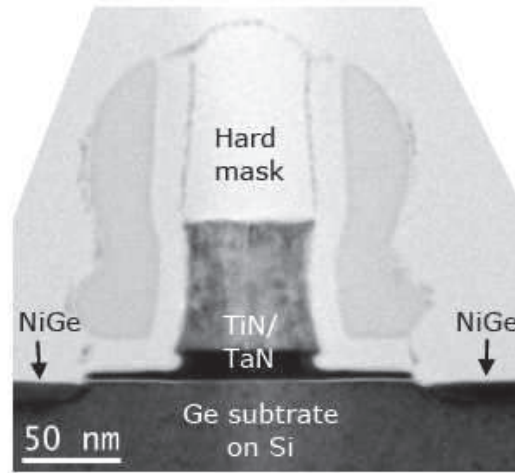


Figure A.14: Image TEM d'un transistor Ge-pMOSFETs [Mitard08]

Cependant, beaucoup de problèmes qui accompagnent le dimensionnement persistent et leur étude est primordiale pour élaborer des composants à base de Ge fiables. Des problèmes comme l'effet tunnel bande à bande font augmenter le courant de fuite, et donc la consommation passive. Nous reviendront sur ces points dans la partie consacrée à l'étude des transistors germanium.

- [La technologie à base de matériaux III/V](#)

Comme nous l'avons évoqué brièvement, l'expérience du Ge-nMOSFETs n'a pas été très concluante. La contrainte sur du Si-nMOSFET ne permet plus d'augmenter la performance. Les matériaux III/V se sont donc révélés comme de potentiels candidats pour augmenter les performances du nMOSFET. Les III/V ont une grande mobilité et saturent moins vite en effet de la contrainte mécanique comparés au silicium. Ces matériaux sont prometteurs pour le futur notamment pour l'intégration III/V-nMOSFET et Ge-pMOSFET sur isolant (BOX) comme illustré sur la [figure A.15](#) [Takagi07].

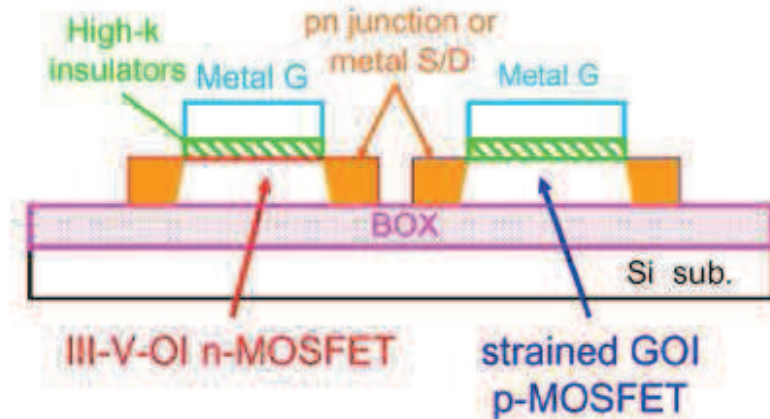


Figure A.15: Figure illustrant la combinaison de III/V-nMOSFET et GE-pMOSFET sur isolant [Takagi07]

Beaucoup de recherches ont eu lieu pour réfléchir à la manière d'augmenter le courant avec moins de dimensionnement. En parallèle, d'autres études se sont focalisées sur la manière de changer l'architecture du transistor pour continuer la réduction des dimensions et en même temps contrôler l'électrostatique. C'est ainsi que des technologies FDSOI (fully depleted silicon on insulator) et finFET ont émergé et attirent aujourd'hui l'attention du monde de la nanoélectronique.

A.I.5.d) Structures à électrostatique améliorée

Isoler le canal du reste du dispositif permet un meilleur contrôle électrostatique, et donc de réduire les effets parasites et de canaux courts.

- La technologie FDSOI

La figure A.16 montre l'architecture du MOS sur SOI [Beranger07-Beranger12]. On peut dire que la présence d'une couche enterrée d'oxyde de silicium SiO_2 (ou HfSiON) est la rupture technologique entre la technologie conventionnelle bulk et le SOI.

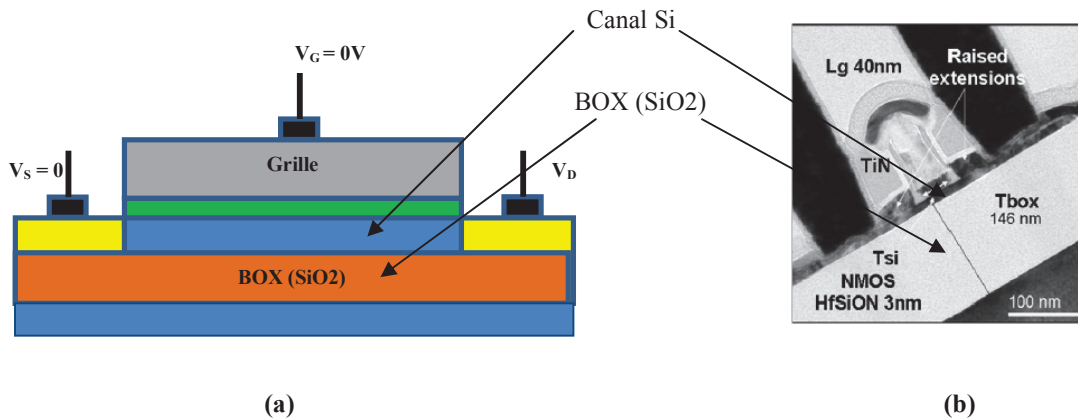


Figure A.16: l'architecture d'un MOS type SOI (a) et coupe TEM d'un transistor MOS SOI (b) [Beranger07]

Le terme « fully depleted » (totalement déserté) vient du fait que le canal se forme dans toute la couche de silicium au-dessus de l'oxyde enterré (couche d'isolant, constituée de SiO_2 ou HfSiON). Cette couche de silicium isolée et non dopée, permet un meilleur contrôle de la tension de seuil V_{TH} . Le FDSOI offre ainsi un avantage colossal par rapport au bulk. La communauté microélectronique peut ainsi compter sur cette technologie pour les nœuds futurs.

Il faut mentionner que des études ont aussi été faites sur du « partially depleted » (PDSOI), c'est-à-dire dans le cas où la couche de Si est assez épaisse pour n'être que partiellement désertée

Un autre concurrent sérieux au FDSOI pour contrecarrer les effets de canaux court est le finFET (autrement appelé trigate).

- La technologie finFET et trigate

L'idée générale reste le contrôle des charges qui se créent dans le canal. Comme nous venons de le voir avec le FDSOI, l'isolement du canal permet d'atteindre cet objectif. Une autre manière d'augmenter cette emprise sur les charges, est d'augmenter la surface de grille autour du canal. On peut voir sur la [figure A.17](#), que le contrôle des charges par la grille dans le cas du trigate [Auth12] se fait désormais des trois côtés. Le contrôle se fait de deux côtés dans le cas du finFET.

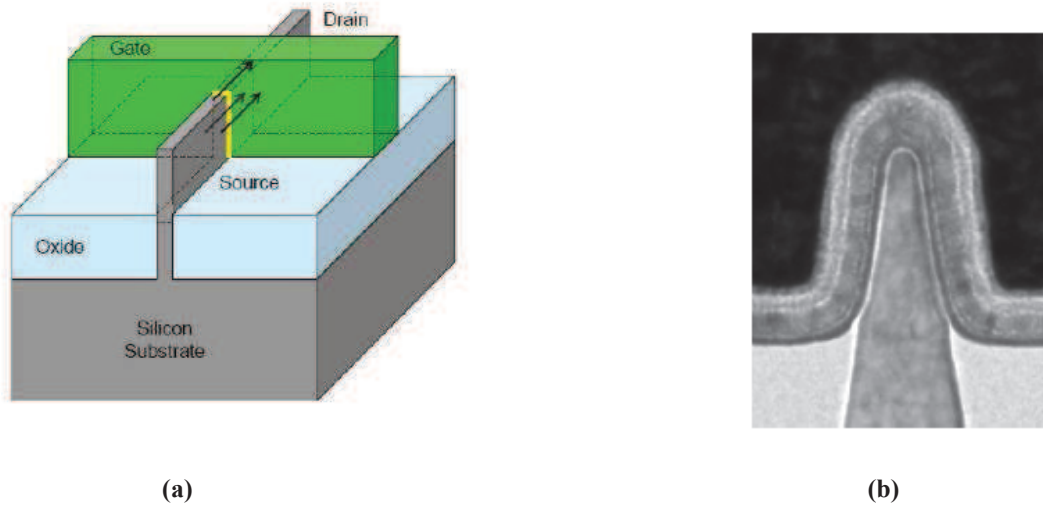


Figure A.17: Schéma simplifié illustrant une architecture trigate (a), coupe TEM d'un transistor trigate (b) [Auth12]

C'est ainsi qu'Intel a fait a annoncé son passage au tri-gate pour le nœud 22nm et au-delà (cf. <http://www.intel.com>).

Pour encore d'avantage de contrôle électrostatique, le finFET sur substrat SOI est envisagé. En résumé, il s'agit de coupler deux manières de contrôler les charges : l'isolement du canal et l'augmentation de la surface de grille sur celui-ci. Sur la [figure A.18](#), un benchmark entre une technologie finFET sur substrat massif (BFF) et finFET sur SOI (SOIFF) est présenté [Chiarella10]. Ces technologies sont considérées par certains comme l'avenir du transistor MOS grâce à leur faculté de dimensionnement.

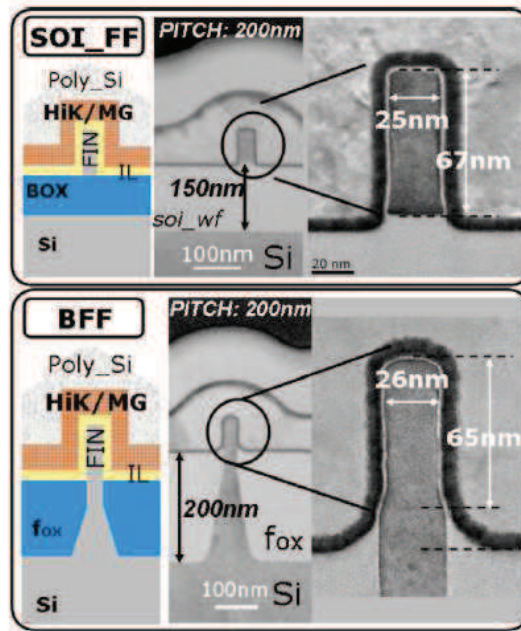


Figure A.18: Benchmark entre du SOIFF et du BFF [Chiarella10]

A.I.5.e) Bilan et conclusion

Nous avons vu tout au long du chapitre I de cette partie que l'architecture MOS a beaucoup évolué. Un besoin de compréhension des phénomènes physiques s'impose au fur et à mesure que l'architecture se complexifie. Le transport électronique qui décrit la manière dont les porteurs (électrons ou trous) se déplacent dans le canal sous l'effet d'un champ électrique (appliqué au drain) est un outil fiable pour comprendre le fonctionnement de ces composants de plus en plus complexes. Le chapitre II de cette partie sera donc consacré à la description du transport électronique.

Chapitre A.II : LE TRANSPORT ELECTRONIQUE

Le transport électronique est le phénomène physique qui définit la manière dont des porteurs se déplacent dans un matériau sous l'effet d'un champ électrique. Dans le cas du transistor MOS, le transport est l'étude du mouvement des électrons ou des trous de la charge d'inversion dans le canal sous l'effet d'un champ électrique longitudinal. Ce champ est obtenu grâce l'application d'une différence de potentiel entre le drain et la source.

A.II.1) Transport des porteurs dans le canal

Pour décrire historiquement le transport des porteurs, nous utiliseront le modèle de Drude et celui dit « dérive diffusion ».

A.II.1.a) Le transport stationnaire : Du model de Drude au transport dérive diffusion

Le concept de collision a été introduit dans les années 1900 par Paul Karl Ludwig Drude (d'où l'appellation « modèle de Drude ») pour rendre compte de la loi d'Ohm locale dans les métaux.

On peut ainsi écrire : $\vec{j} = \sigma \cdot \vec{\mathcal{E}}$ (A.10) avec $\sigma = n \cdot q \cdot \mu$ et $\mu = \frac{q \cdot \tau_{coll}}{m^*}$

Dans cette expression, j , n , q , m^* , μ et τ_{coll} représentent respectivement la densité de courant, la densité d'électrons, la charge de l'électron, sa masse, la mobilité et la durée moyenne entre deux collisions successives (le temps de relaxation). Le paramètre σ est la conductivité du matériau, il permet d'exprimer la densité de courant délivré en fonction du champ électrique latéral.

La prise en compte des collisions dans le traitement du transport revient à considérer que les trajectoires des porteurs en présence d'un champ électrique ne sont pas monotones mais interrompues par des centres de diffusion qui modifient aléatoirement sa direction de déplacement et/ou son énergie (cf. figure A.19).

En l'absence de champ électrique, l'énergie thermique est la seule source d'énergie qui est à l'origine du mouvement des porteurs. La vitesse de dérive s'annule donc et la directionnalité perd tout son sens. L'application d'un champ électrique ajoute une composante de vitesse non nulle au groupe de porteur et la vitesse de dérive n'est pas nulle.

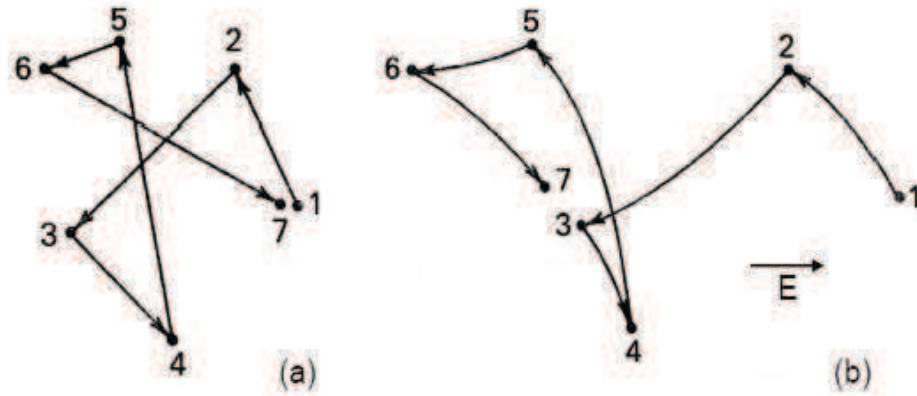


Figure A.19: Illustration de la trajectoire des électrons subissant des collisions dans un cristal. a) sans champ électrique b) sous l'application d'un champ électrique [Grove67]

On parle de transport stationnaire lorsque le système est en équilibre et peut être régi par des équations indépendantes du temps. Le transport est dit stationnaire lorsque le temps de transit des porteurs dans le canal est très supérieur au temps moyen de relaxation en énergie. Dans ce cas, les mécanismes de relaxation en énergie sont fréquents et les porteurs sont dits « thermalisés » avec le réseau : c'est l'équilibre.

Le modèle de transport dérive diffusion prend en compte les équations de Boltzmann (appelées aussi équation de transport dérive diffusion) qui sont ainsi basées sur un calcul à l'équilibre et en considérant que les porteurs subissent un grand nombre de collisions avec les centres de diffusion chargés dans le canal avant de rejoindre le drain.

L'équation de base du modèle de dérive diffusion est la suivante :

$$j = q.(n_{inv} \cdot \mu \cdot \mathcal{E} + D \cdot \nabla n_{inv}) \quad (\text{A.11})$$

Avec n_{inv} la densité volumique des porteurs libres et D est le coefficient de diffusion relié au terme de mobilité μ avec la relation d'Einstein. $D = \frac{\mu \cdot k_B \cdot T}{q}$. Lorsque le transistor est passant, on néglige généralement le terme de diffusion devant celui de dérive.

A.II.2) Notion de mobilité et son utilité pour étudier la performance

La mobilité μ traduit l'aptitude d'un porteur à se mouvoir sous l'effet d'un champ électrique. Elle définit le transport des porteurs dans le canal et influence directement le courant de drain. La mobilité dépend à la fois de la masse effective (m^*) qui est inhérente au matériau du canal et du libre parcours moyen τ_{coll} qui correspond au temps écoulé entre deux collisions.

$$\mu = \frac{q \cdot \tau_{coll}}{m^*} \quad (\text{A.12})$$

Dans le cas d'un transport stationnaire dérive diffusion, le courant I_{dd} est strictement lié à la mobilité (μ_{dd}) et les équations varient selon le régime de fonctionnement du transistor (linéaire et saturation).

$$I_{dd}^{lim} = W \cdot \frac{\mu_{dd} \cdot V_{DS}}{L} \cdot Q_s \cdot \left(1 - \frac{V_{DS}}{2V_{gt}}\right) \approx W \cdot \frac{\mu_{dd} \cdot V_{DS}}{L} \cdot Q_s \quad (\text{A.13}) \text{ si } V_{DS} \ll V_{gt} = V_G - V_{TH} \text{ (linéaire)}$$

$$I_{dd}^{sat} = W \cdot \frac{\mu_{dd} \cdot V_{gt}}{L} \cdot \frac{Q_s}{2} \quad (\text{A.14}) \text{ si } V_{DS} \gg V_{gt} = V_G - V_{TH} \text{ (saturation)}$$

Q_s représente la charge d'inversion et L la longueur du canal.

Ces équations montrent à quel point il est important d'étudier le transport (la mobilité ou la vitesse) pour comprendre les variations de courants et donc de performance.

Les mécanismes de collision qui gouvernent la vitesse de dérive font que les porteurs peuvent avoir des vitesses différentes. On peut donc introduire la notion de mobilité effective μ_{eff} pour rendre compte de la vitesse moyenne des porteurs vue par la couche d'inversion :

$$v_d = \mu_{eff} \cdot \mathcal{E}_{par} \quad (\text{A.15}) \text{ avec } v_d \text{ la vitesse moyenne de dérive et } \mathcal{E}_{par} \text{ le champ parallèle.}$$

A.II.3) La notion de champ effectif et corrélation avec la mobilité effective

A.II.3.a) La notion de champ effectif

Le champ transverse vu par la charge d'inversion (différent du champ à l'interface) varie en fonction de la profondeur x de la couche d'inversion. La mobilité dépendant du champ effectif va donc varier en fonction de la profondeur x de la couche d'inversion.

D'après le théorème de Gauss, le champ électrique $\xi(x)$ en tout point x de la couche d'inversion

$$s'exprime par : \xi(x) = \frac{Q_{dep}}{\epsilon_{si}} + \frac{q}{\epsilon_{si}} \cdot \int_x^{x_{inv}} n_{inv}(u) du \quad (A.16)$$

Q_{dep} représente la densité surfacique de la charge de désertion, n_{inv} la densité de charge d'inversion et ϵ_{si} la permittivité diélectrique du silicium.

Le champ effectif est défini comme étant la moyenne du champ électrique transverse prise sur la distribution de la couche d'inversion [Sabnis79] :

$$\xi_{eff} = \frac{Q_{dep} + \eta \cdot Q_{inv}}{\epsilon_{sc}} \quad (A.17) \quad \text{Où } Q_{inv} \text{ est la charge d'inversion, } \eta \text{ le barycentre électrostatique}$$

propre à la couche d'inversion et ϵ_{sc} la permittivité diélectrique du semiconducteur.

La valeur de η est historiquement et empiriquement donnée égale à 0.5 pour les électrons et 0.33 pour les trous [Takagi94]. Ces valeurs peuvent cependant varier et doivent donc être déterminées pour chaque technologie et selon les conditions d'études.

A.II.3.b) Corrélation mobilité effective versus champ effectif

La corrélation mobilité effective en fonction du champ effectif est d'une importance capitale lorsqu'il s'agit de comprendre le transport dans un transistor. Cette corrélation a permis de dégager une tendance dite « universelle » de la mobilité à l'interface Si/SiO₂ [Takagi94].

Lorsque le paramètre η est bien choisi, la mobilité à champ fort suit une loi universelle en fonction de ξ_{eff} [Takagi94] qui ne dépend pas du dopage (ou de V_{BS}). La déviation à la tendance universelle observée à faible champ sur la figure A.20 provient de la réduction de mobilité causée par des collisions des porteurs avec les impuretés chargées du canal (le dopage) qui se voit principalement à faible charge d'inversion. La réduction à fort champ est attribuée à la

rugosité de surface, alors que la réduction de mobilité à champ moyen est attribuée aux interactions avec les vibrations du réseau (phonons).

Les principaux mécanismes de collision dans le transistor MOS synonyme de réduction de mobilité peuvent être classés en trois catégories dont la description fera l'objet des prochains paragraphes :

- les collisions avec les impuretés chargées (dopant ionisés, charges d'interface et/ou dans l'oxyde),
- les collisions avec les vibrations du réseau (phonons),
- les collisions dues à la rugosité de surface.

Ceci fait l'objet des paragraphes suivants.

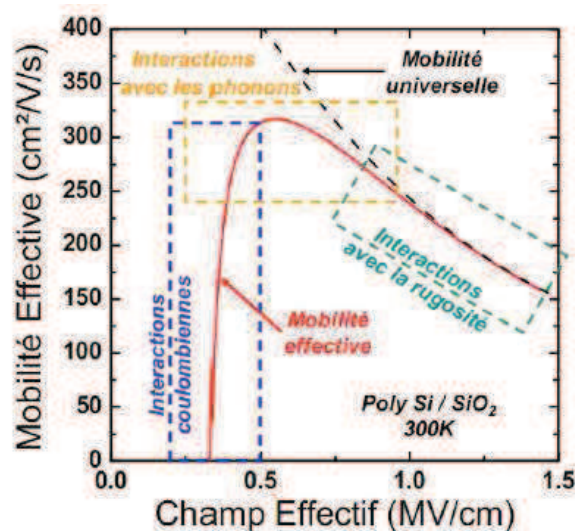


Figure A.20: Principaux mécanismes de réduction de la mobilité dans les transistors MOS avec un empilement de grille de type Si/SiO₂.

A.II.4) Mécanismes de réduction de mobilité

A.II.4.a) Interaction avec les vibrations du réseau (phonons)

A température ambiante, la mobilité est limitée par les vibrations du réseau (les phonons).

Les atomes de silicium formant le réseau ne sont pas strictement fixes. Ils oscillent autour d'une position d'équilibre défini par le minimum de puits de potentiel dans lequel ils se trouvent avec une amplitude d'autant plus grande que la température du réseau T_L est élevée. Les atomes vibrent en groupe selon des modes (longitudinaux et transverses) possédant chacun une pulsation propre ω_0 . Ces vibrations sont assimilées à des quasiparticules susceptibles d'interagir avec les électrons de la couche d'inversion : un phonon peut être absorbé ou émis par un porteur qui échange de ce fait un quantum d'énergie de valeur $\Delta E = \hbar \cdot \omega_0$ et un moment Δp . Ces collisions tendent à limiter la mobilité dans le canal et, puisqu'elles résultent de phénomènes physiques intrinsèques à la structure cristalline du silicium, ne peuvent pas être supprimées ou atténuées simplement. L'influence des phonons sur la mobilité est un des principaux facteurs limitant le transport dans les transistors MOS à température ambiante, le mécanisme d'interaction porteurs-phonons doit donc être compris et modélisé justement.

Leur prise en compte est donc nécessaire pour modéliser le transport dans les transistors MOS.

On distingue cependant 4 types de phonons dans le silicium le long des axes de symétrie [Yu05]:

- longitudinale optique (LO),
- transverse optique (TO),
- longitudinale acoustique (LA),
- transverse acoustique (TA).

Selon [Sze81] la mobilité résultant des interactions des porteurs avec les phonons acoustiques peut être modélisée par la relation (A.18). Le nombre de phonons (i.e. la dégradation du transport) augmente avec la température et résulte en une dépendance théorique de la mobilité phonon en fonction de la température qui respecte $\mu_{PH} \sim T^{-n}$. Dans cette dernière expression, C_{11} est une constante élastique du silicium.

$$\mu_{PH} = \frac{\sqrt{8 \cdot \pi} \cdot q \cdot \hbar \cdot C_{11}}{3 \cdot E_{ds} \cdot m^{*5/2} \cdot (k_B \cdot T)^n} \quad (A.18)$$

Les phonons optiques n'interviennent dans le régime de mobilité que dans une gamme de température inférieure à 100K. La valeur de n est souvent prise entre 1 et 3/2 selon le matériau et la technologie.

A.II.4.b) Interactions de coulomb

L'interaction de Coulomb est suffisamment forte pour que des centres chargés situés à quelques dizaines d'angströms de la couche d'inversion puissent avoir une influence sur le transport électronique dans le canal (figure A.21). La réduction des dimensions, et plus précisément l'amincissement de l'oxyde de grille parallèlement à l'introduction de diélectriques à haute permittivité (high-k) et d'une grille métallique sont ainsi responsables d'un nouveau phénomène: le Remote Coulomb Scattering (RCS). Ce terme anglophone désigne l'interaction additionnelle provenant de centres de diffusion chargés distants, qui proviennent de la forte densité d'états d'interface caractéristique des empilements de grille high-k, mais également de la grille dans le cas d'un oxyde SiO₂ [Cassé06].

La réduction de mobilité correspondante a été mise en évidence expérimentalement en comparant la mobilité de dispositifs ayant un tel empilement de grille avec des transistors pour lesquels l'empilement est standard, c'est-à-dire composé d'un dioxyde de Silicium (SiO₂) et d'une grille en silicium polycristallin. L'origine de cette réduction de mobilité semble provenir d'une densité de charges élevée à l'interface HfO₂/SiO₂, pour les transistors dont la grille est à base de nitrure de titane TiN [Cassé06-Weber04].

Une loi empirique définit la mobilité coulombienne $\mu_C \sim T$ [Sah72] :

$$\mu_C = A \cdot \frac{2 \cdot \hbar k_B \cdot T}{\pi^2 \cdot N_I} \quad (\text{A.19}) \quad \text{où } A \text{ est une constante et } N_I \text{ le dopage.}$$

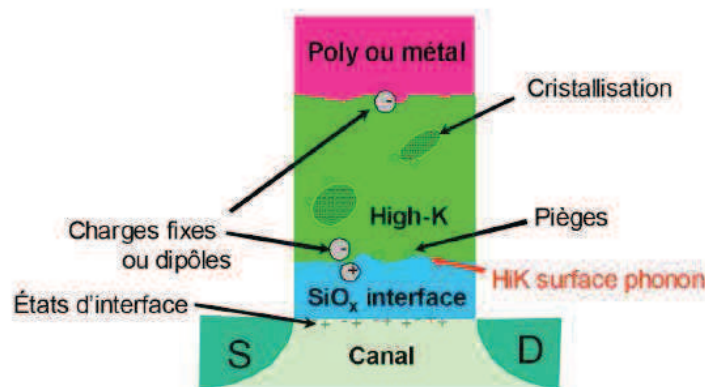


Figure A.21: Illustration des différents mécanismes qui peuvent dégrader la mobilité dans un transistor ayant un empilement de grille high-k métal, d'après [Weber05]

A.II.4.c) La rugosité de surface

L'interface canal/oxyde présentant des défauts, les porteurs de la couche d'inversion confinés à proximité sous l'effet d'un fort champ électrique transverse subissent des collisions qui ne sont pas sans conséquence sur la mobilité. On peut définir la mobilité effective μ_{eff} en fonction du champ effectif par la relation suivante :

$$\mu_{eff} = \frac{L}{W} \cdot \frac{I_D}{V_D} \cdot \frac{\eta}{\xi_{eff} \cdot \epsilon_{SC} \cdot Q_{dep}} \quad (A.20).$$

On met ainsi en évidence la dégradation de la mobilité à fort champ effectif, donc à fort V_G (cf. [figure A.20](#)).

A.II.4.d) Les défauts cristallins non chargés (défauts neutres μ_N)

Une autre cause de dégradation de mobilité présente que sur les transistors courts a été démontrée dans [\[Cros06b\]](#). Les états d'interface n'ont pas pu expliquer cette cause de dégradation.

Ces défauts appelés défauts neutres, seraient de nature interstitielle, résultant d'une guérison incomplète lors de l'implantation [\[Erginsoy50-McGill75\]](#). L'équation (A.21) définit la mobilité défauts neutre μ_N et on voit que la température y est absente:

$$\mu_N = \frac{m^* \cdot q^3}{20 \cdot \epsilon_{SI} \cdot \hbar^3 \cdot N_N} \quad (A.21), \quad N_N \text{ est la densité de défauts neutres}$$

A.II.5) Transport non stationnaire

Avec la réduction des dimensions, le modèle de transport stationnaire, c'est-à-dire un transport à l'équilibre et non dépendant du temps arrive à ses limites. En effet, le libre parcours moyen des porteur (temps écoulé entre deux interactions pendant le transit de la source au drain) devient du même ordre de grandeur que les longueurs de grille : Le transport est dit, dans ce cas, non stationnaire.

Lorsque la longueur de grille est réduite en deçà de 100nm, les phénomènes de transport hors d'équilibre qui étaient négligeables sur les transistors longs deviennent influents et doivent être pris en compte. La théorie de transport dérive diffusion devient obsolète et est alors remplacée par celle de survitesse puis de transport quasi-balistique.

A.II.5.a) Le phénomène de survitesse

Dans le cas d'un transport dérive diffusion, les collisions avec des phonons optiques augmente lorsque le champ électrique latéral est élevé (typiquement $E_{\text{par}} > 10^6 \text{V/m}$). Ce phénomène plafonne la vitesse de dérive des porteurs : c'est le phénomène de vitesse de saturation [Jacoboni77]. Ce concept reste entièrement valable pour les transistors longs, pourvu que le champ latéral soit suffisant. Le profil de vitesse et d'énergie des porteurs en fonction de la position dans le canal peut être obtenu par simulation [Baccarini85], comme le montre la figure A.22.

Le profil de vitesse dans la zone d'entrée du canal attire notre attention, elle constitue une zone de transition dans laquelle la vitesse des porteurs dépasse la valeur à l'équilibre ($v_{\text{sat}}=10^7 \text{cm/s}$). Dans cette zone d'environ 100nm de long, le temps de transit des porteurs est inférieur au temps moyen de relaxation en énergie qui représente le temps moyen nécessaire aux porteurs pour se thermaliser à la température du réseau. En d'autres termes, les porteurs ne subissent pas suffisamment de collisions inélastiques (avec les phonons optiques) pour atteindre l'état d'équilibre dès leur entrée dans le canal et ils acquièrent une vitesse supérieure à v_{sat} . Comme les longueurs de canal des dispositifs modernes sont bien en deçà de 100nm, les champs sont bien supérieurs au champ critique nécessaire pour observer la saturation de la vitesse de dérive dans du silicium massif et le canal est une zone de transition où l'équilibre n'est jamais atteint. On peut donc s'attendre à ce que le phénomène de survitesse soit très présent pour les faibles longueurs de canal [Ge01].

Cette théorie a cependant vite été abandonnée au profit de la théorie du transport quasi-balistique.

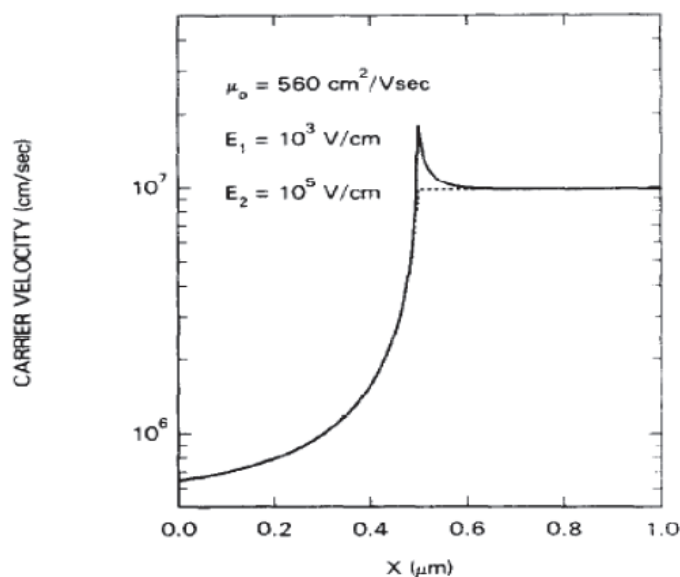


Figure A.22: Vitesse de dérive des porteurs en fonction de la distance du canal à partir de la source [Baccarini85]

A.II.5.b) Modèle de transport balistique (ou quasi-balistique)

Une autre théorie développée pour caractériser le transport dans les transistors courts est la théorie du transport balistique [Lundstrom97-Lundstrom01-Natori01-Natori94]. Ce modèle stipule que les porteurs dans le canal ne subissent aucunes collisions lors de leur transit de la source au drain. On parle de transport quasi-balistique dans le cas où on se trouve à mi-distance entre le modèle dérive diffusion (de nombreuses collisions entre la source le drain) et le modèle balistique (pas de collision entre la source et le drain).

On suppose dans le cas d'un transport balistique la présence d'une barrière de potentiel à la source (appelé aussi source virtuelle), et que seuls les porteurs avec une énergie suffisante franchiront cette barrière pour entrer dans le canal. La vitesse avec laquelle ils franchissent cette barrière de potentiel est appelée vitesse d'injection v_{inj} ou vitesse thermique v_{th} comme illustré sur la figure A.23.

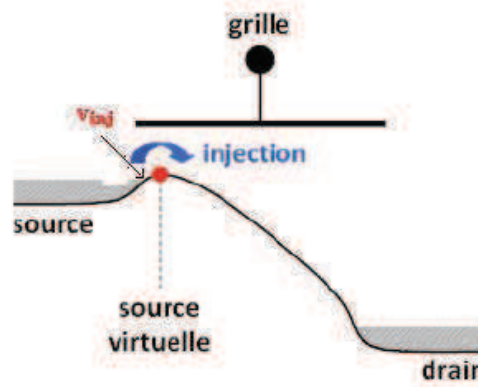


Figure A.23: Schéma simplifié du transistor MOS en régime de conduction : il est composé de deux réservoirs de porteurs reliés par un canal. Ce phénomène se produit au point de maximum de barrière de potentiel coté source appelée « source virtuelle ». [Fleury09a]

Dans le cas où le transport est totalement balistique, tous les porteurs passent directement de la source au drain sans subir aucune interaction : leur temps de transit dans le canal est beaucoup plus court que la durée moyenne entre deux interactions. Dans ce cas, le courant de drain ne dépend plus de la longueur de grille, on a atteint la « limite balistique ». On parle alors de mobilité balistique apparente qui est donnée par :

$$\mu_{bal} = \frac{2.q.L}{m.\pi.v_{th}} \quad (\text{A.22}) \quad [\text{Lee83-Shur02}]$$

Dans l'expression (A.22), m et q sont la masse effective et la charge de l'électron respectivement. v_{th} est la vitesse thermique et est donnée selon [Shur02] par : $v_{th} = \sqrt{\frac{8.k.T}{\pi.m}}$, k est la constante de Boltzmann et T la température.

Les expressions du courant dans ce régime peuvent être données selon le régime de conduction (faible ou fort V_{DS}). On trouvera ainsi dans [Lundstrom01] les expressions du courant quasi-balistiques suivantes :

$$\text{E linéaire : } I_{Dlin} = W.Cox.\frac{v_T}{T} \cdot [1 - r_{lin}] \cdot [V_{GS} - V_{TH}] V_{DS} \quad (\text{A.23})$$

$$2.k.\frac{q}{q}$$

$$\text{En saturation : } I_{Dsat} = W.Cox.v_T \cdot \frac{1 - r_{sat}}{1 + r_{sat}} \cdot [V_{GS} - V_{TH}] \quad (\text{A.24})$$

Dans ces expressions, v_T est la vitesse thermique unidirectionnelle extraite à partir de simulation et varie de 1.2 à 2.10^7 cm.s⁻¹ selon la charge surfacique d'inversion (cf. [figure A.24](#)), r est le coefficient de rétrodiffusion dans le canal. En effet, la théorie de Mark Lundstrom évoque que tous les porteurs qui quittent le canal n'arrivent pas à rejoindre le drain. On définit donc un coefficient r de rétrodiffusion pour quantifier le nombre de porteurs qui retournent à la source virtuelle après avoir franchi la barrière de potentiel. Une illustration de ce phénomène est présentée à la [figure A.25](#) [Lundstrom02].

On définit λ_0 comme le libre parcours moyen entre deux collisions et ℓ la longueur critique (Au-dessus de cette longueur, la rétrodiffusion n'est plus possible) pour une rétrodiffusion à fort V_{DS} ([figure A.25](#)), on définit r comme suit :

$$\text{E linéaire : } r_{lin} = \frac{L}{L + \lambda_0} \quad (\text{A.25})$$

$$\text{En saturation : } r_{sat} = \frac{\ell}{\ell + \lambda_0} \quad (\text{A.26})$$

Il faut noter que λ_0 est défini en considérant un canal long où la mobilité μ est clairement définie :

$$\lambda_0 = \frac{2 \cdot \mu \cdot k \cdot T}{q \cdot v_T} \quad (\text{A.27})$$

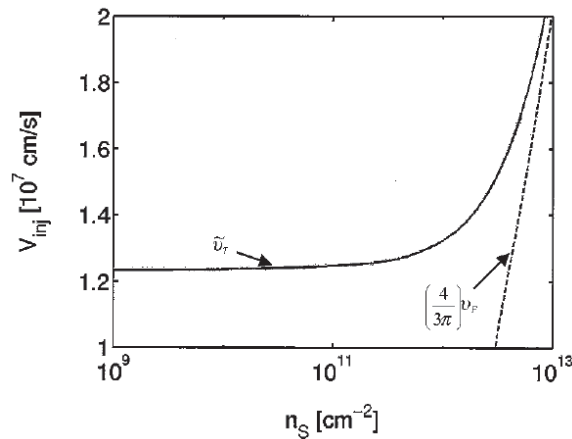


Figure A.24: Variation de la vitesse d'injection (ou thermique) en fonction de la charge surfacique d'inversion [Lundstrom02]

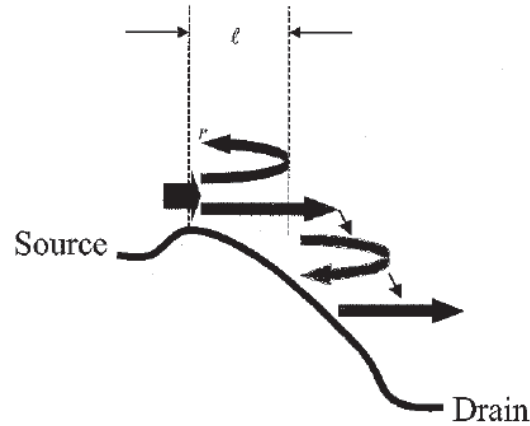


Figure A.25: Illustration de la rétrodiffusion dans un MOSFETs à fort V_{DS} . La longueur critique en dessous de laquelle les porteurs reviennent à la source est illustrée [Lundstrom02].

A.II.5.c) Balisticité et dégradation de mobilité

En égalisant l'expression (A.23) du courant en régime balistique à faible V_{DS} avec l'expression du courant en régime classique de dérive, on démontre les expressions suivantes [Cros06a] :

$$\mu_{bal} = \frac{L \cdot q \cdot v_T}{2 \cdot k \cdot T} \quad (\text{A.28}) \quad \text{avec} \quad v_T = \sqrt{\frac{2 \cdot k \cdot T}{\pi \cdot m}}$$

Dans ces conditions, I_{Dlin} devient proportionnel à $\frac{\mu_{bal}}{L}$. Nous savons que dans le cas d'un transport balistique, on arrive à un moment où le courant n'augmente plus avec la diminution de L . Il faudrait donc que μ_{bal} diminue pour combler la diminution de L dans l'expression du courant. C'est ainsi qu'une dégradation de la mobilité dans les transistors courts a été dans un premier temps attribuée à la balisticité. Ces allégations ont vite été mises à plat par la suite dans [Cros06b]. La limite balistique étant encore loin dans de tels dispositifs (encore assez longs). Il a été démontré que la dégradation de la mobilité observée (à ce moment-là) dans les transistors ultra-courts provenait de la présence de centre non chargés issus de la non-guérison complète de dopants.

A.II.5.d) Unification des mécanismes de transport : dérive diffusion et quasi-balistique

L'idée est de stipuler que les deux mécanismes de transport peuvent être présents dans un dispositif en même temps. Cela se traduit par la loi de Mathiessen. Cette loi peut s'écrire aussi bien en mobilité qu'en courant [Ghibaudo09]:

$$\frac{1}{\mu_{eff}} = \frac{1}{\mu_{dd}} + \frac{1}{\mu_{bal}} \text{ ou bien } \frac{1}{I_{eff}} = \frac{1}{I_{dd}} + \frac{1}{I_{bal}} \quad (\text{A.29})$$

Comme stipulé dans [Fleury09b], on peut partir avec l'hypothèse avec laquelle on ne connaît pas la nature du transport. Dans ce cas, on peut reformuler l'équation (A.29) en régime saturé par :

$$\frac{1}{I_{don}} = \frac{1}{I_{ddon}} + \frac{1}{I_{lim}} \quad (\text{A.30}).$$

I_{don} étant le courant à l'état ON du transistor, I_{ddon} est le courant dérive diffusion à l'état ON du transistor.

I_{lim} est dans ce cas le courant limitant (saturation ou quasi-balistique) et est défini par (une sorte de loi de Mathiessen) :

$$\frac{1}{I_{lim}} = \frac{1}{I_{sat}} + \frac{1}{I_{bal}} \Rightarrow I_{lim} = \min(I_{sat}, I_{bal}) \quad (\text{A.31})$$

De la même manière, en linéaire on peut aussi définir le courant linéaire :

$$\frac{1}{I_{dlin}} = \frac{1}{I_{ddlin}} + \frac{1}{I_{ballin}} \quad (\text{A.32})$$

On peut aussi définir le courant limitant en fonction d'une vitesse limitante v_{lim} , de la charge d'inversion Q_s et de la largeur W du transistor [Fleury09b]:

$$I_{lim} = W \cdot v_{lim} \cdot Q_s \quad (\text{A.33})$$

En combinant les équations (A.30) à (A.32) avec les équations de dérive diffusion (seront montrées dans la partie résultat), il a été démontré dans [Fleury09b] que l'on peut calculer v_{lim} sans hypothèse préalable sur la nature du transport. L'étude en température de ce paramètre v_{lim} permet de conclure quant à la nature du transport. Cette méthode expérimentale dite « méthode de v_{lim} », vient appuyer toutes les méthodes de simulation qui ont été développées sur le sujet de la balisticité du transport. En revanche, la difficulté de cette méthode réside en la connaissance de la vraie charge d'inversion Q_s à fort V_{DS} . Elle est approximée jusqu'à aujourd'hui par la relation :

$$Q_s = C_{ox} \cdot (V_{GS} - V_{TH}) \quad (A.34)$$

Une incertitude pèse non seulement sur la validité de la formule (A.33) (particulièrement dans les transistors courts) mais aussi sur l'extraction de V_{TH} à fort V_{DS} . Nous reviendront sur cette méthode dans la partie «extraction de paramètres» en l'appliquant à une technologie silicium et germanium bulk. Nous verrons les améliorations que nous avons apportées à la méthodologie afin de la rendre plus fiable. Nous verrons aussi les limites réelles de cette technique en comparant les valeurs de v_{lim} extraites sur silicium bulk avec des vitesses d'injection obtenues avec des simulations Monte-Carlo. La critique se portera particulièrement sur la validité de la loi de Mathiessen telle qu'elle a été utilisée dans [Fleury09b].

Table des figures de la partie A

Figure A.1: transistor nMOS à l'état bloqué ($V_g = 0$), valide aussi pour les dispositifs pMOS en inversant les polarisations	9
Figure A.2: transistor nMOS à l'état passant ($V_g > 0$), valide aussi pour les dispositifs pMOS en inversant les polarisations	9
Figure A.3: Schéma de bande d'un semiconducteur dopé N (a) et dopé P (b).....	11
Figure A.4: Architecture d'un transistor nMOS sur silicium massif, d'après [Skotnicki03]	12
Figure A.5 : Illustration des trois régimes de conduction qui caractérisent le fonctionnement d'un transistor nMOS [Skotnicki03]	15
Figure A.6: Schéma électrique équivalent du MOSFET - influence des résistances séries de source et de drain	16
Figure A.7: Illustration de l'abaissement de la barrière de potentiel source-drain induit par les effets de canaux courts (CSE et DIBL).....	17
Figure A.8: Courant de drain en fonction de la tension de grille pour un transistor long ($L=1\mu\text{m}$) et un transistor ultracourt ($L=40\text{nm}$) à faible (10mV) et forte (1V) polarisation de drain.	18
Figure A.9: L'évolution des technologies selon la roadmap ITRS 2011, actualisé pour 2012	19
Figure A.10: Illustration de la contrainte compressive du canal par les sources/drain via un mismatch de paramètre de maille entre le matériau source/drain et le matériau canal [Thompson04].....	21
Figure A.11: Illustration d'un canal silicium contraint dans une heterostructure [Lee05].....	21
Figure A.12: Illustration d'une contrainte extensive via CESL sur le canal [Wang10]	22
Figure A.13: Coupe TEM de pMOSFETs SiGe contraint avec le stack de grille correspondant [Alatise09].....	24
Figure A.14: Image TEM d'un transistor Ge-pMOSFETs [Mitard08].....	25
Figure A.15: Figure illustrant la combinaison de III/V-nMOSFET et GE-pMOSFET sur isolant [Takagi07].....	26
Figure A.16: l'architecture d'un MOS type SOI (a) et coupe TEM d'un transistor MOS SOI (b) [Beranger07]	27

Table des figures de la partie A

Figure A.17: Schéma simplifié illustrant une architecture trigate (a), coupe TEM d'un transistor trigate (b) [Auth12].....	28
Figure A.18: Benchmark entre du SOIFF et du BFF [Chiarella10]	29
Figure A.19: Illustration de la trajectoire des électrons subissant des collisions dans un cristal. a) sans champ électrique b) sous l'application d'un champ électrique [Grove67]	31
Figure A.20: Principaux mécanismes de réduction de la mobilité dans les transistors MOS avec un empilement de grille de type Si/SiO ₂	34
Figure A.21: Illustration des différents mécanismes qui peuvent dégrader la mobilité dans un transistor ayant un empilement de grille high-k métal, d'après [Weber05]	36
Figure A.22: Vitesse de dérive des porteurs en fonction de la distance du canal à partir de la source [Baccarini85]	39
Figure A.23: Schéma simplifié du transistor MOS en régime de conduction : il est composé de deux réservoirs de porteurs reliés par un canal. Ce phénomène se produit au point de maximum de barrière de potentiel coté source appelée « source virtuelle ». [Fleury09a].....	40
Figure A.24: Variation de la vitesse d'injection (ou thermique) en fonction de la charge surfacique d'inversion [Lundstrom02]	41
Figure A.25: Illustration de la rétrodiffusion dans un MOSFETs à for V_{DS} . La longueur critique en dessous de laquelle les porteurs reviennent à la source est illustrée [Lundstrom02]......	42

Liste des tableaux de la partie A

Tableau A.1: Tableau périodique des éléments: Mendeleïev 1870 11

Partie B

Table des matières de la partie B

Partie B	48
Table des matières de la partie B	49
PARTIE B: INFLUENCE DU PROCEDE DE FABRICATION ET DE L'ARCHITECTURE SUR LES PERFORMANCES DES TRANSISTORS	51
CHAPITRE B.I : LE RECUIT HAUTE PRESSION SOUS H ₂ OU D ₂	52
B.I.1) Performances électriques à température ambiante	53
B.I.1.a) Mobilité à faible charge.....	54
B.I.1.b) Résistance source-drain.....	55
B.I.1.c) Le pompage de charge pour la caractérisation de l'interface Si/SiO ₂	56
B.I.1.d) La variation de la tension de seuil après HPFA.....	57
B.I.2) Etude de la mobilité à faible charge à basse température	58
B.I.2.a) Description du modèle utilisé	58
B.I.2.b) Effets du HPFA sur les différentes contributions de mobilité à faible charge.....	60
B.I.3) Hydrogène ou Deutérium: Comparaison en mobilité effective, pompage de charge, HCI (hot carrier injection) et NBTI (Negative bias temperature instability).....	62
B.I.3.a) Mobilité effective.....	62
B.I.3.b) HCI et NBTI.....	63
B.I.4) Bilan et conclusion	64
CHAPITRE B.II : ARCHITECTURE CANAL A BASE DE SILICIUM-GERMANIUM.....	65
B.II.1) Pourquoi introduire une proportion de germanium dans le canal?.....	65
B.II.2) Canal silicium-germanium.....	67
B.II.2.a) Simulations.....	67
B.II.2.b) Analyse à température ambiante.....	68
B.II.2.c) Analyse à basse température	74

Table des matières de la partie B

B.II.2.d) Mesures de bruit basse fréquence.....	78
B.II.3) Discussion et conclusion.....	82
CHAPITRE B.III : L'OXYDE DE GRILLE : LE HIGH- κ ET SON IMPACT	84
B.III.1) Pourquoi l'introduction du high- κ ?.....	84
B.III.2) Les inconvénients du high-k.....	84
B.III.3) Les transistors de type GO1 et GO2.....	86
B.III.4) Etude comparative des transistors GO1 et GO2	87
B.III.5) Bilan et conclusion.....	95
Table des figures de la partie B.....	96
Liste des tableaux de la partie B.....	101

PARTIE B: INFLUENCE DU PROCEDE DE FABRICATION ET DE L'ARCHITECTURE SUR LES PERFORMANCES DES TRANSISTORS

Une solution alternative à la miniaturisation pour augmenter les performances est de faire évoluer le procédé de fabrication ou l'architecture du transistor. De son invention à maintenant, le procédé de fabrication et l'architecture ne cessent d'évoluer. Dans cette partie, l'influence du recuit haute pression est étudiée dans un premier temps.

L'effet d'un matériau alternatif comme le silicium-germanium (SiGe) est aussi étudié au deuxième chapitre.

Le troisième chapitre parle de l'impact du high- κ sur le transport électronique.

CHAPITRE B.I : LE RECUIT HAUTE PRESSION SOUS H₂ OU D₂

Pendant son processus de fabrication, le transistor passe par plusieurs étapes avec des températures plus ou moins élevées. Les hautes températures peuvent avoir des effets négatifs sur l'interface silicium-oxyde. En effet, les liaisons Si-O ne supportent pas les fortes températures [Sahrani00] et des liaisons Si pendantes existent à la fin du processus de fabrication. Pour guérir ces liaisons Si pendantes, des recuits en température en présence d'un gaz comme l'hydrogène sont réalisés (figure B.1).

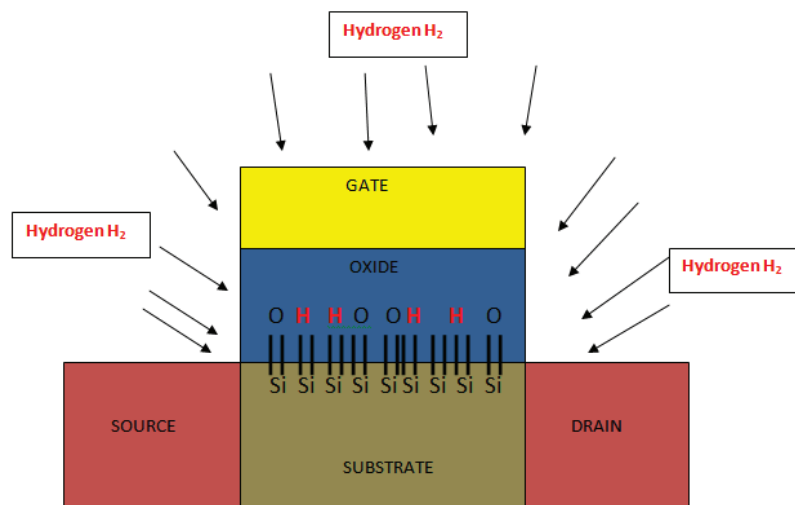


Figure B.1: Principe du recuit standard en présence de l'hydrogène

Ainsi le recuit standard sous atmosphère hydrogène (STDFH) est connu pour améliorer les propriétés électriques du transistor en passivant l'interface Si-SiO₂. Alternativement, sur les dispositifs avec un stack de grille à base de high- κ , le deutérium peut être utilisé pour améliorer la fiabilité (meilleure durée de vie) [Choi02]. Pour bénéficier d'une amélioration supérieure des performances, le recuit à haute pression sous hydrogène (HPH₂) ou deutérium (HPD₂) est utilisé [Song09]. Cependant, l'état de l'art montre que les études se sont focalisées principalement sur les performances de fiabilité ou de bruit [Cho06-Park09].

C'est ainsi que ce chapitre aborde une étude poussée des effets du recuit à haute pression sur le transport électronique en 28nm bulk. En plus de l'impact sur la mobilité et la tension de seuil, l'influence sur les états d'interface rapides, les résistances source- drain, les effets de porteurs chauds, et, la dérive de V_{TH} sont aussi conjointement abordés [Diouf12a].

Les dispositifs étudiés ont un T_{inv} de 1.4nm pour les NMOS et 1.7nm pour les PMOS [Arnaud09]. Ils sont constitués d'une grille métal et d'oxyde de grille de type high- κ au-dessus d'une couche interfaciale, « interfacial layer (IL) », en SiON. Les PMOS ont du SiGe épitaxié comme canal (figure B.2).

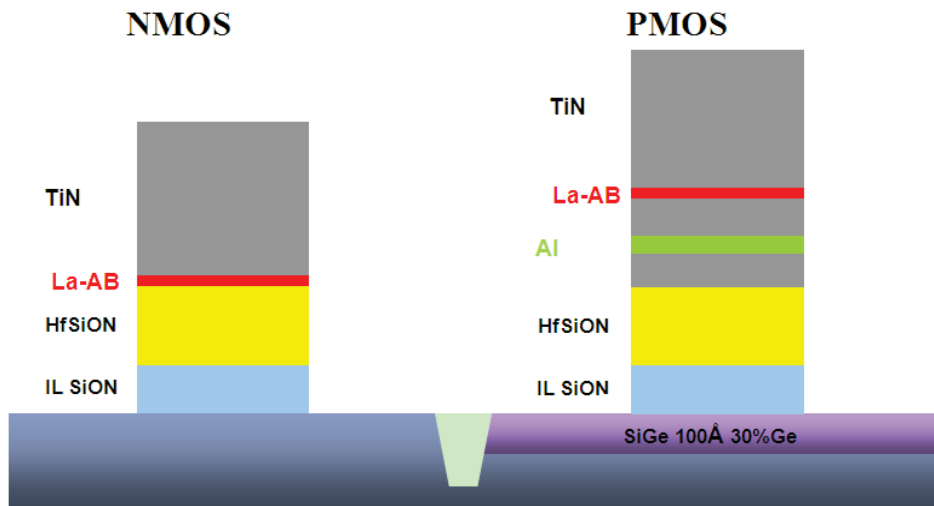


Figure B.2: Empilement du stack de grille des N&PMOS étudiés dans ce chapitre

B.I.1) Performances électriques à température ambiante

Le but ultime du recuit est d'augmenter les performances du transistor. Cette amélioration doit être visible sur le courant I_{ON} à l'état ON du dispositif. La figure B.3 compare les courant I_{ON} après recuit standard (STDFA), sans recuit (NOFA), recuit hydrogène HPH2 et deutérium HPD2 pour N & PMOS.

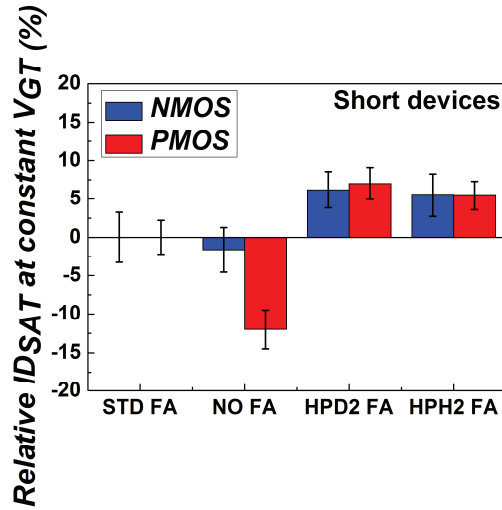


Figure B.3: Illustration de l'amélioration de courant ION après recuit HPH2 et HPD2

Le recuit standard est pris en référence dans cette figure. Comme attendu, l'amélioration du courant à V_{GT} ($V_{GT}=V_G-V_{TH}$) constant (I_{ON}) sans recuit pour N & PMOS est négative. En revanche le HPH2 et HPD2 permettent d'augmenter de façon significative le courant I_{ON} . Aussi bien pour les N et les PMOS, on atteint des améliorations qui avoisinent les 6% sur des dispositifs nominaux. Il faudra maintenant comprendre l'origine de cette amélioration de courant après recuit HPH2 ou HPD2.

B.I.1.a) Mobilité à faible charge

Les figures B.4a & B.4b montrent la mobilité à faible charge μ_0 après recuit pour N & PMOS respectivement.

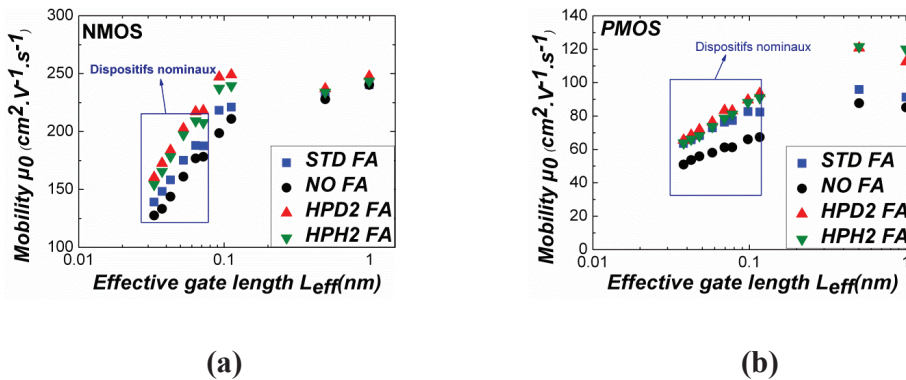


Figure B.4: Mobilité à faible charge après recuit standard, HPH2 et HPD2 pour NMOS (a) et PMOS (b)

Les N & PMOS montrent un comportement différent.

- Les NMOS montrent une amélioration de la mobilité à faible charge après recuit haute pression (HPH2 ou HPD2) pour les dispositifs nominaux. Cela explique l'augmentation de courant I_{ON} observé sur la [figure B.3](#) pour les NMOS. Les transistors longs ne sont pas améliorés en mobilité après le recuit haute pression (HPFA).
- Les PMOS montrent un comportement opposé : les dispositifs nominaux ne sont pas améliorés après HPFA. Seulement la mobilité des transistors longs est sensible au recuit HPFA. Il faudra donc trouver d'autres explications pour l'amélioration du courant I_{ON} dans transistors nominaux des PMOS.

B.I.1.b) Résistance source-drain

Les investigations sur les résistances source-drain extraites selon [\[Fleury09c\]](#) montrent une diminution de 25% des R_{SD} pour les PMOS après HPFA ([figure B.5](#)) alors que celles des NMOS ne sont pas impactées. Sachant que les R_{SD} n'influencent que les transistors courts, cela peut expliquer l'augmentation du courant I_{ON} observée pour les PMOS nominaux sur la [figure B.3](#).

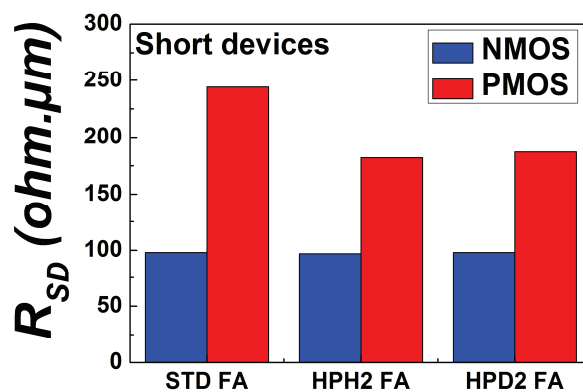


Figure B.5: Comportement des résistances source-drain R_{SD} après HPFA

B.I.1.c) Le pompage de charge pour la caractérisation de l'interface Si/SiO₂

Le pompage de charge est utilisé pour la caractérisation des états d'interface rapides à l'interface SiON/Substrat (cf. [figure B.2](#)).

- Principe

Dans le cas d'un PMOS l'application d'un pulse négatif à la grille du transistor crée des charges minoritaires provenant des régions source-drain P+. Lorsqu'on retire le pulse, certains des porteurs minoritaires se recombinent avec des porteurs majoritaires dans les centres recombinants que constituent les états d'interface SiO₂/substrat. De cette recombinaison découle un courant DC moyenné sur une période appelé I_{CP} . A partir du courant I_{CP} on peut calculer la densité d'état d'interface N_{IT} en fonction de la fréquence F selon [\[Brugler69\]](#):

$$N_{IT} = \frac{I_{CP}}{A.q.F.\Delta V_s} \quad (\text{B.1})$$

ΔV_s est le balayage du potentiel de surface ($\approx 1V$), A et q sont la surface du transistor et la charge de l'électron, respectivement.

Le pompage de charge est souvent mesuré sur des transistors à grande surface ou le courant de grille peut perturber le courant I_{CP} . Pour éviter cela surtout pour les faibles fréquences, on fait la mesure à deux fréquences F_1 et F_2 proches telles que $F_1 = \beta.F_2$. Dans ces conditions, $I_{CP1} = \beta.I_{CP2} + I_G$.

$I_{CP1} - I_{CP2}$ permet ainsi de corriger le courant de grille et on démontre que [\[Masson99\]](#):

$$I_{CP2} = \frac{I_{CP1} - I_{CP2}}{\beta - 1} \quad (\text{B.2})$$

- Pompage de charge avant et après HPFA

Les [figures B.6a & B.6b](#) montrent une amélioration de la qualité de l'interface après HPFA pour les N & PMOS. Le pompage de charge n'explique donc pas la différence d'amélioration en mobilité μ_0 observée sur les transistors longs entre les N & PMOS. Les PMOS montrent

cependant une plus grande densité d'états d'interface. Ceci devrait être dû au fait que les PMOS ont un empilement de grille plus complexe que celui du NMOS (cf. [figure B.2](#)). Malgré la correction de l'impact du courant de grille sur le courant de pompage de charge, le N_{IT} des NMOS n'est pas constant en fonction de la fréquence ([figure B.6a](#)).

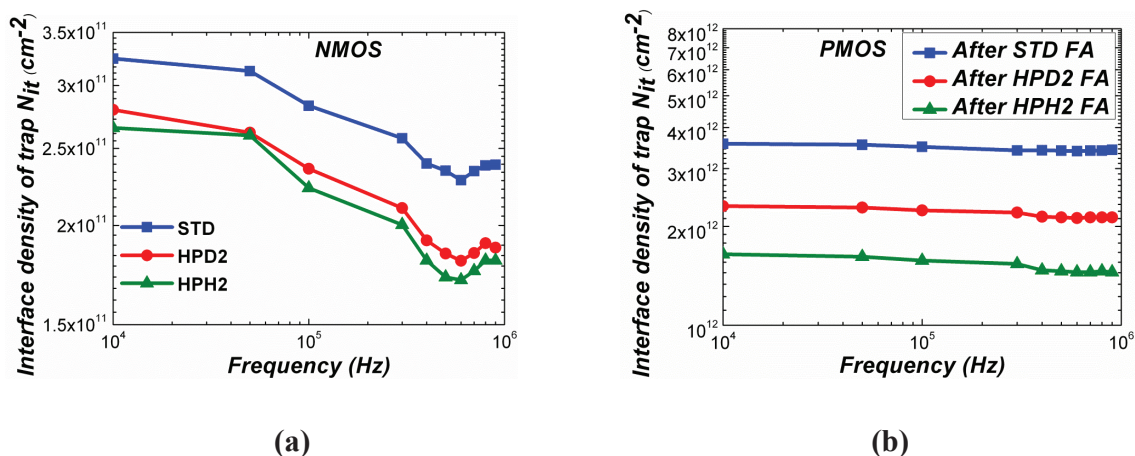


Figure B.6: Extraction de la densité d'états d'interface par pompage de charge après recuit standard, HPH2 et HPD2 pour NMOS (a) & PMOS (b)

B.I.1.d) La variation de la tension de seuil après HPFA

La tension de seuil V_{TH} diminue après HPFA pour les NMOS ([figures B.7a](#)) tandis qu'elle augmente en valeur absolue pour les PMOS ([figures B.7a](#)) (diminue en valeur relative). Ce résultat n'est pas en accord avec ce qu'on peut voir en pompage de charge ([figure B.6](#)). Cela montre qu'en plus de passiver l'interface, le HPFA crée ou déplace le même type de charge (ou défauts) pour les N & PMOS. Plus probablement, le HPFA augmente la quantité de charges positives comme les lacunes en oxygène comme décrit dans [\[Robertson07\]](#), ou réduit la quantité de charges négatives.

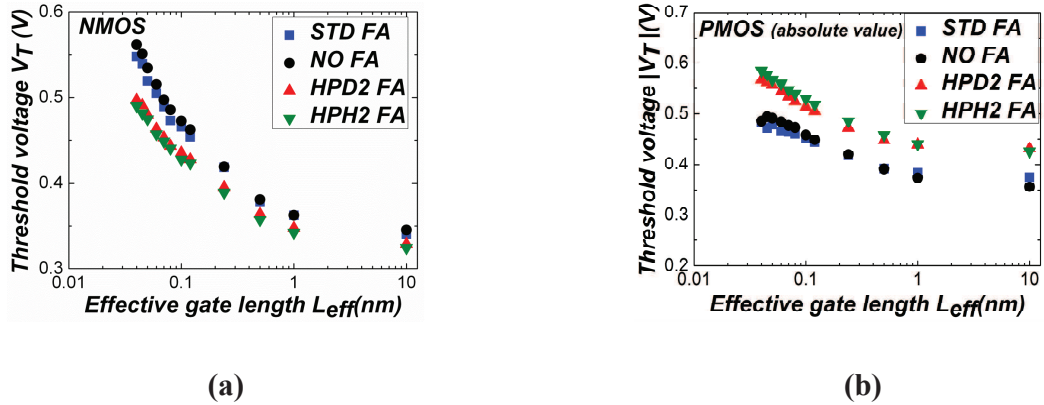


Figure B.7: Variation de la tension de seuil après HPFA pour NOMS (a) & PMOS (b)

B.I.2) Etude de la mobilité à faible charge à basse température

Un modèle empirique de mobilité est utilisé pour l'étude à basse température [Cros06]. Ce modèle permet de séparer les différentes contributions de la mobilité à faible charge : diffusion par les phonons, remote Coulomb scattering et diffusion sur les défauts neutres.

B.I.2.a) Description du modèle utilisé

Une loi de type Mathiessen (B.3) permet de décrire la mobilité à faible charge μ_0 en fonction des différentes contributions citées au début de ce paragraphe.

$$\frac{1}{\mu_0(T)} = \frac{1}{\mu_{0PH}(T)} + \frac{1}{\mu_{0C}(T)} + \frac{1}{\mu_{0N}(T)} \quad (\text{B.3})$$

$\mu_0(T)$ est la mobilité à faible charge,

$$\mu_{0PH}(T) \text{ est la contribution due aux phonons : } \mu_{0PH}(T) = \frac{300 \cdot \mu_{PH}}{T} \quad (\text{B.4})$$

$$\mu_{0C}(T) \text{ est la contribution due au remote Coulomb scattering : } \mu_{0C}(T) = \frac{T \cdot \mu_C}{300} \quad (\text{B.5})$$

$$\mu_{0N}(T) \text{ est la contribution due aux défauts neutres [Cros06] : } \mu_{0N}(T) = \mu_N \quad (\text{B.6})$$

μ_{PH} , μ_C et μ_N sont des constantes indépendantes de la température et relatives aux collisions sur les phonons, remote Coulomb et défauts neutres respectivement.

On peut ainsi ré-exprimer la mobilité à faible charge μ_0 en fonction des différentes contributions par la relation (B.7) :

$$\frac{1}{\mu_0(T)} = \frac{T}{300 \cdot \mu_{PH}} + \frac{300}{T \cdot \mu_C} + \frac{1}{\mu_N} \quad (B.7)$$

L'équation (B.7) est utilisée pour déterminer les constantes μ_{PH} , μ_C et μ_N en ajustant les courbes de μ_0 (valeurs expérimentales) en fonction de la température comme sur la figure B.8. On peut ainsi tracer les contributions $\mu_{0PH}(T)$, $\mu_{0C}(T)$ et $\mu_{0N}(T)$ dans un même graphe pour analyser le poids de chacune d'elles sur μ_0 .

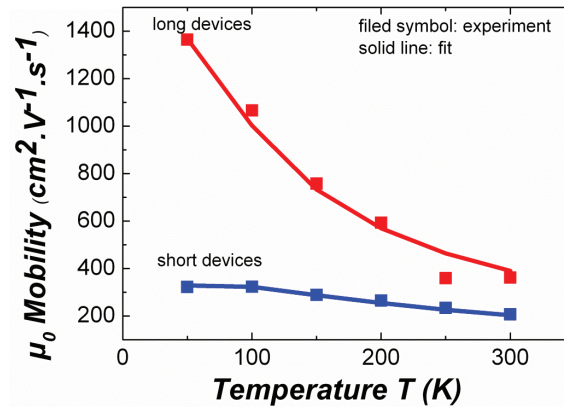


Figure B.8: Mobilité à faible charge μ_0 en fonction de la température

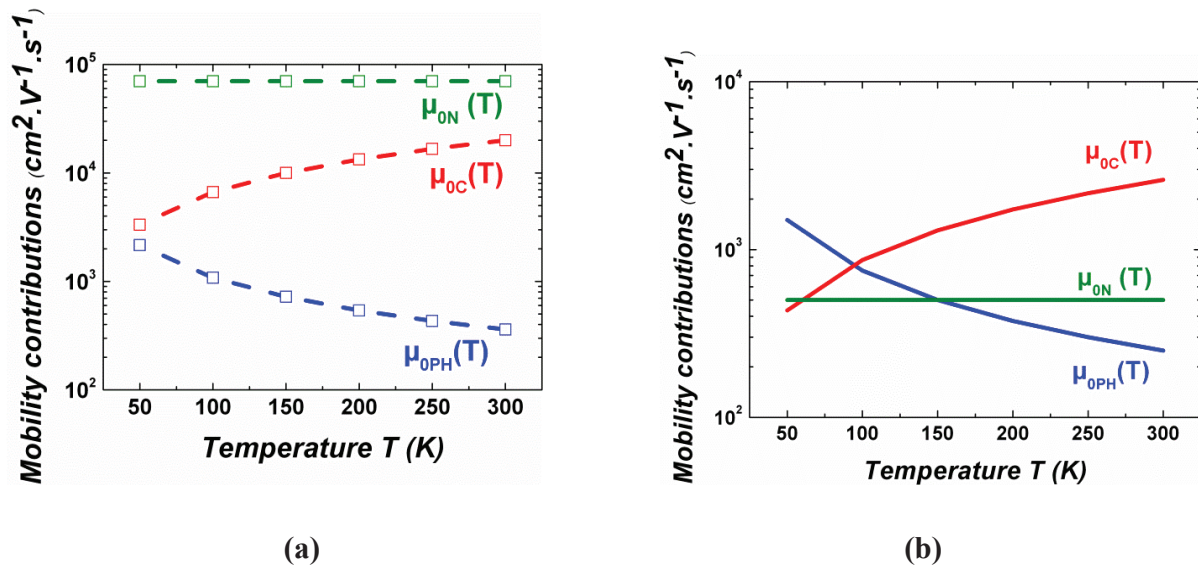


Figure B.9: Représentation des différentes contributions de la mobilité à faible charge pour un transistor long (a) et pour un transistor court (b)

L'analyse des figures B.9a & B.9b et l'équation (B.3) montre une différence entre les transistors longs et courts. En effet la contribution due au phonon scattering $\mu_{OPH}(T)$ domine dans les transistors longs. Dans les transistors courts, les contributions dues aux remote Coulomb scattering $\mu_{OC}(T)$ et aux défauts neutres $\mu_{ON}(T)$ ont plus de poids. Nous verrons maintenant si cette différence de comportement des contributions de mobilité selon les transistors long et courts, N & PMOS peut affecter l'effet du HPFA.

B.I.2.b) Effets du HPFA sur les différentes contributions de mobilité à faible charge

Nous allons étudier N & PMOS séparément.

- NMOS

La figure B.10a montre que sur les NMOS à canal long la mobilité limitée par les phonons est dominante (cf. B.3). Cela explique l'absence d'amélioration de la mobilité à faible charge sur les NMOS longs après HPFA (cf. figure B.4a). En effet, il est compréhensible que l'interaction avec les phonons ne soit pas impactée après recuit.

Pour les NMOS courts, la contribution du remote Coulomb scattering augmente et devient non négligeable (figure B.10a). L'amélioration des paramètres de remote Coulomb scattering après HPFA (figure B.10b) se traduit donc par une amélioration de la mobilité à faible charge observée sur les transistors courts à la figure B.4a.

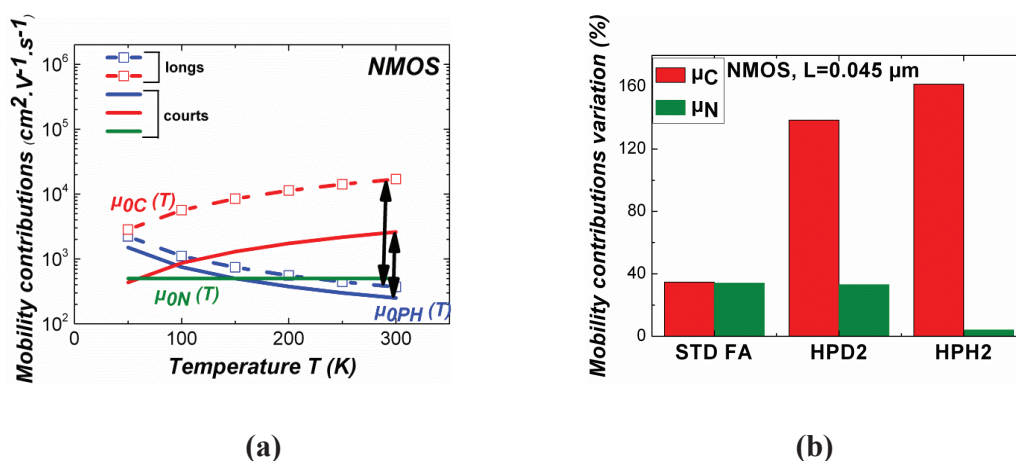


Figure B.10: Illustration des différentes contributions de la mobilité d'un transistor long et court n'ayant subi aucun recuit (a); Evolution des constantes de remote coulomb scattering et des défauts neutres après recuit (b)

- PMOS

Contrairement aux transistors NMOS, les transistors PMOS longs et courts présentent le même niveau de remote Coulomb scattering (figure B.11a). Cette contribution élevée de remote Coulomb scattering dans les PMOS long explique l'amélioration de la mobilité à faible charge observée sur la figure B.4b. La figure B.11b montre clairement l'amélioration de la constante de remote Coulomb scattering. Dans les PMOS courts, l'apparition des défauts neutres masque la contribution de remote Coulomb. Il faut cependant noter que la possible amélioration des constantes de défauts neutres (figure B.11c) n'a visiblement pas eu beaucoup d'impact sur la mobilité à faible charge, d'où la faible amélioration de mobilité observée sur les transistors PMOS courts à la figure B.4b.

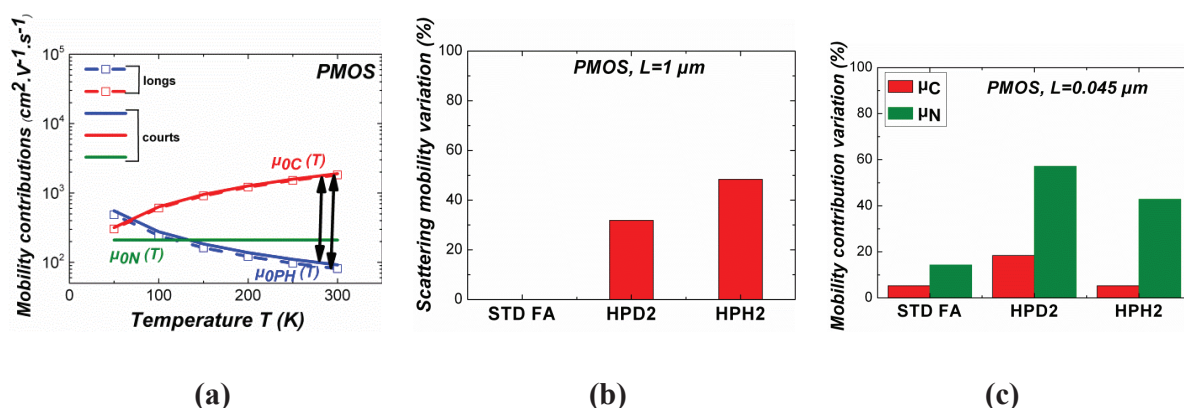


Figure B.11: Illustration des différentes contributions de la mobilité d'un transistor long et court n'ayant subi aucun recuit (a); Evolution des constantes de remote coulomb scattering sur un PMOS long après recuit (b); Evolution des constantes de remote coulomb scattering et de défauts neutres sur un PMOS court après recuit (c)

Nous venons d'étudier les effets du recuit haute pression sans grande distinction entre l'hydrogène et le deutérium. La dernière partie de cette partie va se focaliser sur la comparaison de ces deux gaz utilisés pour le recuit.

B.I.3) Hydrogène ou Deutérium: Comparaison en mobilité effective, pompage de charge, HCI (hot carrier injection) et NBTI (Negative bias temperature instability)

B.I.3.a) Mobilité effective

La [figure B.12a](#) montre la mobilité effective d'un transistor PMOS après un recuit standard, HPH2 et HPD2. Comparé au recuit standard, le recuit HPH2 améliore davantage la mobilité que le recuit HPD2. Cette différence est expliquée par la [figure B.6b](#). Sur cette celle-ci, on voit que le recuit HPH2 passive mieux l'interface que le recuit HPD2 en permettant d'avoir moins d'états d'interface. Il faut noter que le degré de passivation sur les NMOS est le même entre HPH2 et HPD2 ([figure B.6a](#)). Cependant cette amélioration de l'interface n'est pas répercutée sur la mobilité effective ([figure B.12b](#)) à cause des raisons que nous avons évoquées au paragraphe [I.2.b](#).

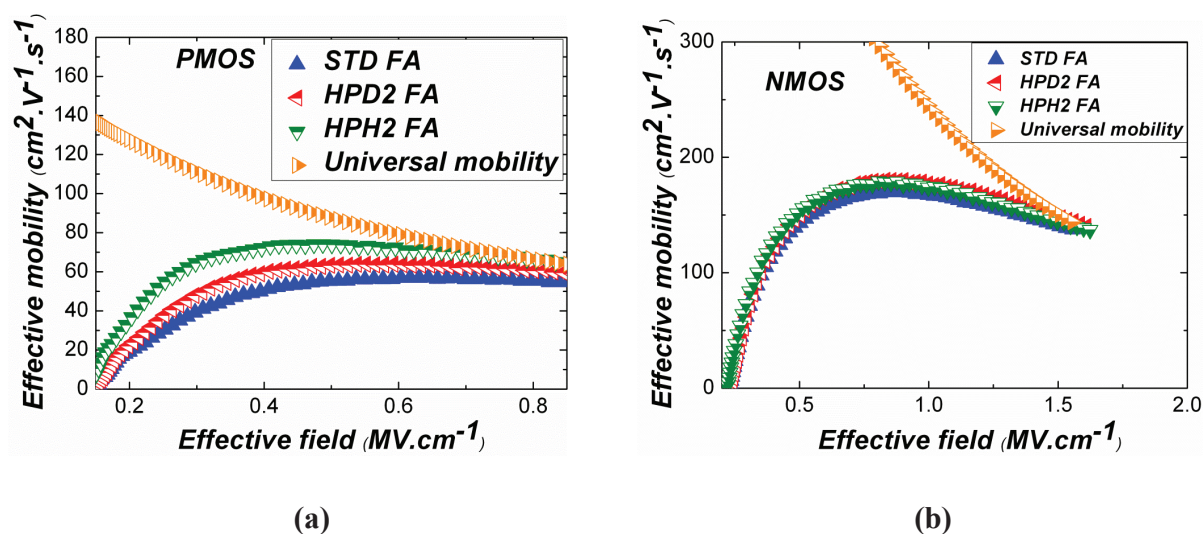
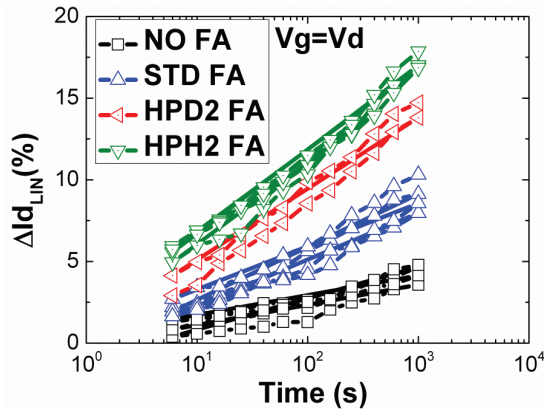


Figure B.12: Mobilité effective après recuit standard, HPH2 et HPD2 pour un PMOS (a) et un NMOS (b)

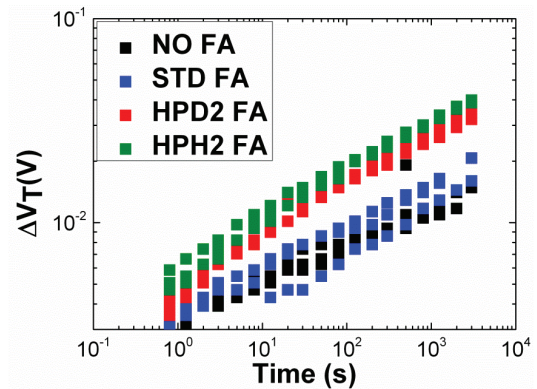
B.I.3.b) HCI et NBTI

En terme de fiabilité, le HPFA montre davantage de dégradation en HCI et en NBTI comparé au recuit standard et sans recuit. Il est important de noter que la dégradation en HCI est bien corrélée avec les états d'interface alors que le NBTI est plutôt associé à une variation du V_{TH} . Alors que la dégradation en HCI est probablement liée à la dé-passivation des liaisons pendantes passivées pendant le recuit, la variation en NBTI est reliée à la dé-passivation des lacunes d'oxygène.

Cependant dans notre cas, le HPH2 montre une plus grande dégradation comparée à celle du HPD2 aussi bien en HCI qu'en NBTI (figure B.13a & B.13b). Cela peut être relié à la meilleure passivation de l'interface par le H_2 comme vu en pompage de charge à la figure B.6b. Le D_2 présente donc un meilleur compromis entre performance électrique et performance en fiabilité pour l'optimisation des dispositifs.



(a)



(b)

Figure B.13: Comparaison entre sans recuit, le recuit standard, HPH2 et HPD2 en dégradation HCI (a) et en NBTI (b)

B.I.4) Bilan et conclusion

Au cours de ce chapitre, nous avons montré que le recuit à haute pression permet d'avoir de meilleures performances électriques avec une réduction de la densité d'états d'interface. Le résultat varie selon qu'on considère N ou PMOS, transistors longs ou courts. Même si le H₂ donnait de meilleures performances électriques, l'étude en fiabilité a permis de montrer que le deutérium permet un meilleur compromis entre performances électriques et durée de vie du composant.

Nous avons pu voir que le PMOSFET semblait présenter plus de remote coulomb scattering que le NMOSFET. Ceci n'est pas étonnant vu les différences qui existent entre l'empilement de grille du N et PMOSFET. La couche de SiGe épitaxiée dans le cas du PMOSFET vient s'ajouter à cette différence. Le chapitre suivant étudie l'impact de cette couche de SiGe sur les performances du transistor.

CHAPITRE B.II : ARCHITECTURE CANAL A BASE DE SILICIUM-GERMANIUM

B.II.1) Pourquoi introduire une proportion de germanium dans le canal?

Pour continuer la réduction des dimensions, l'architecture du transistor a radicalement changé. Ainsi, les diélectriques de grille à haute permittivité « high- κ » sont introduits pour atteindre de faibles EOT (equivalent oxide thickness) et de faibles courants de fuite [ITRS03]. Cependant, les empilements de grille à base de high- κ mènent souvent à de fortes tensions de seuil aussi bien en N qu'en PMOSFET. Ceci est dû à la haute densité de pièges et/ou de charges fixes dans l'empilement de grille du high- κ . Pour le nœud 32 nm CMOS, du TiN et du HfSiON sont souvent utilisés comme métal et oxyde de grille respectivement. L'introduction de l'oxyde d'aluminium (Al_2O_3) comme oxyde de grille ou par dépôt au-dessus du high- κ permet de contrecarrer l'augmentation de la tension de seuil V_{TH} en décalant positivement la tension de bandes plates V_{FB} [Kita09-Charbonnier10a-Charbonnier10b]. Cependant cette méthode induit un effet indésirable appelé "roll-off" sur la courbe de V_{FB} en fonction du « EOT » [Akiyama07]. Du lantane (La) est ainsi inséré dans le NMOS pour contrer le « roll-off ». Dans le PMOS, c'est de l'aluminium (Al) qui est inséré pour contrecarrer cet effet indésirable [Baudot11]. La diffusion des atomes d'aluminium peut générer des dipôles à l'interface IL (interfacial layer)/high- κ , dégradant la tension de seuil. En plus de cela, les charges fixes, les pièges et/ou les dipôles générés dans une architecture si complexe peuvent diminuer les performances du transistor comme la mobilité des porteurs dans le canal.

Les matériaux à haute mobilité comme le silicium-germanium (SiGe) ont été introduit dans le PMOS du nœud 28 nm pour compenser de tels décalages de V_{TH} et de dégradations de mobilité causés par l'utilisation de gate stack à base de high- κ [Takagi07-Skotnicki08].

Nous avons étudié les effets de l'incorporation du germanium (Ge) dans le canal du PMOS à base de high- κ [Diouf13b], avec de l'Al pris en sandwich par du TiN (figure B.14). La couche de SiGe est épitaxié (figure B.15).

Une référence Si assujettie aux mêmes étapes de procédé de fabrication sans germanium est aussi utilisée.

Des simulations Poisson-Schrödinger sont utilisées pour mieux interpréter les résultats expérimentaux.

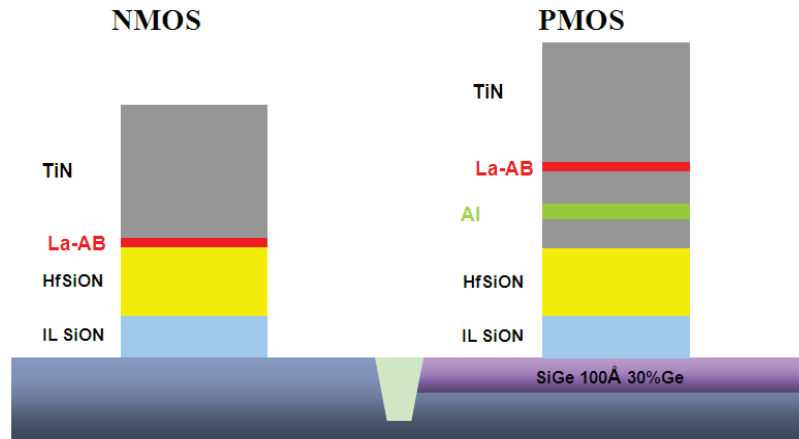


Figure B.14: Empilement du stack de grille des N&PMOS étudiés dans ce chapitre

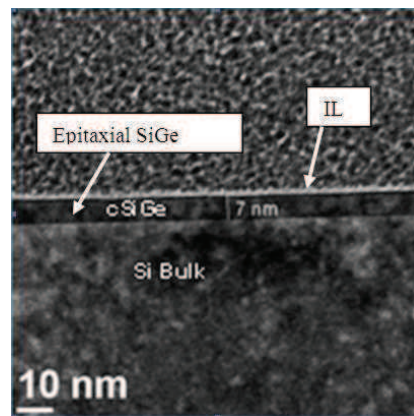


Figure B.15: Image TEM illustrant la couche de SiGe épitaxiée

B.II.2) Canal silicium-germanium

B.II.2.a) Simulations

Les simulations sont basées sur des solutions auto-consistantes à une dimension des équations de Poisson et de Schrödinger. Dans ce modèle, la densité de charge pour l'hétérostructure SiGe/Si est calculée en utilisant le modèle k.p à 6-bandes [Soussou10]. La tension de seuil V_{TH} est extraite à partir de 40% du maximum de la capacité (correspondant au maximum de la dérivée de la capacité, mais les incertitudes et bruits de mesures issus de la dérivée sont évités). Tous les travaux de sortie sont ajustés sur celui de la référence silicium. Cela permet de n'inclure que les effets du germanium lorsque le pourcentage de germanium augmente dans le canal. En fin le calcul de la structure de bande du SiGe prend en compte les effets de la contrainte. Pour plus de détails à propos de la simulation Poisson-Schrödinger, se référer à [Soussou10].

Lorsqu'on change le pourcentage du germanium de 0 à 30%, les valeurs expérimentales montrent que V_{TH} est réduit de 300mV (figure B.16). Il existe cependant un décalage de 150 mV entre la réduction de V_{TH} obtenue en mesures expérimentales et la réduction prédite par les simulations Poisson- Schrödinger (figure B.16).

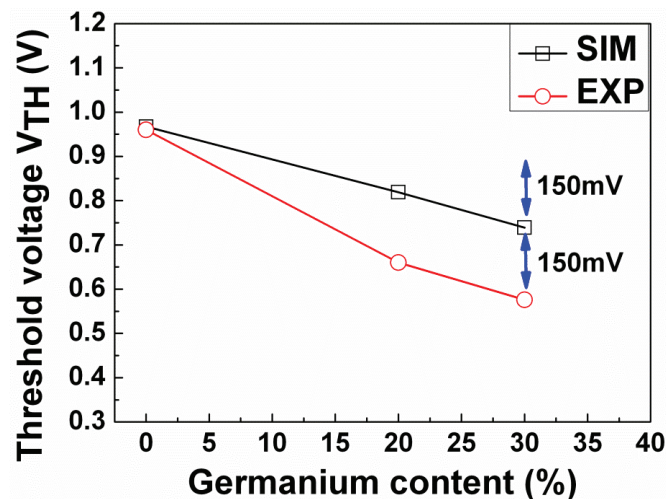


Figure B.16: V_{TH} simulé et V_{TH} expérimental en fonction du pourcentage du germanium

Pour reproduire par simulation la capacité totale du dispositif avec du $\text{Si}_{0.7}\text{Ge}_{0.3}$, il a fallu ajuster le travail de sortie du métal de 0.15 eV, confirmant le shift supplémentaire obtenu en V_{TH} .

La suite de ce chapitre consistera à chercher à comprendre l'origine de ces 150 mV de shift inattendue. Des défauts ou charges dans le high- κ et/ou à l'interface IL/canal peuvent être à l'origine de ce décalage supplémentaire et inattendu, ainsi qu'une différence de dopage entre les deux types de dispositifs.

B.II.2.b) Analyse à température ambiante

L'augmentation du courant de drain est notée dans les transistors SiGe comparé au transistor Si (figure B.17)

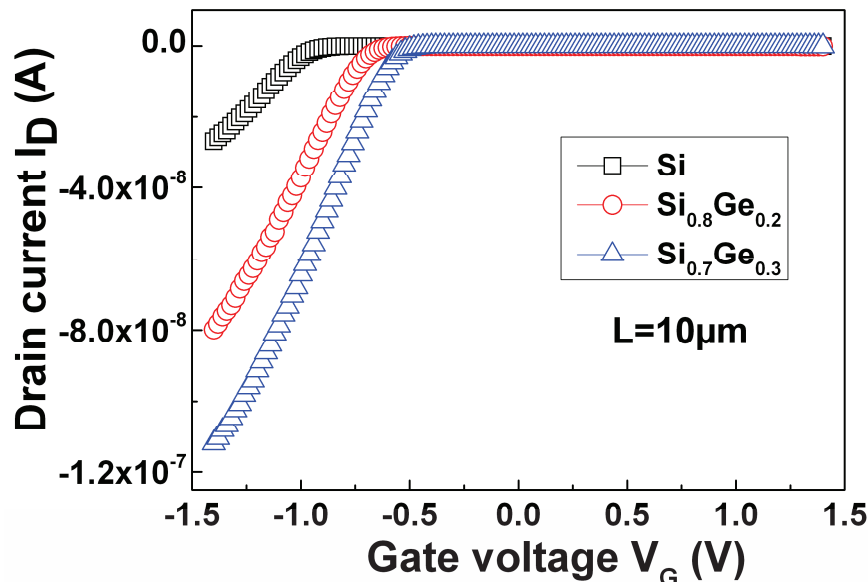


Figure B.17: Représentation du courant de drain en fonction de la tension de grille pour Si, $\text{Si}_{0.8}\text{Ge}_{0.2}$ et $\text{Si}_{0.7}\text{Ge}_{0.3}$

La mobilité à faible charge μ_0 (B.8) est extraite en utilisant la fonction Y permettant d'éliminer l'influence des résistances source-drain R_{SD} [Ghibaudo88]. La mobilité effective μ_{eff} (B.9) est obtenue pour les transistors longs par la technique du split-CV [Koomen73-Sodini82].

$$\mu_0 = \frac{\beta L_{eff}}{C_{ox} W_{eff}} \quad (B.8)$$

$$\mu_{eff} = \frac{L_{eff} \cdot I_D}{W_{eff} \cdot Q_i \cdot V_D} \quad (B.9)$$

Une comparaison entre les dispositifs Si et Si_{0.7}Ge_{0.3} montre sur la **figure B.18a** une augmentation de 160% de la mobilité à faible charge μ_0 . Sur la **figure B.18b**, la mobilité effective est améliorée de 120% environ entre Si et Si_{0.7}Ge_{0.3}. Une partie de cette amélioration du transport électronique qui se répercute sur le courant de drain (**figure B.17**), vient de la plus faible masse effective du germanium comparé au silicium. Cependant le gain en mobilité est plus important que ce qui est montré dans la littérature [Alatise09-Krishnan11]. La mobilité à faible charge prend en compte des mécanismes de réduction de mobilité comme le remote Coulomb scattering (RCS) et le phonon-scattering (PS). Elle est cependant moins sensible à la diffusion par les rugosités de surface (SR) qui apparait à fort V_G ou à fort champ effectif. La mobilité effective, quant à elle, prend en compte tous les mécanismes de réductions de mobilité.

L'impact de la longueur effective ne peut être mis en cause. En effet, l'erreur possible sur la détermination de la longueur effective n'a d'impact que sur les transistors courts et on voit clairement sur la **figure B.19** que le décalage de V_{TH} est le même pour les transistors longs et courts. De plus, le **tableau B.1** montre la faible variation de la longueur effective entre Si et SiGe.

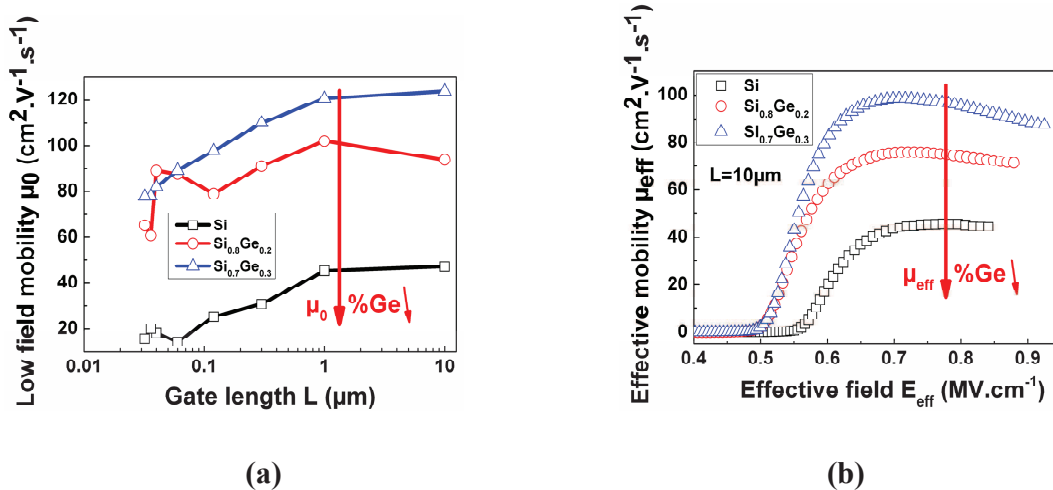


Figure B.18: Illustration de la mobilité à faible charge (a) et de la mobilité effective (b) pour Si, Si_{0.8}Ge_{0.2} et Si_{0.7}Ge_{0.3}

$L_{\text{drawn}} (\mu\text{m})$	1	0.3	0.12	0.06	0.04
$L_{\text{eff Si}} (\mu\text{m})$	0.906	0.278	0.110	0.058	0.041
$L_{\text{eff Si}_{0.7}\text{Ge}_{0.3}} (\mu\text{m})$	0.907	0.281	0.112	0.058	0.041

Tableau B.1: Description de l'impact de la longueur effective entre Si et $\text{Si}_{0.7}\text{Ge}_{0.3}$

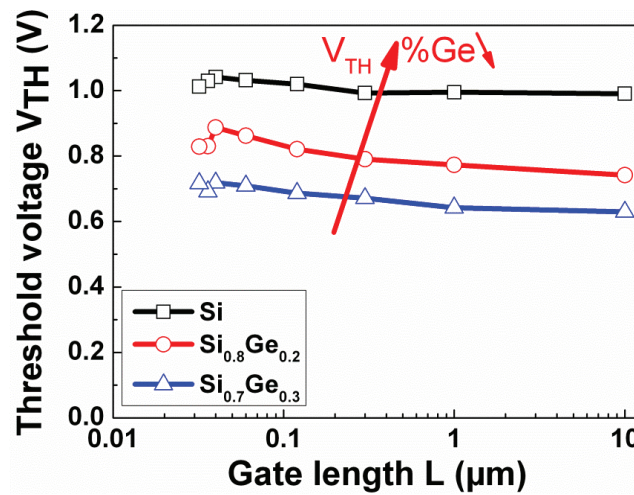


Figure B.19: Illustration de la tension de seuil en fonction de la longueur pour Si, $\text{Si}_{0.8}\text{Ge}_{0.2}$ et $\text{Si}_{0.7}\text{Ge}_{0.3}$

En considérant les observations sur la mobilité à faible charge et sur la mobilité effective, il est fort probable que l'amélioration supplémentaire de la mobilité vienne des mécanismes de réduction de mobilité évoqués plus haut. Le RCS et/ou le PS restent des pistes plus que crédibles. Il est important de noter que l'amélioration inattendue de mobilité peut être interprétée par une faiblesse inattendue de la mobilité de nos dispositifs silicium. La figure B.20 illustre la localisation des mécanismes de scattering susceptible d'influencer négativement la mobilité dans un MOSFET.

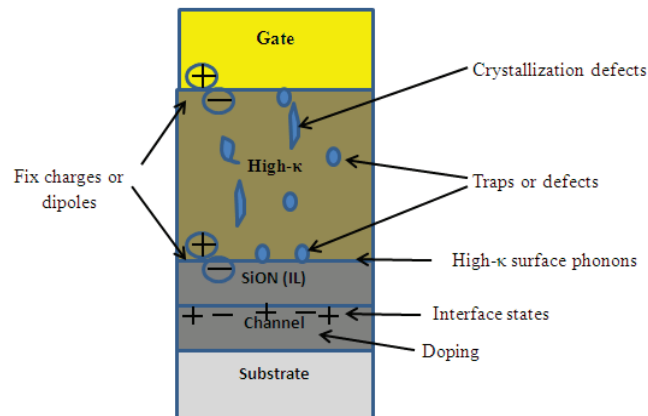


Figure B.20: Illustration de la localisation des différents mécanismes susceptibles d'influencer négativement la mobilité

Sur la **figure B.20** on peut distinguer l'impact des facteurs suivants :

- Le dopage, via son influence sur la charge de désertion et donc le champ transverse, ainsi que sur le Coulomb scattering.
- Les états d'interface qui peuvent avoir un impact sur la mobilité lorsqu'ils sont importants.
- Les phonons qui sont inhérents au matériau et qui sont généralement un facteur prédominant parmi les facteurs de limitation de la mobilité.
- Les dipôles ou charges fixes qui se forment à l'interface $\text{SiO}_2/\text{high-}\kappa$ ou dans le volume du high- κ très proche de l'interface $\text{SiO}_2/\text{high-}\kappa$.
- Les lacunes d'oxygène dans le diélectrique qui peuvent jouer un rôle dans le comportement électronique du transistor.
- Les pièges et les défauts de cristallisation qui sont localisés dans le volume du high- κ et/ou aux interfaces.

Les mesures de capacité grille canal (CGC) et grille bulk (CGB) n'ont pas pu apporter plus d'informations compte tenu de la différence de densité d'états et de «dark space» entre Si et Ge [Wu12].

❖ Le dopage :

Une première piste consistera à explorer une possible différence de dopage entre nos dispositifs Si et SiGe. Pour cela, la fonction de Maserjian est utilisée [Maserjian94]. Comme illustré dans l'équation (B.10), elle est exclusivement basée sur des mesures capacitives.

$$Y_{mas} = \frac{1}{C^3} \cdot \frac{\partial C}{\partial V_G} = -\frac{1}{2} \cdot \frac{\partial(\frac{1}{C^2})}{\partial V_G} \quad (\text{B.10})$$

Cette fonction Y_{mas} donne accès à la fois à la tension de bande plate V_{FB} et au dopage N_A . Le tracé $Y_{mas}(V_G)$ donne lieu à un plateau appelé Y_{masmin} .

Dans le cas d'une statistique Maxwell-Boltzmann, $V_G = V_{FB}$ quand $Y_{V_{FB}} = \frac{Y_{masmin}}{3}$ (B.11) avec

$$Y_{masmin} = -\frac{1}{q \cdot \epsilon \cdot N_A} \quad (\text{B.12}).$$

La figure B.21 montre un décalage de V_{FB} entre Si et Si_{0.7}Ge_{0.3}. Il est important de noter que ce décalage est du même ordre de grandeur que celui du V_{TH} observé lorsque la proportion de Ge passe de 0 à 30%, c'est-à-dire de l'ordre de 300 mV.

Le plateau Y_{masmin} qui symbolise le niveau de dopage reste quasiment le même entre Si et Si_{0.7}Ge_{0.3}. Il n'est donc pas possible d'expliquer la faible mobilité de nos dispositifs Si (ou l'amélioration inattendue observée entre Si et SiGe) par le dopage.

Il faut noter que l'incertitude sur la valeur de ϵ qui varie de Si à Si_{0.7}Ge_{0.3} n'est pas assez forte pour impacter la conclusion.

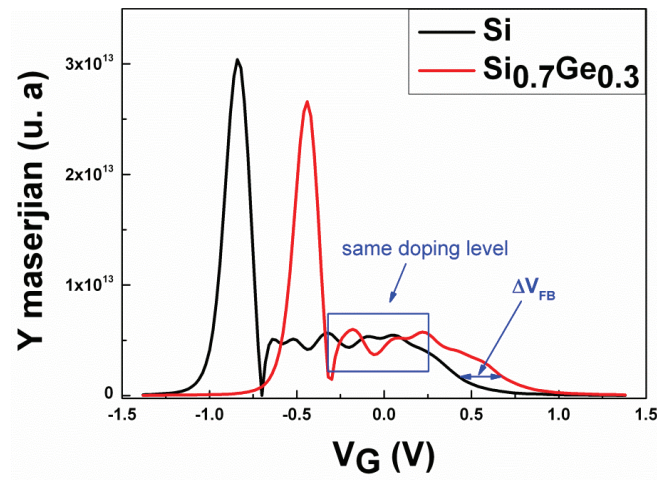


Figure B.21: La fonction de Maserjian en fonction de la tension de grille: extraction de V_{FB} et du niveau de dopage

❖ Les états d'interface

Une technique pour caractériser les états d'interface est le pompage de charge. Cette technique est décrite dans le chapitre I de cette partie. Sur la [figure B.22](#), la densité d'états d'interface N_{IT} est extraite pour Si et $Si_{0.7}Ge_{0.3}$. On peut apercevoir que la différence de N_{IT} est faible entre Si et $Si_{0.7}Ge_{0.3}$, excepté à basse fréquence où la précision de la méthode diminue. Il s'avère donc impossible d'expliquer le scattering supplémentaire dans nos dispositifs Si par les états d'interface.

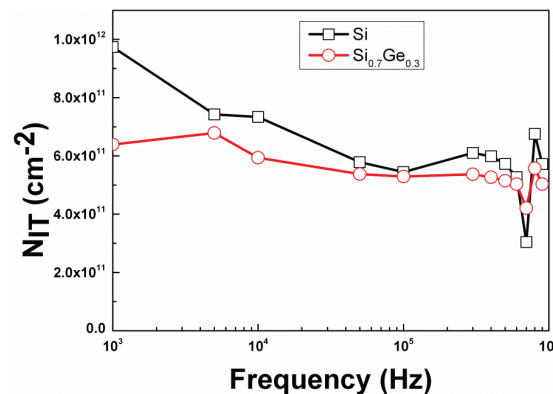


Figure B.22: Illustration de la densité d'états d'interface en fonction de la fréquence pour Si et $Si_{0.7}Ge_{0.3}$

Comme nous venons de le voir, l'analyse à température ambiante ne permet pas de tirer des conclusions claires. Nous allons passer à l'analyse à basses températures afin d'en apprendre davantage sur le rôle du germanium dans nos PMOSFETs bulk.

B.II.2.c) Analyse à basse température

Pour faire la lumière sur le mécanisme de réduction de mobilité (RSC ou PS) responsable de la faible mobilité dans nos dispositifs silicium, des mesures à basses températures sont menées. La loi empirique déjà décrite au chapitre I de cette partie est à nouveau utilisée [Cros06]. Une loi type Mathiessen décrit la mobilité à faible charge en fonction de la température T (B.13).

$$\frac{1}{\mu_0}(T) = \frac{1}{\mu_{0PH}(T)} + \frac{1}{\mu_{0C}(T)} \quad (B.13)$$

Dans l'équation B.13, μ_0 est la mobilité à faible charge, μ_{0PH} est la contribution de mobilité reliée au phonon-scattering, μ_{0C} est celle reliée au Coulomb scattering.

Ces contributions de mobilité peuvent être décrites comme suit:

$$\mu_{0PH}(T) = \frac{300 \cdot \mu_{PH}}{T} \quad (B.14) \quad , \quad \mu_{0C}(T) = \frac{T \cdot \mu_C}{300} \quad (B.15), \mu_{PH} \text{ et } \mu_C \text{ sont des constantes.}$$

En fittant l'équation B.13 (combinée avec les équations B.14 et B.15) avec les valeurs de μ_0 expérimentales comme sur la figure B.23, on peut déterminer la valeur des constantes μ_{PH} et μ_C .

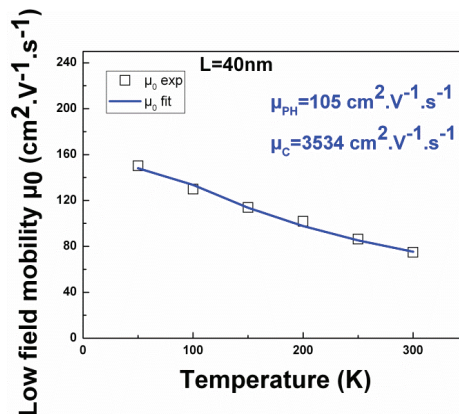


Figure B.23: Mobilité à faible charge expérimentale et théorique obtenue par l'équation B.6

Les dispositifs Si et SiGe sont ainsi séparément étudiés afin de clarifier les différences.

❖ pMOSFETs Si pur

Le courant de drain à $V_{GT} = V_G - V_{TH}$ constant et la mobilité à faible charge μ_0 sont tracés sur les figures B.24a et B.24b respectivement.

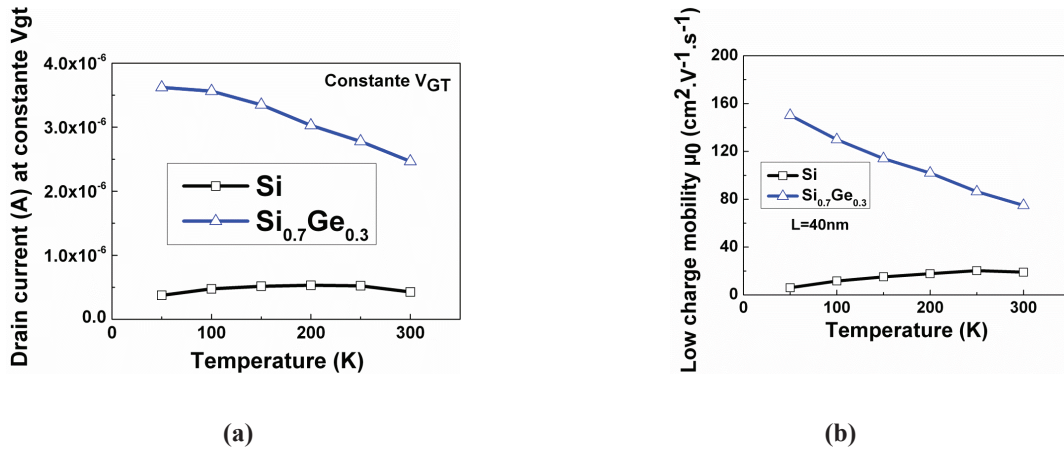


Figure B.24: Variation du courant de drain à V_{GT} constant en fonction de la température pour Si et Si_{0.7}Ge_{0.3} (a) ; mobilité à faible charge en fonction de la température pour Si et Si_{0.7}Ge_{0.3} (b)

On remarque que pour le silicium, le courant à V_{GT} constant est quasiment stable en température, voire décroissant à très basse température. La mobilité à faible charge montre un comportement similaire et est décroissante lorsque la température diminue. Ce comportement de la mobilité est caractéristique du RCS, qui est linéairement croissant avec la température. Pour confirmer ces conclusions sur la mobilité, on peut tracer les différentes contributions ($\mu_{0PH}(T)$ et $\mu_{0C}(T)$) en fonction de la température (figure B.25) comme donné dans les équations 7 et 8.

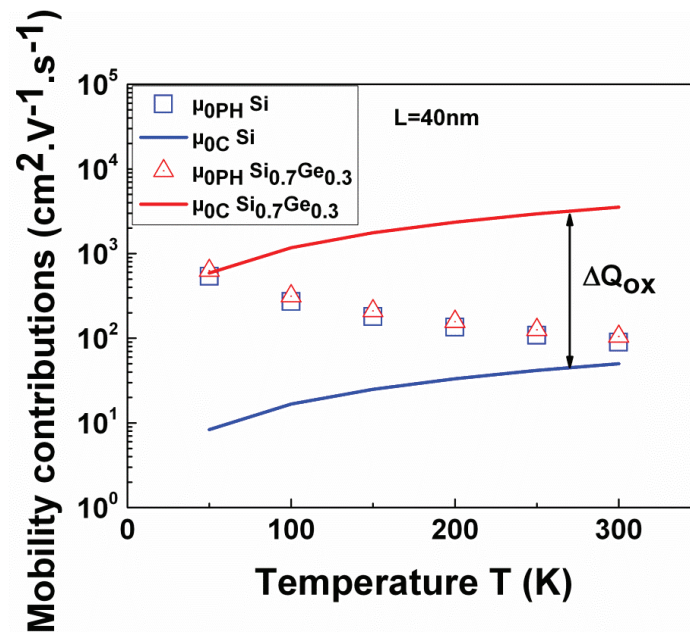


Figure B.25: Variation des différentes contributions de mobilité en fonction de la température pour Si et $Si_{0.7}Ge_{0.3}$

Lorsqu'on reporte ces valeurs de $\mu_{0PH}(T)$ et $\mu_{0C}(T)$ dans l'équation B.6 pour Si, on peut voir que la mobilité à faible charge μ_0 est dominée par le RCS.

❖ pMOSFETs $Si_{0.7}Ge_{0.3}$ pur

Si on fait la même analyse que le Si sur la figure B.24a & B.24b, on remarque que le SiGe montre un comportement opposé à celui du Si. Le courant à V_{GT} constant croît lorsque la température diminue. La mobilité μ_0 du SiGe montre aussi un comportement différent à celui du Si. Contrairement à celle du silicium, μ_0 de SiGe varie comme T^{-1} avec la température. Ce comportement de la mobilité est caractéristique d'une mobilité gouvernée par le phonon-scattering. Cela se confirme avec la séparation des contributions sur la figure B.25. Si on reporte les contributions de mobilité dans l'équation B.6 pour SiGe, on se retrouve avec une mobilité dominée par le PS.

Une analyse commune entre Si et SiGe, montre que le niveau de phonon est identique dans les deux dispositifs. Le PS n'est donc pas la cause de la faible mobilité dans nos transistors silicium. Ces conclusions sont en désaccord avec ce qui a été précédemment publié dans la littérature [Haartman06]. Cette surprenante et faible valeur de mobilité peut cependant être reliée au

processus de fabrication. La présence de charge dans le stack de grille est suspectée. On peut calculer la quantité de charges ΔQ_{ox} responsable de la mobilité additionnelle (9) entre Si et $Si_{0.7}Ge_{0.3}$ via (B.16).

$$\Delta Q_{ox} = \frac{\mu_{add}^{-1}}{\alpha \cdot q} \quad (B.16); \quad \frac{1}{\mu_{add}} = \frac{1}{\mu_{Si}} - \frac{1}{\mu_{Si_{0.7}Ge_{0.3}}} \quad (B.17); \quad \frac{1}{\mu_{add}} = \frac{1}{\mu_{0CSi}} - \frac{1}{\mu_{0CSi_{0.7}Ge_{0.3}}}$$

(B.18)

α est le coefficient de remote Coulomb scattering. Il a été évalué à 4100 et 73000 V.s.C⁻¹ pour les électrons et les trous respectivement [Bennamane13], q est la charge de l'électron.

En utilisant (B.13) et du fait que le niveau de phonon est le même entre Si et $Si_{0.7}Ge_{0.3}$, la mobilité additionnelle est donnée par (B.17 & B.18). On peut ainsi obtenir la différence de charge à l'origine de la mobilité supplémentaire entre Si et $Si_{0.7}Ge_{0.3}$ en fonction de la température (figure B.26).

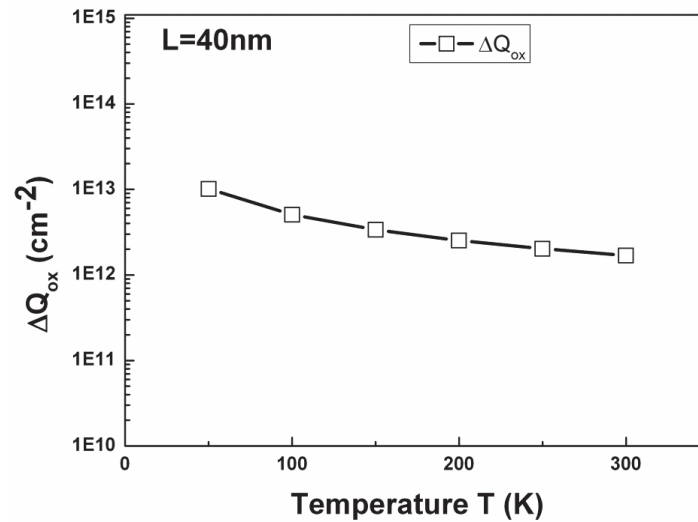


Figure B.26: Illustration de la quantité de charge responsable de la mobilité additionnelle entre Si et $Si_{0.7}Ge_{0.3}$

Il semble donc que l'introduction du germanium diminue la quantité de charges contenues dans l'empilement de grille. Une interaction physico-chimique entre le Ge et le stack de grille est suspectée.

Cette hypothèse d'interaction du germanium avec le stack de grille de type high- κ n'est pas tout à fait nouvelle. En effet, dans [Tsuchiya10], une implantation ionique de Ge dans le canal décale le V_{TH} et le V_{FB} sans améliorer la mobilité, ne contribuant donc pas à la conduction (figure B.27a & B.27b).

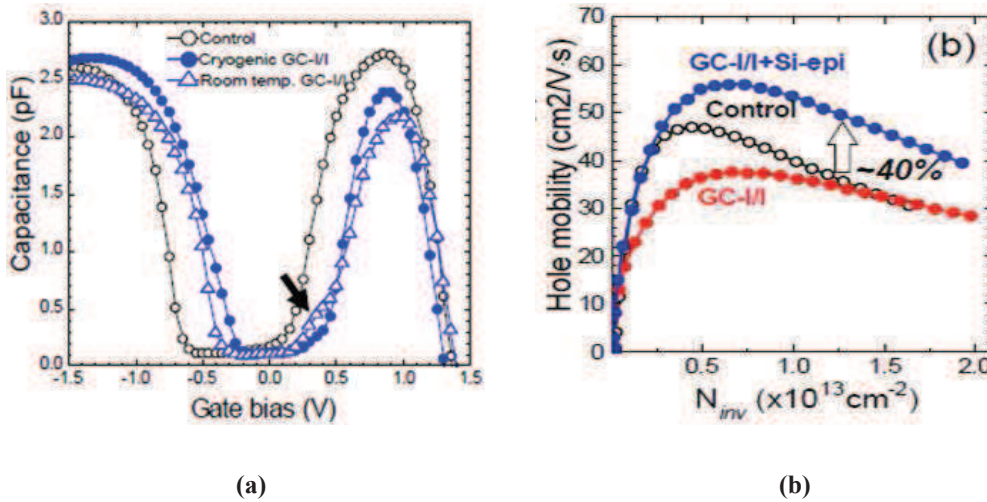


Figure B.27: Capacité totale en fonction de la tension de grille avant et après implantations ioniques de Ge dans le canal (a); mobilité effective en fonction de la charge d'inversion avant et après implantations ioniques de Ge dans le canal et après dépôt de silicium épitaxié (b) [Tsuchiya10]

Pour améliorer la mobilité, il a fallu déposer une couche de silicium épitaxiée au-dessus du canal après implantation ionique. Tout s'est passé comme s'il y avait du SiGe épitaxié : La mobilité est améliorée (figure B.27b). Tous ces effets sont expliqués dans [Tsuchiya10] par une interaction chimique entre le Ge et le stack de grille.

Pour en apprendre plus sur cette interaction entre empilement de grille et canal, des mesures de bruit basse fréquence ont été faites.

B.II.2.d) Mesures de bruit basse fréquence

- ❖ Généralités sur le bruit basse fréquence

Le bruit est défini comme étant la fluctuation d'une grandeur autour de sa valeur moyenne dans le domaine temporel (figure B.28). On parle de bruit de fond des dispositifs lorsque les fluctuations sont intrinsèques au composant et non pas générées par une source extérieure.

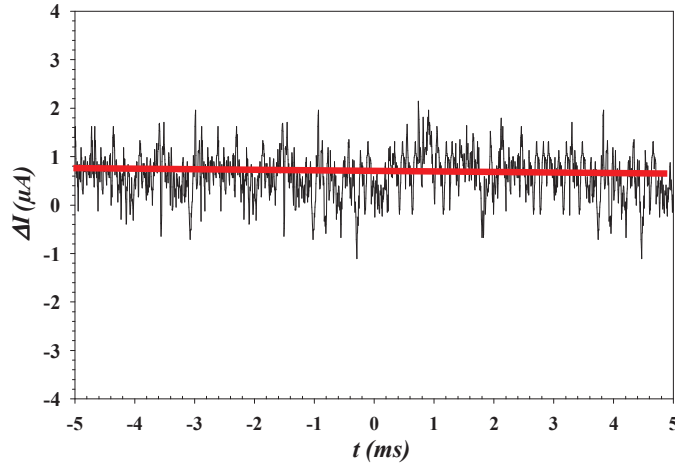


Figure B.28: Illustration de la fluctuation d'un paramètre autour de sa valeur attendue dans le domaine temporel

Pour séparer les sources du bruit, on l'étudie dans le domaine fréquentiel à travers des mesures de densité spectrale de puissance des fluctuations S_I . En effet S_I et la variance σ du signal de bruit $i(t)$ sont reliés par :

$$\sigma^2 = \int_{-\infty}^{+\infty} S_I(f, T) df \quad (\text{B.19})$$

$$\sigma = \frac{1}{T} \cdot \int_0^T i^2(t) dt \quad (\text{B.20}),$$

T est la période du signal $i(t)$ et f la fréquence.

Sur la figure B.29, S_I est tracé en fonction de f pour un MOS. On peut y faire la différence entre le bruit en 1/f et le bruit thermique.

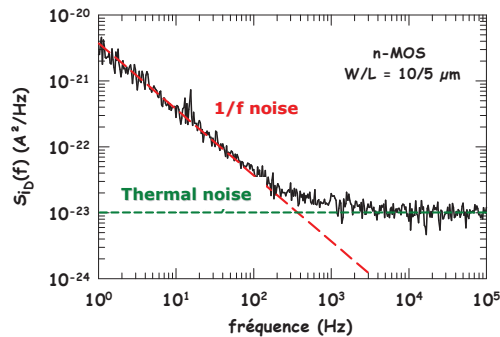


Figure B.29: Densité spectrale de puissance en fonction de la fréquence: illustration du bruit 1/f et du bruit thermique

Dans les transistors MOSFETs, on s'attend à ce que le bruit de fond à, très basse fréquence soit de type 1/f. Le bruit est aussi associé à la fluctuation de la conductivité $\Delta\sigma_c$:

$$\sigma_c = q \cdot \mu \cdot N \quad (\text{B.21})$$

N est densité de porteurs dans le canal et μ la mobilité. Une fluctuation de la conductivité sera donc associée à une fluctuation du nombre de porteurs ΔN et de la mobilité $\Delta\mu$. Une fluctuation du nombre de porteurs peut être engendrée par la présence de défauts, de charges et/ou de pièges (figure B.30). Quand à celle de la mobilité elle peut être engendrée par des interactions coulombiennes par exemple.

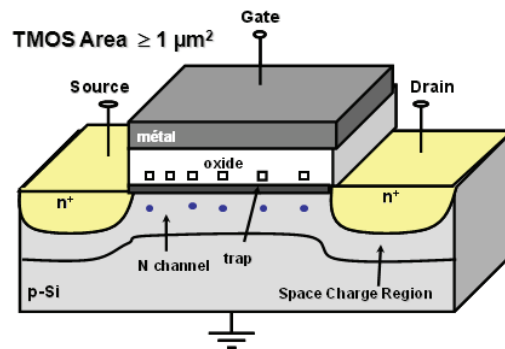


Figure B.30: Illustration de l'interaction entre le nombre de porteurs N et les pièges dans l'oxyde sources de bruit

❖ Etudes du bruit basses fréquences dans nos dispositifs : confrontation entre Si et SiGe

La densité spectrale de puissance de courant S_{ID} normalisée avec le carré du courant de drain I_D est tracée en fonction du courant de drain. La figure B.31 montre une plus grande densité spectrale de puissance pour le Si comparé au SiGe. Cela veut dire qu'il existe plus de bruit dans nos transistors pMOSFET Si pur. Le bruit dans les MOSFETs peut être bien corrélé à la fluctuation du nombre de porteurs et à la fluctuation de la mobilité comme illustré sur les équations B.22 & B.23 [Boutchacha99].

$$\frac{S_I}{I_D^2} = \left(\frac{g_m}{I_D} \right)^2 \left(1 + \alpha_{SC} \cdot \mu_{eff} \cdot C_{ox} \cdot \frac{I_D}{g_m} \right)^2 \cdot S_{VFB} \quad (\text{B.22}); \quad S_{VFB} = \frac{q^2 \cdot k \cdot T \cdot \lambda \cdot N_t}{W \cdot L \cdot C_{ox}^2 \cdot f} \quad (\text{B.23})$$

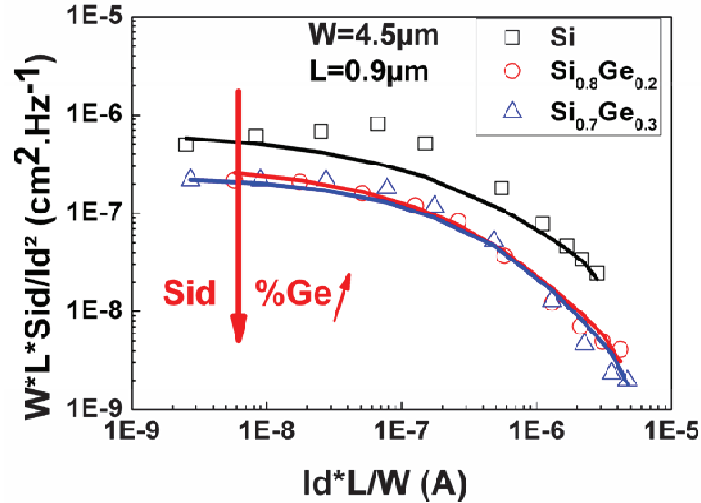


Figure B.31: Bruit basses fréquences normalisée avec le carré du courant en fonction du courant de drain pour Si, Si_{0.8}Ge_{0.2} et Si_{0.7}Ge_{0.3}

Dans (B.22) le paramètre α_{sc} est le coefficient de corrélation entre fluctuation de mobilité et bruit. Il peut arriver que le terme de l'équation contenant α_{sc} ne soit pas nécessaire pour bien ajuster S_{ID} avec le modèle. Dans notre cas, il a fallu prendre en compte ce terme de l'équation pour bien fitter S_{ID} . Les valeurs de α_{sc} sont synthétisées dans le [tableau B.2](#) pour Si, Si_{0.8}Ge_{0.2} et Si_{0.7}Ge_{0.3}. On peut s'apercevoir que le Si a le plus grand paramètre α_{sc} , permettant de mettre en évidence une plus grande contribution de mobilité corrélée. Il est à noter que cette plus grande fluctuation de mobilité est associée à une plus grande variabilité statique de la mobilité ou plus précisément du paramètre de gain β (proportionnel à μ_0) pour les dispositifs Si comparés aux dispositifs SiGe ([figure B.32](#)) [Rahhal13].

	Si	Si _{0.8} Ge _{0.2}	Si _{0.7} Ge _{0.3}
$\alpha_{sc}(\text{s.F}^{-1})$	70000	20000	12000
$N_{ST}(\text{cm}^{-2})$	$2.17 \cdot 10^{11}$	$1.5 \cdot 10^{11}$	$1.25 \cdot 10^{11}$

Tableau B.2: Coefficient de corrélation entre mobilité et bruit et densité de pièges

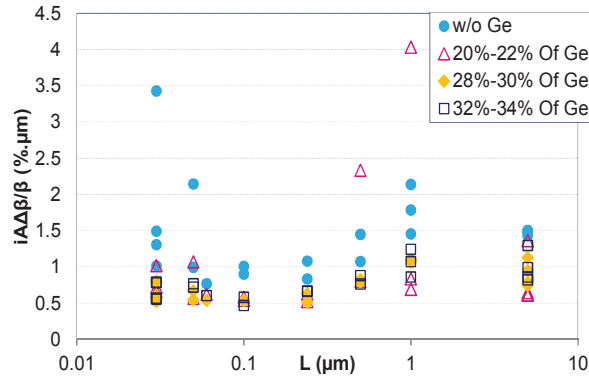


Figure B.32: Variabilité de β pour Si et SiGe [Rahhal13]

Dans (B.22) S_{VFB} est représentatif du plateau obtenu en figure B.31 pour les faibles I_D . S_{VFB} est relié à la densité de pièges dans l'oxyde N_{ST} par l'équation B.23. W et L sont la largeur et la longueur de grille respectivement, C_{ox} est la capacité de l'oxyde, T est la température, k est la constante de Boltzmann et λ est la longueur d'atténuation de tunnel dans l'oxyde de grille. La figure B.33 montre les densités de pièges extraites pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$. Sans surprise comme synthétiser sur le tableau B.2, N_{ST} est quasiment deux fois plus grand pour le Si comparé au $Si_{0.7}Ge_{0.3}$.

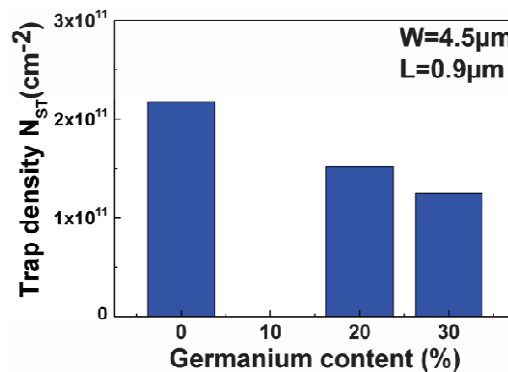


Figure B.33: Densité de piège dans l'oxyde pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$

B.II.3) Discussion et conclusion

L'origine d'un décalage supplémentaire de tension de seuil inattendu de 150 mV dans les dispositifs SiGe a été investiguée. Des simulations Poisson-Schrödinger ont montré la nécessité

d'un ajustement de 150 mV du travail de sortie du métal du transistor $\text{Si}_{0.7}\text{Ge}_{0.3}$ pour reproduire la capacité totale mesurée expérimentalement. Il faut signaler que ces simulations prennent en compte le décalage de la bande de valence du Ge par rapport au Si, l'effet de la contrainte amené par l'introduction du Ge et l'effet de la faible masse effective du Ge.

Les mesures de mobilité (faibles charges et effectives) montrent une plus grande amélioration comparée à la littérature pour les dispositifs SiGe. Ce phénomène n'a pu être expliqué ni par le dopage ni par les états d'interface. Les mesures à basses températures ont permis de révéler un taux important de remote Coulomb scattering dans les dispositifs Si dû à une plus grande quantité de charges dans son empilement de grille. Les mesures de bruit ont aussi révélé une plus grande contribution de fluctuations de mobilité et une plus grande densité de piège dans le pMOSFETs Si.

Toutes ces conclusions laissent penser que le PMOSFET étudié dans ce chapitre présente un high- κ avec des charges et/ou pièges qui agissent négativement sur la mobilité. On suspecte cependant une interaction physico-chimique entre le Ge et le stack de grille. Cette interaction tend à améliorer la mobilité en plus de l'amélioration due à la plus faible masse effective du Ge, en réduisant la quantité de charges.

En utilisant des IL (SiO_2) d'épaisseurs différentes, on modifie la position du high- κ dans le stack de grille. Dans ce cas, l'impact du high- κ peut être étudié. Dans le chapitre suivant, des transistors GO1 (SiO_2 standard) et GO2 (SiO_2 plus épais) seront comparés afin de discriminer le rôle du high- κ sur les performances du PMOSFET 28 nm bulk.

CHAPITRE B.III : L'OXYDE DE GRILLE : LE HIGH-κ ET SON IMPACT

B.III.1) Pourquoi l'introduction du high-κ ?

L'empilement ou « stack » de grille a été historiquement constitué de poly-silicium surdopé comme matériau de grille et de dioxyde de silicium comme oxyde de grille. Il a été difficile de maintenir cette architecture pour contrecarrer les effets de la miniaturisation. Aujourd'hui le stack de grille est constitué de multiples couches de métaux et d'oxydes.

L'oxyde de grille initialement utilisé pour la fabrication des transistors MOS a été le dioxyde de silicium SiO₂. L'utilisation du SiO₂ a atteint sa limite avec la réduction continue des dimensions. Les épaisseurs d'oxyde atteignent les alentours de 1.5 nm et cela engendre des problèmes tels que l'effet tunnel des électrons à travers ce même oxyde. Il s'en suit donc une augmentation inévitable et non souhaitable du courant de fuite (I_{off}) dans les dispositifs, et donc de la consommation électrique à l'état off (lorsque le dispositif n'est pas en phase d'utilisation) :

$$P_{off} = V_{DD} \cdot I_{off} \quad (\text{B.24})$$

P_{off} est la puissance consommée passivement et I_{off} est le courant débité par le transistor au repos (par l'effet tunnel évoqué plus haut par exemple).

C'est ainsi qu'une solution technologique a été proposée pour parer à ces difficultés. Cela consiste à utiliser un autre type de diélectrique de grille avec une constante diélectrique plus grande que le celle du SiO₂. Ces oxydes appelés high-κ (le plus utilisé est le HfO₂) permettent l'utilisation d'épaisseurs plus grandes pour un même champ électrique, tout en évitant les fuites de grille [Taur97-Liu01].

B.III.2) Les inconvénients du high-κ

L'inconvénient majeur du high-κ est la réduction de mobilité dans les dispositifs qu'il compose. Les mécanismes susceptibles d'affecter la mobilité dans les dispositifs à base de high-κ sont listés sur la figure B.34. Dans [Saito03], la prise en considération des charges fixes et de la cristallisation dans le high-κ paraît suffisante pour expliquer la dégradation de mobilité par le

remote Coulomb scattering (RCS). Selon [Cassé06-Weber04], l'origine de cette réduction de mobilité proviendrait de la densité de charge élevée à l'interface $\text{HfO}_2/\text{SiO}_2$ pour les transistors dont la grille est à base de nitrure de titane (TiN).

En outre, les « phonons optiques mous » (soft phonon) ont été considérés comme jouant un rôle non négligeable dans la réduction de mobilité dans les transistors à base de high- κ [Fischetti01-Zhu04]. Cette hypothèse a été affaiblie par la suite. Le rôle des phonons optique de surface a été minimisé comparé à celui joué par le RCS dû aux charges ou dipôles aux interfaces [Ota107-Toniutti08].

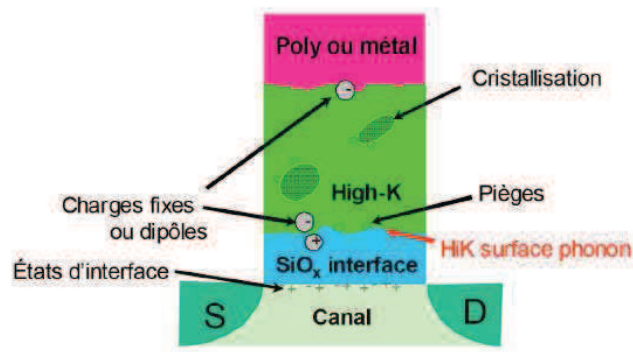


Figure B.34: Illustration des différents mécanismes qui peuvent dégrader la mobilité dans un transistor ayant un empilement de grille high-k métal, d'après [Weber05]

D'après [Kita09], l'origine de ces dipôles qui affectent aussi la tension de bandes plates VFB est due à la différence de densité d'atome d'oxygènes qui existe à l'interface high- κ/SiO_2 . Une manière d'étudier l'impact du high- κ sur la performance des transistors est de disposer de deux épaisseurs d'oxyde différentes. D'où l'utilisation des transistors GO1 (faible épaisseur de SiO_2) et GO2 (grande épaisseur de SiO_2) comme structures de test pour la mobilité.

B.III.3) Les transistors de type GO1 et GO2

Pour éviter la formation d'une couche de SiO_2 non contrôlée au moment de déposer le high- κ (HfO_2) sur le silicium, on réalise d'abord un dépôt d'une fine couche d'oxyde de silicium nitruré (SiON) contrôlée. Cette couche plus communément appelée IL pour « interfacial layer » en anglais, forme une meilleure interface avec le high- κ . Dans le cas d'un transistor type GO2, une couche de SiO_2 est premièrement déposée avant la formation de l'IL. Grâce à leur oxyde plus épais ($\text{SiO}_2 + \text{SiON}$), les transistors GO2 sont capables de supporter de grandes tensions sans atteindre la tension de claquage contrairement aux transistors type GO1. Ainsi, ils sont principalement utilisés en entrée et sortie des circuits.

En regardant les figures B.35a & B.35b, on peut s'apercevoir que le high- κ est plus distant du canal dans le cas du transistor GO2 (figure B.35b). On s'attend donc à ce que l'impact du high- κ ou celui de son interface avec l'IL soit peu influent sur le transport des porteurs dans le canal.

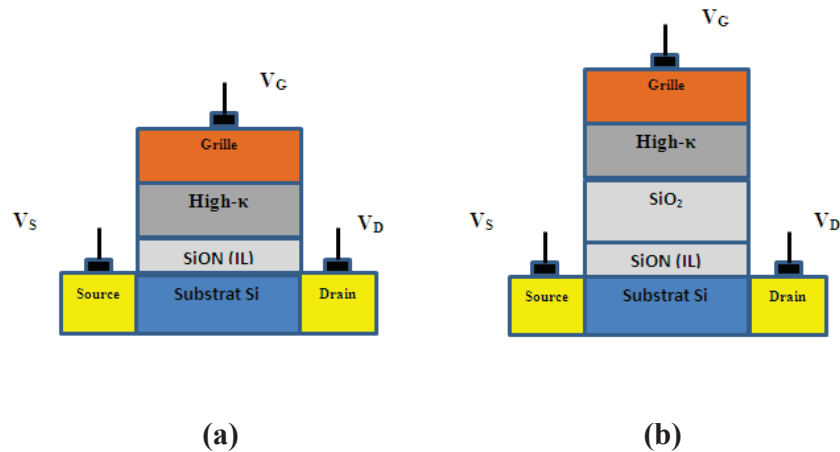


Figure B.35: Illustration des transistors type GO1 (a) et GO2 (b)

B.III.4) Etude comparative des transistors GO1 et GO2

Une étude comparative de la mobilité a été réalisée sur les transistors GO1 et GO2. Comme attendu, les dispositifs N & PMOS montrent une augmentation de la mobilité à faible charge des transistors GO2 comparés aux GO1. Cette amélioration est cependant plus importante à très basses températures (figures B.36a & B.36b). Les dispositifs FDSOI montrent le même comportement comme illustré sur les figures B.37a & B.37b.

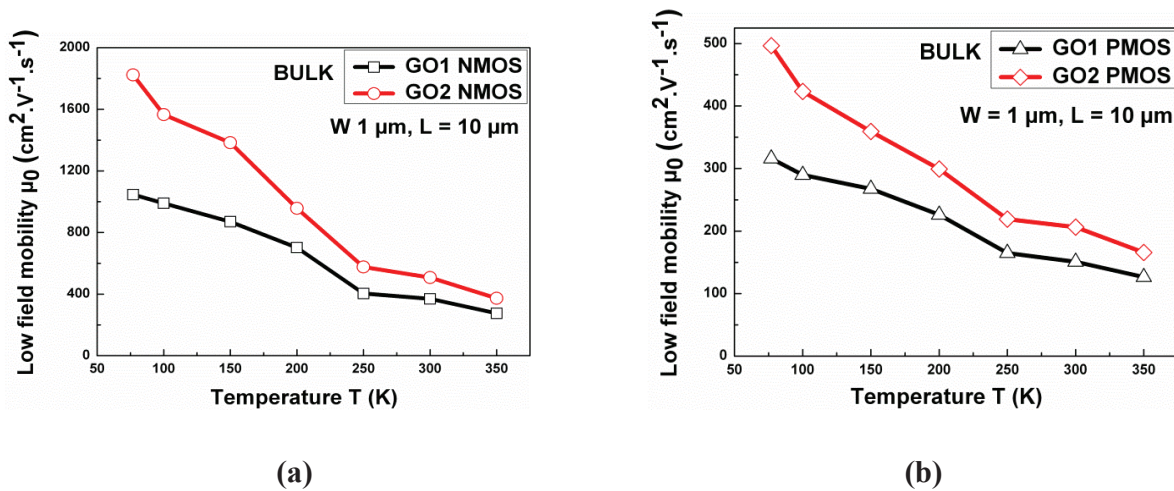


Figure B.36: Comparaison de mobilité à faible charge de transistors GO1 et GO2 bulk pour NMOS (a) et PMOS

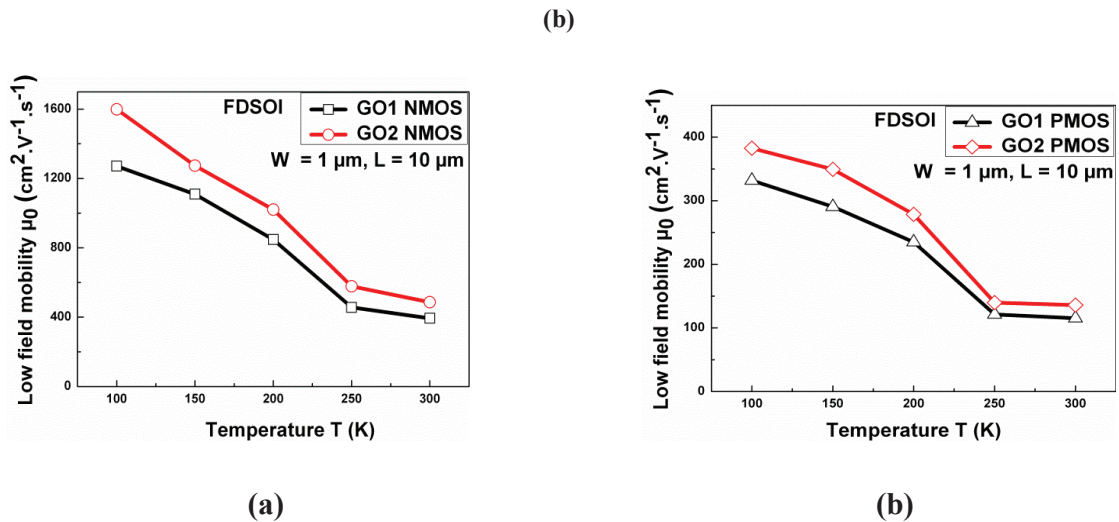


Figure B.37: Comparaison de mobilité à faible charge de transistors GO1 et GO2 sur FDSOI pour NMOS (a) et PMOS (b)

Comme la mobilité à faible charge, la mobilité effective montre une augmentation en fonction du champ effectif pour le bulk (figure B.38a) et le FDSOI (figure B.38b). Il faut noter que les courbes de mobilité effective de GO1 et GO2 ont tendance à se rejoindre à fort champ effectif pour lequel la mobilité est limitée par la rugosité de surface. Cette tendance est plus accentuée pour les dispositifs FDSOI. En effet, à fort champ effectif les porteurs dans le canal sont très proches de l'interface Si/oxyde et l'écrantage par le mécanisme de diffusion sur la rugosité de surface devient majoritaire et est identique pour les deux types de transistors

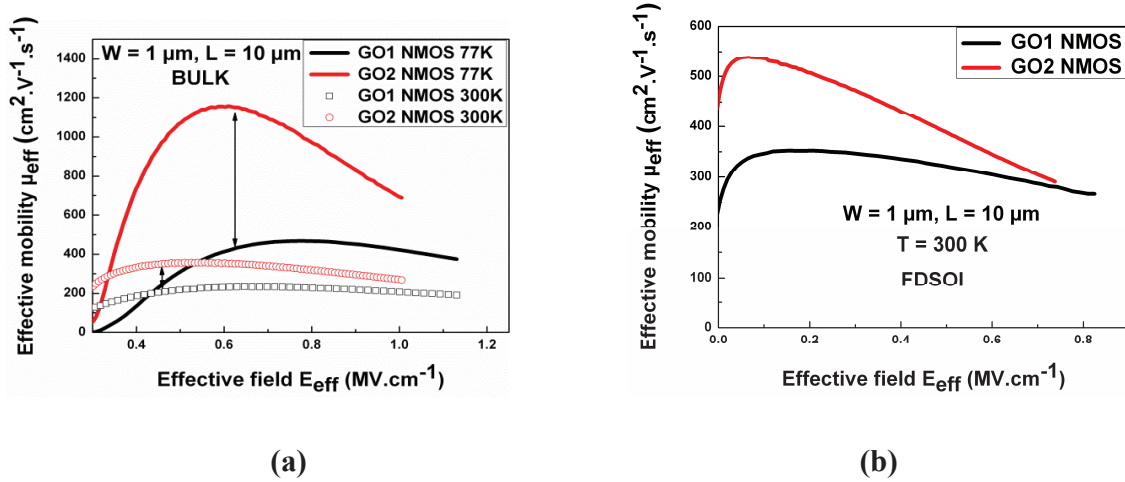


Figure B.38: Comparaison de mobilité effective en fonction du champ effectif de transistors GO1 et GO2 pour du bulk (a) et du FDSOI (b)

De l'amélioration de mobilité effective, la mobilité additionnelle sur les transistors GO2 par rapport aux GO1 μ_{add} peut être obtenue en utilisant (B.25) :

$$\frac{1}{\mu_{add}} = \frac{1}{\mu_{GO1}} - \frac{1}{\mu_{GO2}} \quad (B.25)$$

Sur les NMOS à la figure B.39a, μ_{add} est croissant avec le champ effectif, mettant en évidence un comportement typique du RCS. Cela veut tout simplement dire que le fait de repousser le high-κ réduit l'impact du RCS sur le transport des porteurs dans le canal.

Si pour les NMOS, l'apparition de l'impact de la rugosité de surface sur la mobilité additionnelle tarde à apparaître, sur les PMOS, celui-ci apparaît à moyen champ effectif comme le montre le comportement de μ_{add} sur la [figure B.39b](#).

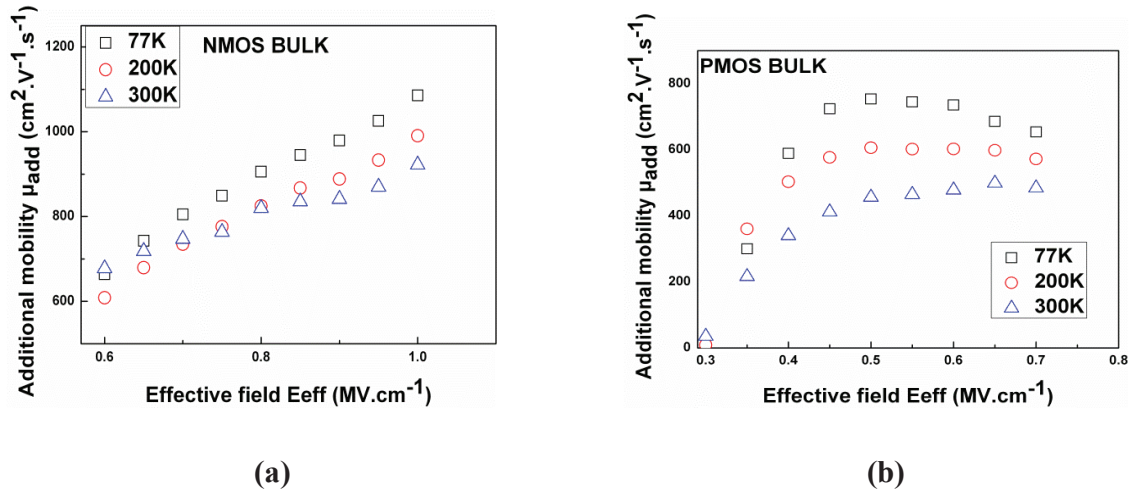


Figure B.39: Mobilité additionnelle en fonction du champ effectif pour NMOS **(a)** et PMOS **(b)** bulk

Le raisonnement servant à comprendre le comportement de la mobilité additionnelle μ_{add} en fonction du champ effectif est inspiré de la [figure B.40](#) de [Takagi94]. De ce graphe on apprend que lorsque la mobilité croît avec le champ effectif, cela veut dire qu'elle est gouvernée par du RCS. Lorsqu'en revanche elle décroît avec le champ effectif, on dira qu'elle est dominée par la rugosité de surface.

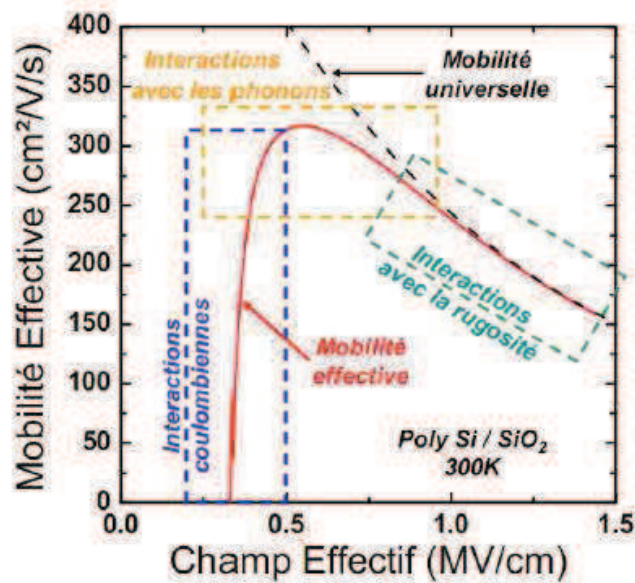


Figure B.40: Principaux mécanismes de réduction de la mobilité dans les transistors MOS [Takagi94]

Il est intéressant de noter que μ_{add} dans le cas du FDSOI montre le même comportement aussi bien en NMOS qu'en PMOS (figure B.41a & B.41b).

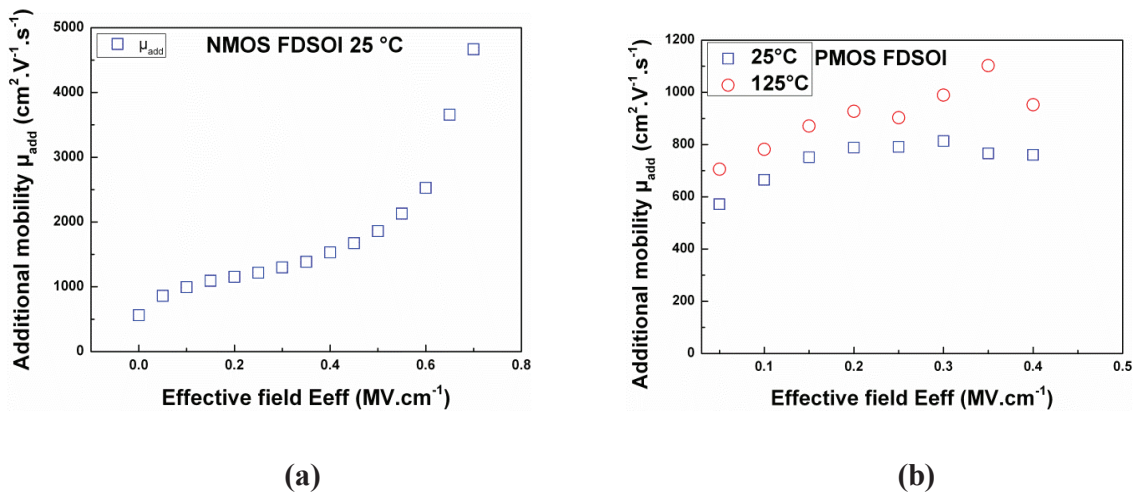


Figure B.41: Mobilité additionnelle en fonction du champ effectif pour NMOS (a) et PMOS (b) FDSOI

L'analyse de l'amélioration de la mobilité effective à champ effectif donné (figure B.42a) et à charge d'inversion donnée (figure B.42b) vient consolider l'hypothèse évoquée plus haut. L'amélioration de mobilité est plus visible à faible champ effectif ou à plus faible charge

d'inversion lorsque le RCS gouverne la mobilité. A forte inversion les porteurs sont plus proches de l'interface et l'impact de la rugosité de surface qui augmente à cette occasion est identique quelque soit l'épaisseur de l'oxyde.

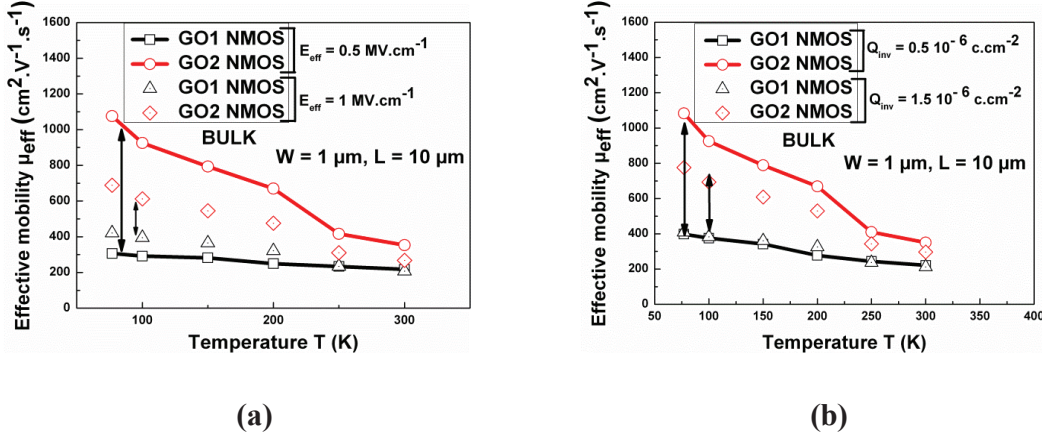


Figure B.42: Mobilité effective en fonction de la température pour GO1 et GO2 à faible et fort champ effectif donné (a), à faible et forte charge inversion donnée (b)

La mobilité à faible charge μ_0 peut être séparée en deux contributions : le remote Coulomb μ_C et le phonon scattering μ_{PH} :

$$\frac{1}{\mu_{0,GO1}} = \frac{1}{\mu_{C,GO1}} + \frac{1}{\mu_{PH,GO1}} \quad (B.26)$$

$$\frac{1}{\mu_{0,GO2}} = \frac{1}{\mu_{C,GO2}} + \frac{1}{\mu_{PH,GO2}} \quad (B.27)$$

En supposant que l'origine de la mobilité additionnelle est purement coulombienne et en combinant (B.25), (B.26) et (B.27) on obtient (B.28). On peut calculer la quantité de charges ΔQ_{ox} responsable de la mobilité additionnelle entre GO1 et GO2 (B.29) (figures B.43a, B.43b & B.44).

$$\frac{1}{\mu_{add}} = \frac{1}{\mu_{C,GO1}} - \frac{1}{\mu_{C,GO2}} \quad (B.28) \quad \Delta Q_{ox} = \frac{\mu_{add}^{-1}}{\alpha \cdot q} \quad (B.29)$$

L'architecture du GO2 avec un oxyde plus volumineux (high- κ + IL + SiO₂) peut laisser penser qu'on aurait plus de charges dans ces dispositifs. Mais comme le high- κ s'éloigne du canal (à

cause de la diminution des interactions dues au high- κ), tout se passe en réalité comme s'il y avait moins de charges.

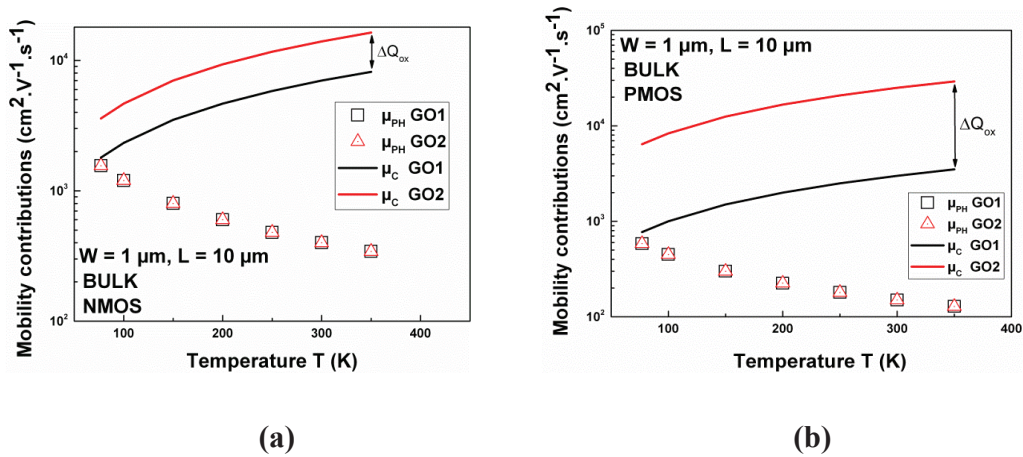


Figure B.43: Variation des contributions de mobilité en fonction de la température pour NMOS (a) et PMOS (b)

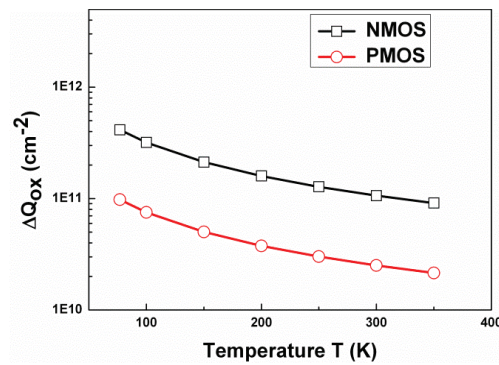


Figure B.44: Illustration de la quantité de charge responsable de la mobilité additionnelle entre GO1 et GO2

La question de savoir si cette supposée charge additionnelle est assez grande pour influencer la tension de bandes plates V_{FB} ou la tension de seuil V_{TH} se pose. Les équations de V_{FB} et de V_{TH} sont données par (B.30) et (B.31):

$$V_{FB} = \Phi_{MS} - \frac{Q_{ox}}{C_{ox}} \quad (B.30) \quad V_{TH} = V_{FB} + 2.\Phi_F + \frac{Q_{dep}}{C_{ox}} \quad (B.31)$$

$\Phi_{MS} = \Phi_M - \Phi_S$, avec Φ_M et Φ_S les travaux de sortie du métal et du semiconducteur respectivement. Φ_F est le niveau de fermi du semiconducteur. Q_{ox} est la charge dans l'oxyde, Q_{dep} la charge de déplétion et C_{ox} la capacité de l'oxyde.

La figure B.45 montre que C_{ox} est quasiment deux fois plus faible dans les transistors GO2. Cela veut dire que ces dispositifs devraient avoir un V_{FB} plus faible en valeur relative (B.30) ($V_{FB} < 0$ pour les NMOS). Or tel n'étant pas le cas (figure B.46), on peut en déduire que ΔQ_{ox} s'oppose à la réduction du V_{FB} . La figure B.47 est en adéquation avec l'équation (B.31). En effet une augmentation de V_{FB} (en valeur relative) accompagnée d'une diminution de C_{ox} va augmenter le V_{TH} des transistors GO2.

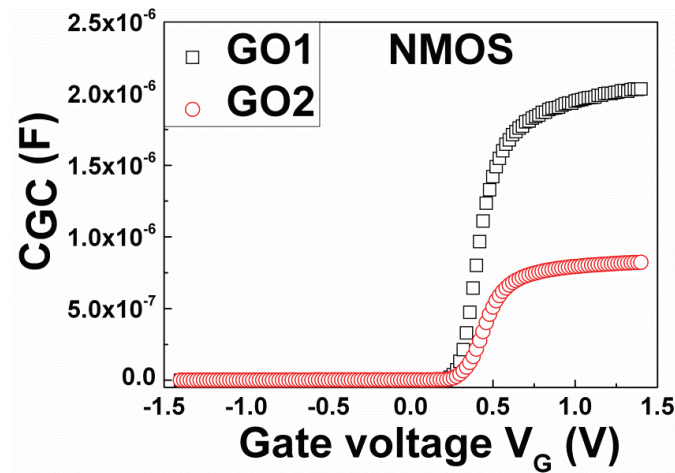


Figure B.45 : Capacité-grille canal C_{GC} de transistors GO et GO2

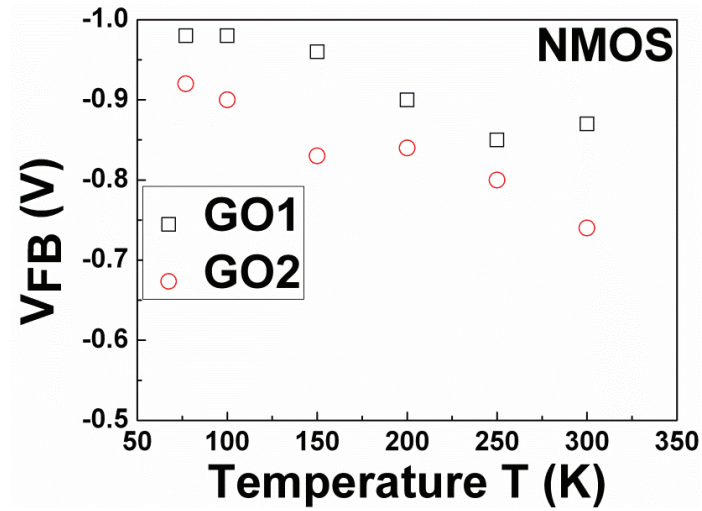


Figure B.46: Evolution de la tension de bandes plates V_{FB} en fonction de la température pour GO1 et GO2

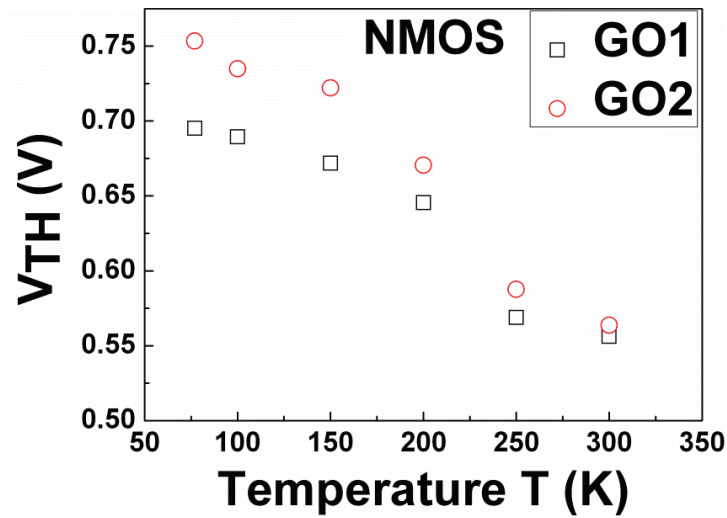


Figure B.47: Evolution de la tension de seuil V_{TH} en fonction de la température pour GO1 et GO2

B.III.5) Bilan et conclusion

Un gain en mobilité est constaté sur les transistors type GO2 avec un oxyde de grille plus épais. Cela confirme que le high- κ qui dans un tel cas est repoussé plus loin a un impact négatif sur la mobilité. L'analyse de la mobilité additionnelle a permis de montrer que cet impact sur le transport est majoritairement de nature coulombienne. Il en découle avec certitude que même si le high- κ présente un réel avantage en terme de miniaturisation, son impact sur le transport électronique est non négligeable.

Table des figures de la partie B

Figure B.1: Principe du recuit standard en présence de l'hydrogène.....	52
Figure B.2: Empilement du stack de grille des N&PMOS étudiés dans ce chapitre.....	53
Figure B.3: Illustration de l'amélioration de courant ION après recuit HPH2 et HPD2	54
Figure B.4: Mobilité à faible charge après recuit standard, HPH2 et HPD2 pour NMOS (a) et PMOS (b).....	54
Figure B.5: Comportement des résistances source-drain après HPFA.....	55
Figure B.6: Extraction de la densité d'états d'interface par pompage de charge après recuit standard, HPH2 et HPD2 pour NMOS (a) & PMOS (b)	57
Figure B.7: Variation de la tension de seuil après HPFA pour NOMS (a) & PMOS (b)	58
Figure B.8: Mobilité à faible charge μ_0 en fonction de la température ...	59
Figure B.9: Représentation des différentes contributions de la mobilité à faible charge pour un transistor long (a) et pour un transistor court (b) .	60
Figure B.10: Illustration des différentes contributions de la mobilité d'un transistor long et court n'ayant subi aucun recuit (a); Evolution des constantes de remote coulomb scattering et des défauts neutres après recuit (b).....	61
Figure B.11: Illustration des différentes contributions de la mobilité d'un transistor long et court n'ayant subi aucun recuit (a); Evolution des constantes de remote coulomb scattering sur un PMOS long après recuit	

Table des figures de la partie B

(b); Evolution des constantes de remote coulomb scattering et de défauts neutres sur un PMOS court après recuit (c).....	62
Figure B.12: Mobilité effective après recuit standard, HPH2 et HPD2 pour un PMOS (a) et un NMOS (b)	63
Figure B.13: Comparaison entre sans recuit, le recuit standard, HPH2 et HPD2 en dégradation HCI (a) et en NBTI (b)	64
Figure B.14: Empilement du stack de grille des N&PMOS étudiés dans ce chapitre.....	66
Figure B.15: Image TEM illustrant la couche de SiGe épitaxiée	66
Figure B.16: V_{TH} simulé et V_{TH} expérimental en fonction du pourcentage du germanium.....	67
Figure B.17: Représentation du courant de drain en fonction de la tension de grille pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$	68
Figure B.18: Illustration de la mobilité à faible charge (a) et de la mobilité effective (b) pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$	69
Figure B.19: Illustration de la tension de seuil en fonction de la longueur pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$	70
Figure B.20: Illustration de la localisation des différents mécanismes susceptibles d'influencer négativement la mobilité	71
Figure B.21: La fonction de Maserjian en fonction de la tension de grille: extraction de V_{FB} et du niveau de dopage	73
Figure B.22: Illustration de la densité d'états d'interface en fonction de la fréquence pour Si et $Si_{0.7}Ge_{0.3}$	73

Table des figures de la partie B

Figure B.23: Mobilité à faible charge expérimentale et théorique obtenue par l'équation B.6.....74

Figure B.24: Variation du courant de drain à V_{GT} constant en fonction de la température pour Si et $Si_{0.7}Ge_{0.3}$ (a) ; mobilité à faible charge en fonction de la température pour Si et $Si_{0.7}Ge_{0.3}$ (b)75

Figure B.25: Variation des différentes contributions de mobilité en fonction de la température pour Si et $Si_{0.7}Ge_{0.3}$ 76

Figure B.26: Illustration de la quantité de charge responsable de la mobilité additionnelle entre Si et $Si_{0.7}Ge_{0.3}$ 77

Figure B.27: Capacité totale en fonction de la tension de grille avant et après implantations ioniques de Ge dans le canal (a); mobilité effective en fonction de la charge d'inversion avant et après implantations ioniques de Ge dans le canal et après dépôt de silicium épitaxié (b) [Tsuchiya10]78

Figure B.28: Illustration de la fluctuation d'un paramètre autour de sa valeur attendue dans le domaine temporel.....79

Figure B.29: Densité spectrale de puissance en fonction de la fréquence: illustration du bruit $1/f$ et du bruit thermique.....80

Figure B.30: Illustration de l'interaction entre le nombre de porteurs N et les pièges dans l'oxyde sources de bruit.....80

Figure B.31: Bruit basses fréquences normalisée avec le carré du courant en fonction du courant de drain pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$ 81

Figure B.32: Variabilité de σ pour Si et SiGe [Rahhal13]82

Figure B.33: Densité de piège dans l'oxyde pour Si, $Si_{0.8}Ge_{0.2}$ et $Si_{0.7}Ge_{0.3}$ 82

Table des figures de la partie B

Figure B.34: Illustration des différents mécanismes qui peuvent dégrader la mobilité dans un transistor ayant un empilement de grille high-k métal, d'après [Weber05]	85
Figure B.35: Illustration des transistors type GO1 (a) et GO2 (b).....	86
Figure B.36: Comparaison de mobilité à faible charge de transistors GO1 et GO2 bulk pour NMOS (a) et PMOS (b)	87
Figure B.37: Comparaison de mobilité à faible charge de transistors GO1 et GO2 sur FDSOI pour NMOS (a) et PMOS (b).....	87
Figure B.38: Comparaison de mobilité effective en fonction du champ effectif de transistors GO1 et GO2 pour du bulk (a) et du FDSOI (b).....	88
Figure B.39: Mobilité additionnelle en fonction du champ effectif pour NMOS (a) et PMOS (b) bulk.....	89
Figure B.40: Principaux mécanismes de réduction de la mobilité dans les transistors MOS [Takagi94].....	90
Figure B.41: Mobilité additionnelle en fonction du champ effectif pour NMOS (a) et PMOS (b) FDSOI.....	90
Figure B.42: Mobilité effective en fonction de la température pour GO1 et GO2 à faible et fort champ effectif donné (a), à faible et forte charge inversion donnée (b)	91
Figure B.43: Variation des contributions de mobilité en fonction de la température pour NMOS (a) et PMOS (b).....	92
Figure B.44: Illustration de la quantité de charge responsable de la mobilité additionnelle entre GO1 et GO2.....	92
Figure B.45 : Capacité-grille canal C_{GC} de transistors GO et GO2	93

Table des figures de la partie B

Figure B.46 : Evolution de la tension de bandes plates V_{FB} en fonction de la température pour GO1 et GO294

Figure B.47 : Evolution de la tension de seuil V_{TH} en fonction de la température pour GO1 et GO294

Liste des tableaux de la partie B

Tableau B.1: Description de l'impact de la longueur effective entre Si et $\text{Si}_{0.7}\text{Ge}_{0.3}$ 70

Tableau B.2: Coefficient de corrélation entre mobilité et bruit et densité de pièges81

Partie C

Table des matières de la partie C

Partie C	102
Table des matières de la partie C	103
PARTIE C: METHODOLOGIE D'EXTRACTION DES PARAMETRES: ETAT DES LIEUX ET AMELIORATIONS	105
CHAPITRE C.I : L'EXTENSION DE LA « FONCTION Y » POUR LE REGIME DE SATURATION	106
C.I.1) La « fonction Y » en linéaire.....	106
C.I.1.a) Définition de la « fonction Y »	107
C.I.1.b) Extraction de paramètres	107
C.I.2) La « fonction Y » en régime de saturation.....	112
C.I.2.a) La fonction Y du régime linéaire en régime de saturation.....	112
C.I.2.b) Mobilité apparente	116
C.I.2.c) Vitesse de saturation des porteurs en régime de saturation	116
C.I.2.c) Vitesse moyenne des porteurs en régime de saturation.....	118
C.I.3) Application : Comparaison de performance entre silicium-germanium et germanium massif..	119
C.I.4) Bilan et conclusion	121
CHAPITRE C.II : A LA RECHERCHE DE BALISTICITE DANS LE GERMANIUM : LA METHODE DE LA « VITESSE LIMITE ».....	122
C.II.1) Les dispositifs étudiés.....	123
C.II.1.a) Procédé de fabrication [Mitard08].....	123
C.II.1.b) Les courants de fuites dans le germanium	125
C.II.2) Vérification des propriétés de transport sur nos dispositifs Ge.....	128
C.II.2.a) Mobilité effective.....	128
C.II.2.b) Estimation de la balisticité dans le Ge en régime linéaire.....	129

PARTIE C: METHODOLOGIE D'EXTRACTION DES PARAMETRES: ETAT DES LIEUX ET AMELIORATIONS

C.II.3) La méthode de la vitesse limitante: le v_{lim}	131
C.II.4) Amélioration de la méthode de « v_{lim} »	132
C.II.4.a) Extraction de la tension de seuil en saturation.....	132
C.II.4.b) Confirmation de la méthode d'extraction de V_{THsat} par des simulations TCAD.....	134
C.II.5) Application de la méthode de « v_{lim} » sur Ge-pMOSFET	138
C.II.5.a) Détermination de la nature du transport	138
C.II.6) Bilan et conclusion	141
CHAPITRE C.III : MESURES HF POUR LE TRANSPORT ELECTRONIQUE	142
C.III.1) Contexte	142
C.III.2) Problématique	143
C.III.3) Solutions techniques.....	143
C.III.3.a) Mesures basses fréquences.....	143
C.III.3.b) Mesures hautes fréquences (HF)	145
C.III.3.c) Obtention des capacités et de la charge d'inversion	148
C.III.4) Extraction de mobilité et de vitesse.....	151
C.III.5) Application à l'extraction du v_{lim}	155
C.III.6) Bilan et conclusion.....	156
Table des figures de la partie C.....	157
Liste des tableaux de la partie C.....	161

PARTIE C: METHODOLOGIE D'EXTRACTION DES PARAMETRES: ETAT DES LIEUX ET AMELIORATIONS

Nous avons évoqué dans la partie A des paramètres tels que la mobilité, la tension de seuil ou encore la vitesse des porteurs. Nous allons maintenant voir comment il est possible de remonter jusqu'à ces paramètres à partir de mesures de courant. C'est ce que l'on appelle l'extraction de paramètres.

L'extraction de paramètres de transistors est primordiale pour comprendre l'évolution des technologies d'un nœud à un autre. Les paramètres de transport restent particulièrement des outils incontournables pour la compréhension physique du transistor MOS. Le besoin d'extraction de paramètres de transport intervient par exemple lorsqu'on ajoute une étape supplémentaire dans le processus de fabrication du transistor ou lorsqu'on modifie l'architecture du MOS. On est ainsi confronté à un besoin accru d'adaptation des techniques d'extraction. Dans certains cas il faut mettre en place de nouvelles méthodes pour suivre la mutation technologique dans un but double : compréhension des phénomènes physiques mis en jeu et métrique de suivi des performances de la technologie.

Au premier chapitre de cette partie, nous avons étendu la « fonction Y » jusqu'au régime de saturation pour extraire des paramètres dans ce régime.

Dans le deuxième chapitre, nous avons amélioré la technique dite de « v_{lim} » développée lors d'une thèse précédente. Elle a ensuite été appliquée à une technologie sur germanium massif.

Le dernier chapitre de cette partie est consacré à l'obtention de la charge d'inversion en utilisant des mesures hautes fréquences.

CHAPITRE C.I : L'EXTENSION DE LA « FONCTION Y » POUR LE REGIME DE SATURATION

La caractérisation de transistor MOS devient de plus en plus problématique à cause de la réduction des dimensions et de l'introduction de nouveaux matériaux. La fonction Y permet d'accéder à un certain nombre de paramètres électriques très utiles pour la compréhension de la physique du transistor. Cependant ces extractions ne sont faites que dans le régime linéaire. Nous avons étendu la fonction Y dans le régime d'opération de saturation du transistor (i.e. dans la zone où $V_{DS}=V_{DD}=1V$) pour y extraire des paramètres d'une manière plus représentative [Diouf12b-Diouf13a].

C.I.1) La « fonction Y » en linéaire

Un nombre important de techniques de caractérisation a été développé mais la plupart d'entre elles explorent uniquement le régime linéaire à faible tension de drain V_{DS} . L'une d'entre elles est celle dite « de la fonction Y » [Ghibaudo88], qui a l'avantage d'éliminer l'influence des résistances séries dans l'extraction.

Tout d'abord, commençons par présenter l'équation du courant de drain I_D en fonction de la tension de grille V_G (figure C.1) en régime ohmique :

$$I_D = \frac{W}{L} \cdot \mu_0 \cdot C_{ox} \cdot \frac{V_G - V_{TH}}{1 + \theta_1 \cdot (V_G - V_{TH}) + \theta_2 \cdot (V_G - V_{TH})^2} \cdot V_D \quad (C.1)$$

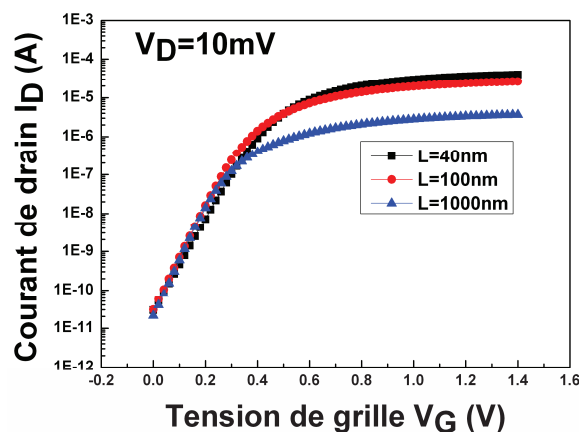


Figure C.1: Exemple de courant de drain en fonction de la tension de grille de nMOSFETs. $V_D = 10\text{mV}$

Certains paramètres de cette équation sont nouveaux et méritent d'être élucidés. On y trouvera ainsi les paramètres θ_1 et θ_2 qui sont des paramètres de réduction de mobilité.

A partir de (C.1), plusieurs méthodes dont la fonction Y, la méthode de Hamer [Hamer86] et celle de McLarty [McLarty95] permettent de remonter aux paramètres électriques tels que μ_0 , V_{TH} , θ_1 et θ_2 ...etc.

Nous nous concentrerons sur la fonction Y, majoritairement utilisée tout au long de cette thèse pour l'extraction de paramètres.

C.I.1.a) Définition de la « fonction Y »

Cette fonction Y est définie comme suit :

$$Y = \frac{I_D}{\sqrt{G_m}} \quad (C.2)$$

Avec G_m la transconductance définie par :

$$G_m = \frac{\partial I_D}{\partial V_G} = \frac{W}{L} \cdot C_{ox} \cdot \mu_0 \cdot \frac{V_D}{(1 + \theta_1 \cdot (V_G - V_{TH}))^2} \quad (C.3)$$

On obtient donc

$$Y = \sqrt{\mu_0 \cdot C_{ox} \cdot V_D \cdot \frac{W}{L} \cdot (V_G - V_{TH})} \quad (C.4) \text{ si } \theta_2 = 0$$

C.I.1.b) Extraction de paramètres

En regardant (C.4), apparaît l'avantage majeur de la méthode « fonction Y » qui est son indépendance vis-à-vis de la résistance série source-drain R_{SD} . En effet, le premier facteur d'atténuation de mobilité θ_1 , contenant la résistance R_{SD} (cf. C.5), est absente dans la formule de la fonction Y. Ceci s'avère crucial pour les transistors de dernières générations, c'est-à-dire de longueur de grille inférieure à 100 nm, car pour ceux-ci la résistance R_{SD} perturbe fortement le

courant de drain. De plus, cette méthode n'utilise qu'une seule dérivation du courant de drain ce qui limite l'impact du bruit de mesure.

$$\theta_1 = \theta_{1,0} + \beta * R_{SD} \text{ avec } \beta = \frac{W}{L} \cdot \mu_0 \cdot C_{ox} \quad (\text{C.5}) \quad (\beta \text{ représente le gain en transconductance}).$$

Ainsi en forte inversion ($V_G > V_{TH}$) la fonction $Y(V_G)$ varie linéairement avec V_G ce qui permet d'obtenir aisément la tension de seuil V_{TH} en extrapolant la valeur à zéro de la partie linéaire de la courbe $Y(V_G)$ comme le montre la **figure C.2**.

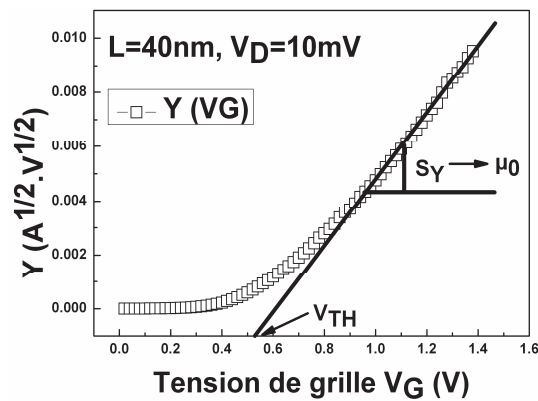


Figure C.2: Figure d'illustration de la fonction Y en fonction de la tension de grille

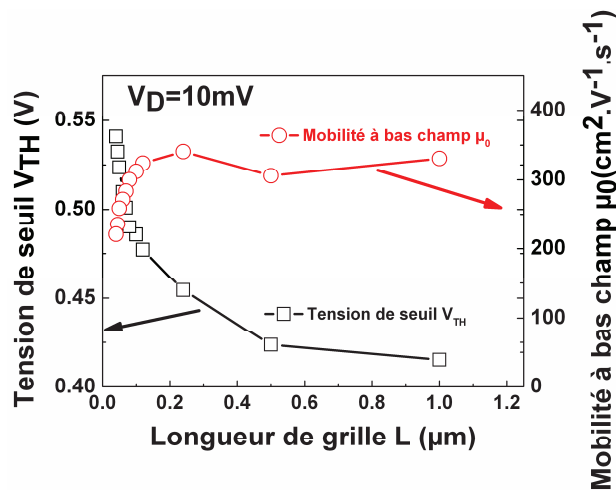


Figure C.3: Illustration de la tension de seuil et de la mobilité en fonction de la longueur de grille extraits par la « fonction Y »

(C.4) montre que $V_G=V_{TH}$ lorsque $Y=0$. La pente de $Y(V_G)$ notée S_Y ($S_Y = \sqrt{\beta V_D}$) permet de remonter à la mobilité μ_0 :

$$\mu_0 = \frac{S_Y^2 \cdot L}{W \cdot C_{ox} \cdot V_D} \quad (C.6)$$

Des exemples de $V_{TH}(L)$ et de $\mu_0(L)$ sont donnés sur la **figure C.3**.

Pour obtenir le coefficient d'atténuation θ_1 , on définit la fonction $X(V_G)$:

$$X = \frac{1}{\sqrt{G_m}} = \frac{1 + \theta_1 \cdot (V_G - V_{TH})}{\sqrt{\beta \cdot V_D}} \quad \text{si } \theta_2 = 0 \quad (C.7)$$

De la même manière que la fonction Y , la fonction X est linéaire avec V_G comme le montre (C.7). Si S_X est la pente de $X(V_G)$ lorsque $V_G > V_{TH}$, On en déduit ainsi le paramètre θ_1 qui est donné par :

$$\theta_1 = S_Y \cdot S_X \quad (C.8)$$

Il ne nous manque à ce stade du raisonnement que le paramètre θ_2 pour pouvoir réécrire le courant de drain comme décrit dans (C.1). Pour déterminer le paramètre θ_2 manquant, on cherchera par la suite à ajuster le courant recalculé avec le courant mesuré jusqu'à obtenir θ_2 avec une erreur fixée à l'avance. La **figure C.4** montre le courant de drain recalculé sans et avec le paramètre θ_2 . On peut remarquer que comme dans certains cas, ce paramètre peut être négligé. Cependant il faut toutefois vérifier si on est proche du courant mesuré, même sans θ_2 .

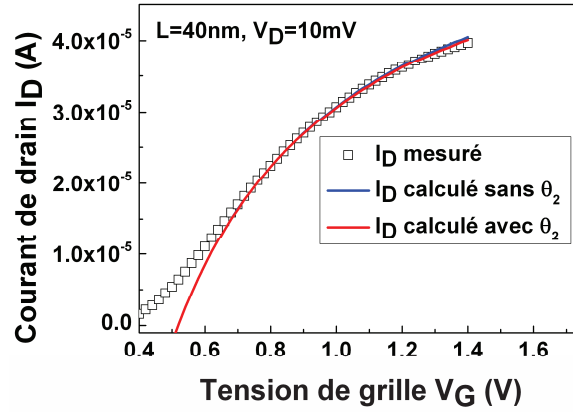


Figure C.4: Illustration du courant de drain mesuré et du courant de drain calculé avec la formule 30

Il est possible d'extraire d'autres paramètres comme la résistance source-drain R_{SD} ou encore la mobilité effective. La résistance source-drain peut être obtenue en traçant la résistance totale R_{TOT} en fonction de l'inverse du gain en transconductance ($\frac{1}{\beta}$) pour plusieurs longueurs

[Fleury09c]. En effet :

$$R_{TOT} = \frac{V_D}{I_D} = R_{SD} + R_{CH} = R_{SD} + \frac{1}{\beta} \cdot \left(\frac{1 + \theta_1 \cdot (V_G - V_{TH}) + \theta_2 \cdot (V_G - V_{TH})^2}{V_G - V_{TH}} \right) \quad (C.9)$$

Ainsi comme illustré sur la **figure C.5**, l'ordonnée à l'origine de R_{TOT} ($\frac{1}{\beta}$) permet d'accéder à R_{SD} .

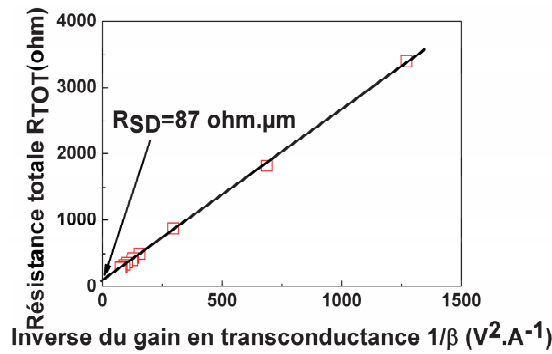


Figure C.5: Méthode d'extraction de la résistance série $R_{TOT} \left(\frac{1}{\beta} \right)$

La mobilité effective, généralement extraite par la méthode dite « split CV » (que nous verrons au [chapitre III de cette partie](#)) peut être approximée via les paramètres que nous venons d'évoquer. Elle est approximée par la formule suivante :

$$\mu_{eff} = \frac{\mu_0}{1 + \theta_1 \cdot (V_G - V_{TH}) + \theta_2 \cdot (V_G - V_{TH})^2} \quad (C.10)$$

La mobilité effective permet d'avoir des informations sur le dispositif à forte inversion contrairement à la mobilité à bas champ qui lui donne des informations autour du V_{TH} . La formule [C.10](#) montre que $\mu_{eff} = \mu_0$ lorsque $V_{TH} = 0$. Expérimentalement, cela n'est pas toujours vrai lorsqu'on extrait μ_{eff} et μ_0 séparément. Ceci est dû aux incertitudes de mesure.

[\(C.10\)](#) est d'autant plus importante qu'elle permet d'avoir une idée de la mobilité effective sur des dispositifs ultra-courts car la méthode du split CV y est difficilement applicable. Une illustration est donnée sur la [figure C.6](#).

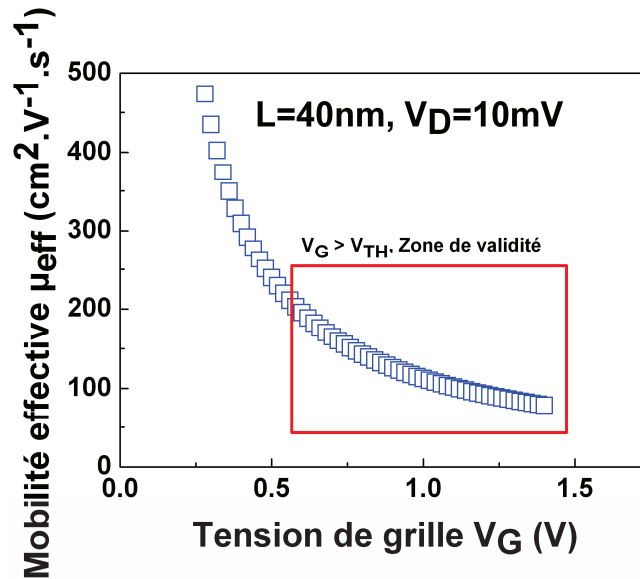


Figure C.6: Mobilité en fonction de la tension de grille: extraction avec les paramètres extraits via la fonction Y

Beaucoup d'améliorations ont été apportées à cette technique pour la rendre plus précise dans le régime linéaire [Fleury08a-Subramanian10], mais sans jamais vraiment exploré son application en régime de saturation.

C'est ainsi que nous avons étendu la fonction dans le régime saturé. Grâce à cela, une mobilité apparente en saturation est extraite ainsi que la vitesse de saturation. La méthode est ensuite appliquée sur des transistors à base de substrat massif germanium.

C.I.2) La « fonction Y » en régime de saturation

Tout ce qui suit dans cette sous partie est fait en régime de saturation.

C.I.2.a) La fonction Y du régime linéaire en régime de saturation

Dans [Ghibaudo88], la fonction Y est définie dans le régime linéaire en fonction du courant de drain et de la transconductance (C.2). On peut aussi relier la résistance totale à la puissance totale dissipée dans le dispositif comme dans (C.11) [Xu11].

En combinant (C.2), (C.3) et (C.11), tout en supposant que R_{SD} ne varie pas beaucoup avec V_G , on peut définir une fonction Y étendue, valable quel que soit V_D (C.12) et éliminant l'influence de la résistance série comme dans le cas linéaire.

$$P = R_{TOT} \cdot I_D^2 \quad (C.11) \quad \frac{dR_{tot}}{dV_G} = \frac{V_D}{Y^2} \quad (C.12)$$

La variation de R_{SD} avec V_G ne dépasse généralement pas 10%. Cette variation est très faible dans nos dispositifs comme illustré sur le [tableau C.1](#).

VGT (V)	0.4	0.5	0.6	0.65
RSD (ohm.µm)	83.09	83.82	83.14	83.44

Tableau C.1: Illustration de la faible fluctuation de la résistance série avec la tension de grille dans nos dispositifs

En plus de la prise en compte de la variation de R_{SD} avec V_G , la gamme de linéarité de $Y(V_G)$ dans laquelle l'extraction est faite doit aussi être bien choisie.

Cette gamme d'extraction peut affecter le résultat si elle n'est pas bien choisie. Dans notre cas, on s'assure de rester dans une zone de faible champ et d'inversion suffisamment forte. Pour prendre en compte la fluctuation de cette gamme avec la longueur de grille, nous calculons une tension de seuil primaire, noté V_{TH}^* pour chaque longueur de grille. Pour le calcul de V_{TH}^* , nous utilisons la technique d'extrapolation avec le maximum de la transconductance et le courant de drain. En effet, on définit V_{t+} et V_{t-} comme les valeurs de V_G entourant V_{t^*} qui est la valeur de V_G correspondant au maximum de G_m . La régression linéaire qui passe par V_{t+} et V_{t-} coupe l'axe V_G en V_{TH}^* . Cette méthode d'extrapolation est illustrée sur la [figure C.7](#).

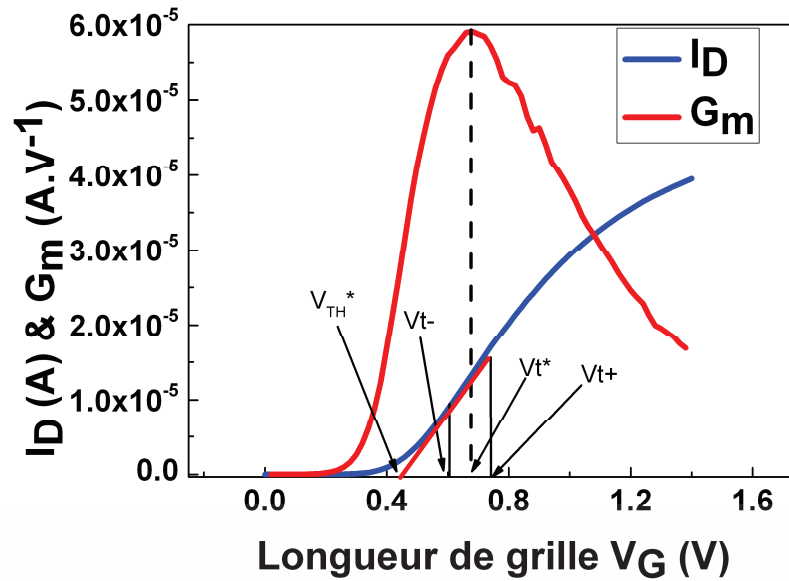


Figure C.7: Illustration de l'extraction de V_{TH}^* avec la méthode d'extrapolation

Cette valeur de V_{TH}^* est utilisée pour mieux définir la plage d'extraction du vraie V_{TH} comme dans (C.13). Dans cette équation, V_{gmin} et V_{gmax} serviront de points de limite de la régression linéaire de la fonction $Y(V_G)$.

$$\begin{pmatrix} V_{g \min} \\ V_{g \max} \end{pmatrix} = V_{TH}^* + \begin{pmatrix} 0.35 \\ 0.55 \end{pmatrix} \quad (C.13)$$

En respectant toutes ces procédures, comme dans le cas du régime ohmique, on peut alors extraire une mobilité apparente en saturation $\mu_{0satapp}$ en utilisant la racine carrée de la pente de la droite $Y(V_G)$.

Sur la figure C.8, on peut vérifier que la fonction Y reste linéaire avec V_G même à fort V_D . Cette linéarité est vérifiée pour deux longueurs de 40 et 120 nm, ce qui confirme la validité de l'extension de la méthode en régime de saturation.

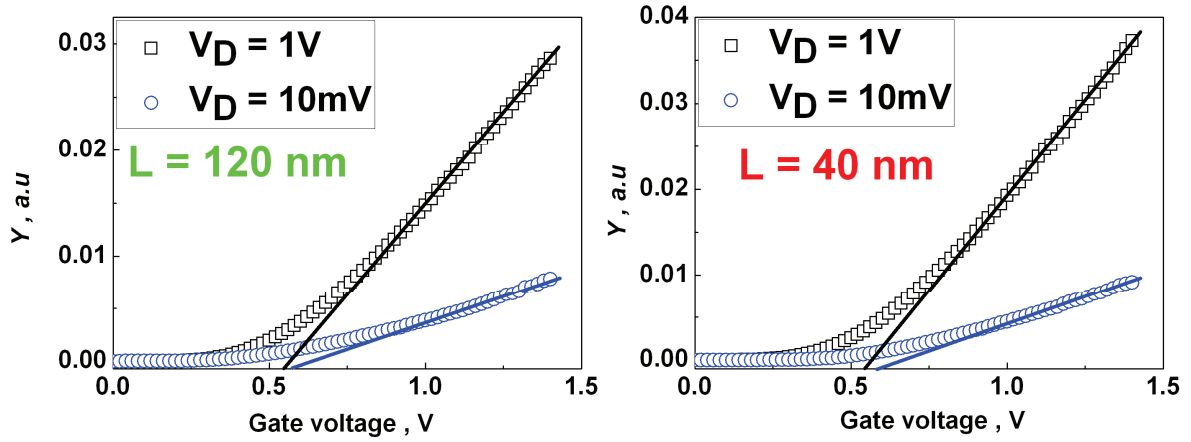


Figure C.8: Illustration de la fonction Y en fonction de VG en saturation

A partir de la partie linéaire, on peut calculer comme dans le cas du régime linéaire, la tension de seuil V_{TH} , le gain en transconductance β et la mobilité apparente en saturation comme dans (C.14).

$$\beta = \frac{W_{eff} \cdot C_{ox} \cdot \mu_{0satapp}}{L_{eff}} \quad (C.14)$$

Pour encore rendre plus fiable l'extraction, comme dans le cas à faible V_D , on cherche à reproduire le courant au mieux avec l'équation (C.1) et en le comparant avec le courant mesuré (figure C.9).

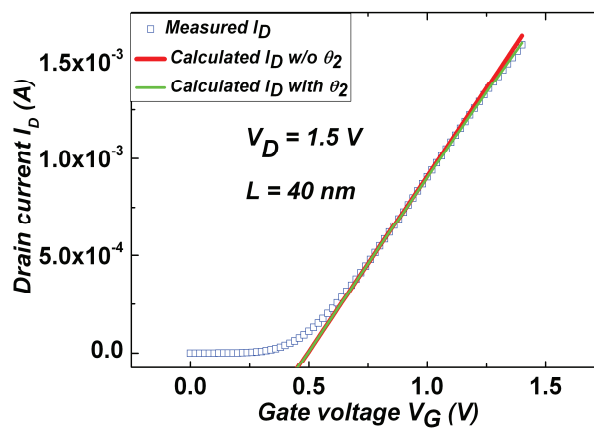


Figure C.9: Courant mesuré et calculé avec l'équation C.30 en saturation

La **figure C.9** montre que même sans la prise en compte du paramètre θ_2 , on arrive à bien reproduire (avec l'**expression 1**) le courant mesuré.

C.I.2.b) Mobilité apparente

Aussi bien pour les nMOSFETs que les pMOSFETs, la mobilité apparente en saturation a été extraite (**C.14**) et les valeurs obtenues sont fiables (**figure C.10**). Même s'il n'a pas de sens physique, la mobilité apparente est un indicateur de performance de transport en régime de saturation. Il faut noter que sa dégradation avec la diminution de la longueur est confirmée à fort V_D . Une autre dégradation de la mobilité apparente cette fois ci quand V_D augmente est aussi à signaler : c'est la dégradation de la mobilité en fonction du champ longitudinal (**C.15**).

$$\mu_{0satapp} = \frac{\mu_0}{1 + \frac{V_D \mu_0}{L_{eff} v_{sat}}} \quad (\text{C.15}),$$

En régime linéaire la mobilité apparente $\mu_{0satapp} \approx \mu_0$

C.I.2.c) Vitesse de saturation des porteurs en régime de saturation

Dans l'hypothèse d'un transport dérive diffusion, nous avons proposé de calculer la vitesse de saturation sans le besoin de corriger l'influence des R_{SD} en utilisant la mobilité apparente en saturation $\mu_{0satapp}$ comme outil mathématique. Pour cela on utilise l'équation (**C.15**) qui donne la dégradation de la mobilité en fonction du champ longitudinal. En **figure C.11**, la pente du tracé de $\mu_{0satapp}^{-1}$ versus V_D nous donne v_{sat} . On peut ainsi voir que notre méthode donne des valeurs acceptables (**figure C.12**).

Il faut aussi noter que cette méthode d'extraction de v_{sat} par la fonction Y peut être facilement implémentée industriellement à cause du fait qu'on n'a pas besoin d'extraire ni de corriger l'influence des R_{SD} sur la tension de grille V_G .

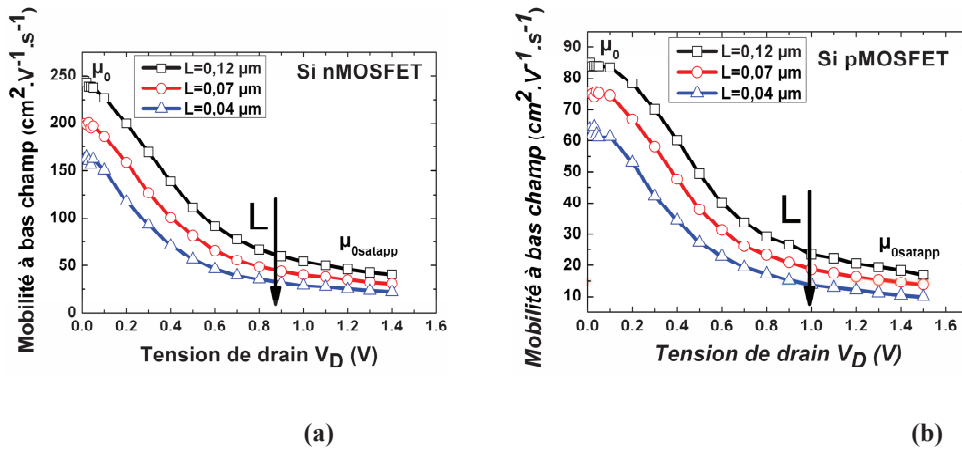


Figure C.10: Illustration de la mobilité apparente en saturation en fonction du champ longitudinal. (a) nMOSFET & (b) pMOSFET.

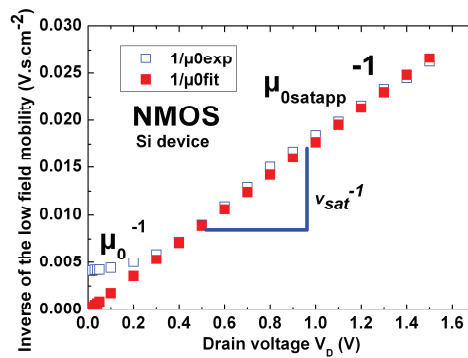


Figure C.11: Procédure d'extraction de la vitesse de saturation en utilisant la fonction Y en saturation

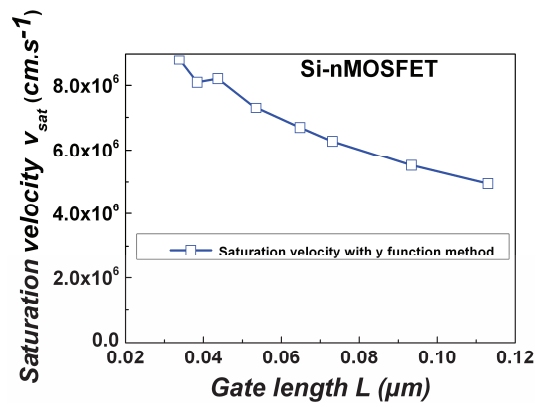


Figure C.12: Vitesse de saturation extraite avec la fonction Y et la méthode standard, avec et sans correction des R_{SD} (faibles valeurs de R_{SD})

C.I.2.c) Vitesse moyenne des porteurs en régime de saturation

Contrairement à la vitesse de saturation, la vitesse moyenne ne fait aucune hypothèse préalable sur la nature du transport. Elle donne la vitesse moyenne acquise par les porteurs quel que soit le régime de transport considéré. Une manière standard de déterminer la vitesse moyenne est donnée dans (C.16).

$$v_{moy} = \frac{I_{dsat}}{W \cdot Q_{inv}} \quad (C.16)$$

Dans celle-ci, V_{TH} peut être obtenu en utilisant la fonction Y. Comme dans le cas en linéaire, $V_G = V_{TH}$ quand $Y(V_G)$ s'annule. Cependant, la limitation de cette méthode reste le besoin d'approximer la charge d'inversion par $Q_{inv} \approx C_{ox} \cdot (V_G - V_{TH})$. Cette approximation devient de plus en plus incertaine à cause de la miniaturisation. Une solution est proposée au chapitre III de cette partie pour éviter cette approximation. Il faut aussi corriger l'influence des R_{SD} sur la tension de grille pour rendre l'approximation plus juste. Les dispositifs utilisés en figure C.13 ont de faibles valeurs de R_{SD} . Par conséquent, la vitesse moyenne avec ou sans correction des R_{SD} reste quasiment le même. En revanche, dans le cas où R_{SD} est plus important comme pour les dispositifs présentés en figure C.14, on voit que la vitesse moyenne peut être sous-estimée sans la correction des R_{SD} . La méthode R_{tot} ($1/\beta$) est utilisée pour l'extraction des R_{SD} car donnant plus de garantie quant à la précision comparée à la méthode R_{tot} (L_{eff}) à cause de l'incertitude sur l'extraction de la longueur effective.

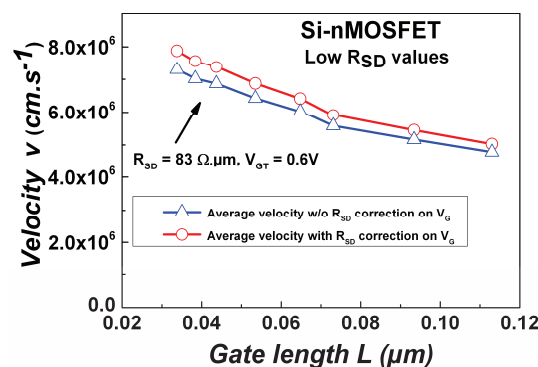


Figure C.13: Vitesse de saturation extraite avec la fonction Y avec de faibles valeurs de R_{SD}

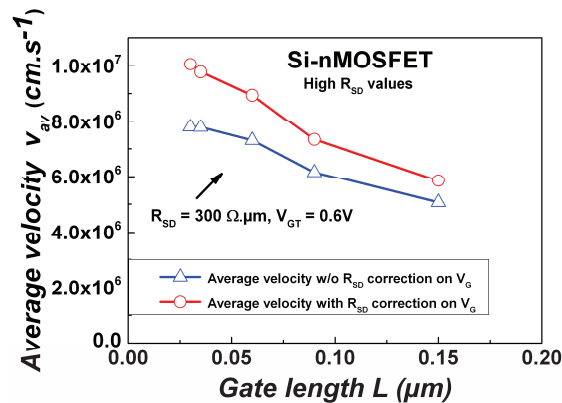


Figure C.14: Vitesse de saturation extraite avec la fonction Y avec de grandes valeurs de R_{SD}

C.I.3) Application : Comparaison de performance entre silicium-germanium et germanium massif

Caractériser le transport de dispositifs à base de germanium (Ge) en régime de saturation reste un sujet alléchant à cause des bonnes propriétés de transport du Ge. Nous allons appliquer la nouvelle méthode d'extraction de v_{sat} à des transistors pMOSFETs à base de germanium fabriqués à l'IMEC dans le cadre du projet DUALLOGIC.

De plus grandes valeurs de mobilité split-CV sur transistor Ge comparé au Si ont été démontrées dans une de nos études [Diouf11]. Cependant, l'amélioration de mobilité linéaire diffère de celle du courant I_{ON} lorsqu'on passe du silicium-germanium (30% de germanium en proportion) au germanium. La figure C.15 montre une augmentation de 160% de la mobilité linéaire lorsqu'on passe du silicium-germanium au germanium. Contrairement à l'amélioration en mobilité linéaire, celle en vitesse de saturation (58%, figure C.16) explique mieux le gain obtenu en courant I_{ON} (52%, figure C.17).

Des doutes subsistent donc quant à la balisticité attendue dans le germanium malgré ses bonnes propriétés de transport. Cela veut aussi dire que pour commencer à avoir de la balisticité (en régime de saturation notamment) dans le germanium, il faudra fabriquer des transistors beaucoup plus courts que 65 nm [Diouf11].

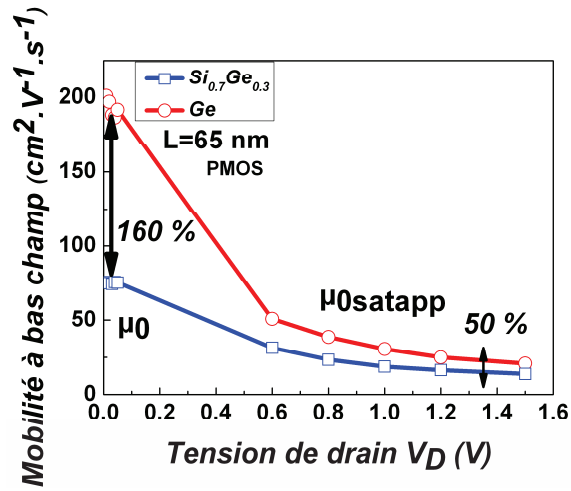


Figure C.15: Extraction de la mobilité faible charge et de la mobilité apparente en saturation sur du $\text{Si}_{0.7}\text{Ge}_{0.3}$ et Ge

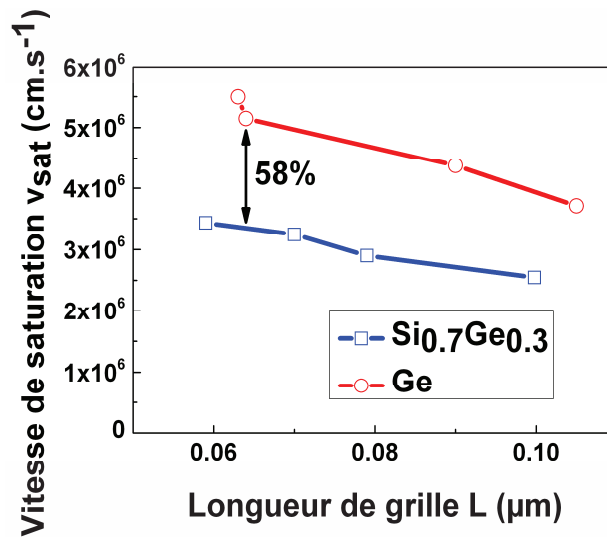


Figure C.16: Comparaison de vitesse de saturation (extraite par la méthode de la fonction Y) entre $\text{Si}_{0.7}\text{Ge}_{0.3}$ et Ge

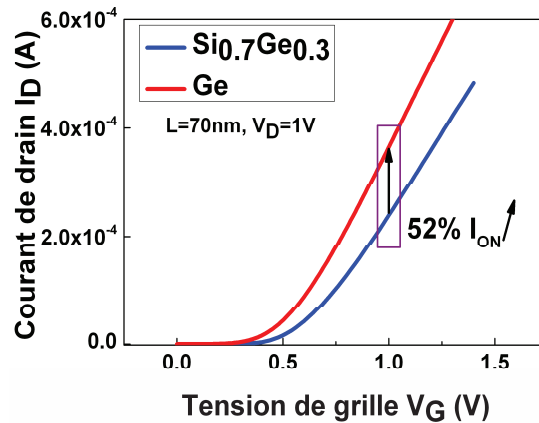


Figure C.17: Courant de drain en fonction de la tension de grille: illustration de l'amélioration entre $\text{Si}_{0.7}\text{Ge}_{0.3}$ et Ge

C.I.4) Bilan et conclusion

Nous avons mis en évidence tout au long de ce chapitre à quel point il était utile de maîtriser l'extraction de paramètres pour comprendre le fonctionnement du transistor. La fonction Y est une technique qui permet d'extraire des paramètres en s'affranchissant de l'influence des résistances source-drain. Nous l'avons étendue en régime de saturation pour obtenir les paramètres en régime de fonctionnement du transistor.

Une vitesse de saturation a été calculée en se servant de la mobilité apparente comme outil mathématique. Il en a découlé grâce à l'application faite sur du SiGe et du Ge, que la mobilité en linéaire ne donnait pas une image directe du courant à l'état ON lorsqu'il s'agit de prévoir des améliorations de performances.

La difficulté liée à la non-convergence entre la mobilité linéaire et le courant I_{ON} a ainsi été surmontée grâce à la vitesse de saturation.

Ces paramètres que sont la mobilité linéaire et la vitesse de saturation sont très intéressants lorsqu'il faut faire un diagnostic rapide du transport. Pour aller plus loin et étudier la nature du transport, une extraction plus poussée s'impose. On va ainsi à la quête de la balisticité dans les transistors courts. Nous utiliserons dans le chapitre suivant la méthode dite de v_{lim} pour caractériser le transport dans des transistors à base de germanium où une meilleure mobilité est attendue.

CHAPITRE C.II : A LA RECHERCHE DE BALISTICITE DANS LE GERMANIUM : LA METHODE DE LA « VITESSE LIMITE »

Dans les transistors MOS longs, la vitesse de dérive des porteurs est proportionnelle au champ électrique avant de tendre vers la vitesse de saturation, et le modèle de transport correspondant est appelé dérive diffusion. Lorsqu'on réduit les dimensions, le modèle de transport quasi-balistique devrait faire son apparition, et la vitesse des porteurs est alors limitée par la vitesse d'injection à la source.

Cette théorie du transport dite « quasi-balistique », caractérisé par un nombre réduit d'interactions des porteurs dans le canal, a été introduite par Natori et Lundstrom [Natori94-Lundstrom97]. A travers des simulations, il a été démontré que ce modèle de transport permettrait d'atteindre des augmentations de performances comme le courant I_{ON} ou la fréquence de coupure [Palestri05-Eminente07]. Dans la littérature, l'évaluation de la balisticité se fait selon le taux de balisticité extrait en régime linéaire ou saturé [Lochtefeld01-Lochtefeld02-Pappas09].

Dans ce chapitre, nous avons utilisé une méthode expérimentale pour extraire la vitesse limitante dans des dispositifs germanium, sans hypothèse préalable sur le modèle de transport. Cette méthode dite de « v_{lim} » présentée dans [Fleury09b], a été améliorée pour mieux prendre en compte la variation de la charge d'inversion avec la tension de drain. Le comportement en fonction de la température nous a permis de conclure quant à la nature du transport électronique [Diouf11]. Avec les bonnes propriétés de transport du germanium [Mitard08-Hellings09-Jaeger07], la chance d'obtenir un transport quasi-balistique dans des transistors de longueur de grille inférieure à 70 nm est peut-être plus importante que dans le silicium.

C.II.1) Les dispositifs étudiés

Les recherches sur les transistors germanium ont été abandonnées à cause des problèmes de fuites inhérents à ce matériau. Les dispositifs étudiés dans ce chapitre ont cependant été optimisés pour minimiser ces fuites, combinant un faible EOT (effective oxide thickness) et des longueurs de grilles inférieures à 70 nm.

C.II.1.a) Procédé de fabrication [Mitard08]

Le processus de fabrication est résumé dans le [tableau C.2](#).

1. Ge-on-Si wafers	9. BF ₂ /pure B extension
2. P well / channel implant	10. spacer formation
3. Box isolation	11. B S/D implant
4. Epitaxial Si passivation	12. Junction anneal
5. 4 nm HfO ₂ dielectric	13. NiGe S/D
6. TaN/TiN metal gate	14. Back end dielectric
7. Dry etch metal gate	15. TiN / Al contact pads
8. As halos	16. Post metallization anneal

Tableau C.2: Processus de fabrication des dispositifs Ge étudiés dans ce chapitre [Mitard08]

D'importantes optimisations ont été apportées, notamment l'ajout d'un cap Si et des étapes technologiques pour contrecarrer les effets de canaux courts. Le stack de grille est constitué d'oxyde high- κ (HfO₂) et de TaN/TiN comme grille métal. Une image TEM est donnée en [figure C.18](#).

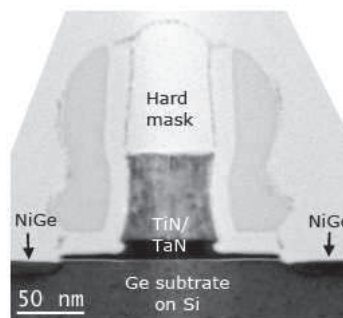


Figure C.18: Coupe TEM illustrant un transistor Ge de longueur de grille 65 nm [Mitard08]

Un gain considérable en performance a été enregistré en comparaison de l'état de l'art. La [figure C.19](#) montre des valeurs record en corrélation I_{OFF}/I_{ON} comparées à celles du germanium sur bulk (Ge bulk) et du germanium sur isolant (GeOI) de la littérature.

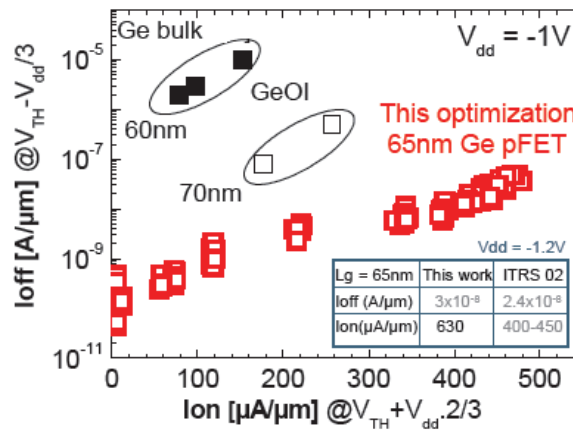


Figure C.19: Valeurs record en corrélation I_{OFF}/I_{ON} de transistor Ge étudiés dans ce chapitre [[Mitard08](#)]

La méthode R_{TOT} ($1/\beta$) évoquée dans le chapitre I de cette partie est utilisée pour extraire les résistances source-drain R_{SD} ([figure C.20](#)). Elle donne plus de précision que les autres méthodes. L'incertitude sur l'extraction de L_{eff} et la variation de la mobilité n'affectent pas la valeur de R_{SD} obtenue. La valeur de R_{SD} obtenue vient confirmer les efforts d'optimisation.

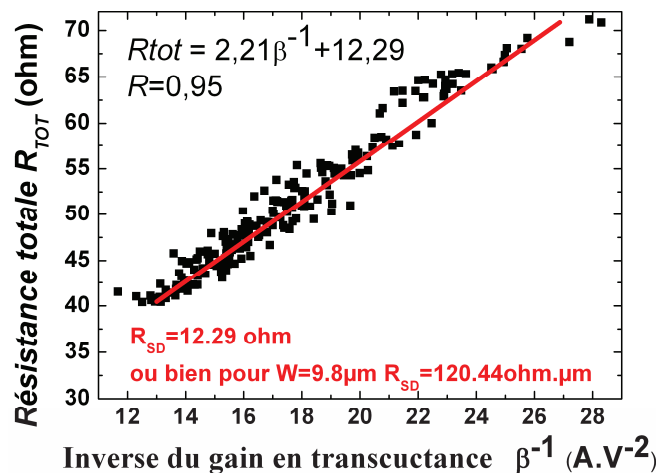


Figure C.20: Extraction de la résistance source-drain dans les dispositifs Ge-pMOSFETs

C.II.1.b) Les courants de fuites dans le germanium

Les courants de fuite dans un MOSFET proviennent (cf. [figure C.21](#)) :

- du courant thermoïonique au-dessus de la barrière source-canal (I_{th}),
- des fuites de jonction des diodes sources-substrat et drain-substrat (I_{junc}),
- Du courant de tunnel de la grille (I_g) à travers le diélectrique,
- du courant de tunnel direct de la source au drain (I_{STD}), (dans le cas de dimensions très agressives ($L < 10$ nm),
- et du courant Band-To-Band-Tunneling (I_{BTBT}) dans les zones à fort champ électrique (près du drain).

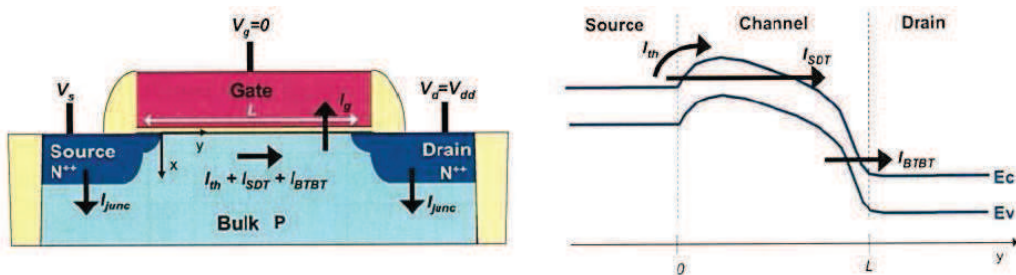


Figure C.21: Illustration des courants de fuite dans un nMOSFET bulk: vue schématique (à gauche) et structure de bande (à droite) [[Batail09](#)]

Le courant tunnel bande à bande ou en Anglais le band to band tunneling (BTBT) est l'une des limitations les plus critiques des transistors Ge à cause du petit gap du germanium comparé au silicium.

Le BTBT peut apparaître en polarisation inverse d'une jonction PN. Lorsque le champ à la jonction augmente, la bande de valence (BV) de la région dopée P (le canal par exemple) peut devenir plus haute que la bande de conduction (BC) de la région dopée N (le drain par exemple).

Conséquemment, les électrons de la région N auront tendance à vouloir passer par effet tunnel dans la région P à travers la barrière : c'est le Band-To-Band Tunneling (BTBT). Les trous de la zone P peuvent aussi subir le même effet et passer dans la zone N. Ce phénomène apparaît à la jonction drain-substrat pour des fortes tensions de drain et génère un courant de fuite au niveau du substrat. Pour que les porteurs puissent passer d'un coté à l'autre dans la jonction, la chute de potentiel doit être suffisante : $V > \frac{E_g}{q} - \Phi_d$ ou E_g est le gap et Φ_d la barrière de potentiel. Cette condition est nécessaire mais pas suffisante pour observer le BTBT. La hauteur de barrière de tunnel et la largeur de tunnel doivent être suffisamment faibles cf. [figure C.22](#).

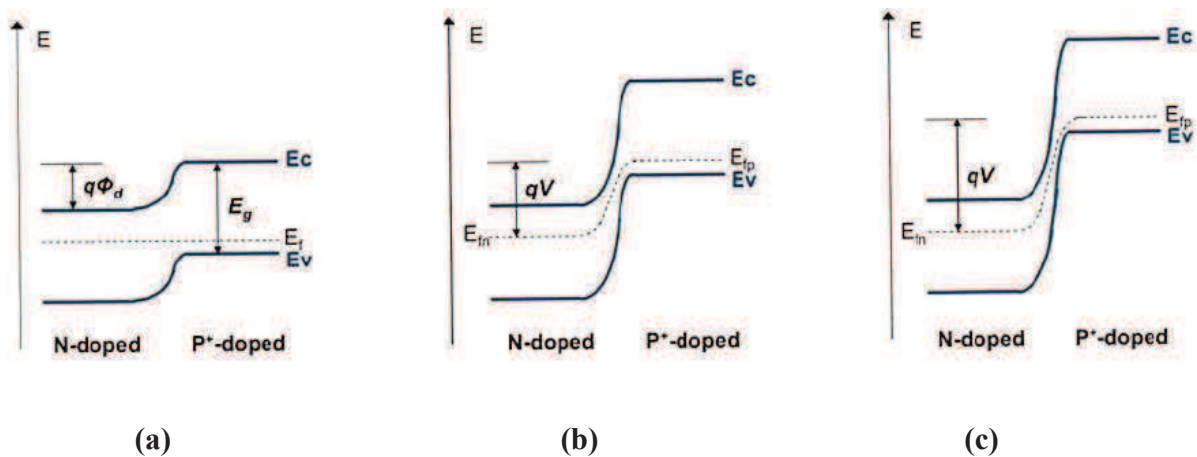


Figure C.22: Coupe d'une structure de bande d'une jonction PN à l'équilibre : pas de courant de tunnel possible (a) ; Coupe d'une structure de bande d'une jonction PN en polarisation inverse modérée : la largeur de tunnel est assez importante pour empêcher les porteurs de passer par effet tunnel (b) ; Coupe d'une structure de bande d'une jonction PN en forte polarisation inverse : Les porteurs ont de forte chance de passer d'une zone à l'autre (c)

[Batail09].

Quand les porteurs transitent par effet tunnel d'une bande à une autre, la transition peut être directe ou indirecte selon que le gap soit direct ou indirect. Dans les matériaux comme le *GaAs*, la transition sera directe (matériau à gap direct). Par contre dans le silicium ou le germanium la transition sera indirecte. Cependant une transition indirecte peut être assistée par des phonons ou

par des impuretés qui modifient le vecteur k . Dans ce cas, si le bas de la vallée de conduction est proche du minimum de la BC au point Γ , la transition directe est favorisée en particulier à forte tension inverse. Par conséquent, même si le gap est indirect, le passage par effet tunnel se fera d'une manière directe de la BV à la BC : C'est le cas du germanium (cf. [figure C.23](#)).

En plus du tunnel direct dans la bande interdite, les niveaux profonds des centres de génération recombinaison peuvent assister la transition de la BV à la BC. Ce phénomène est appelé effet tunnel assisté par piégeage (ou TAT pour trap assisted tunneling). Le BTBT et TAT sont très présents dans le germanium. Dans les structures MOS germanium le courant de fuite n'est plus dû à la diffusion mais plutôt aux fuites par tunneling

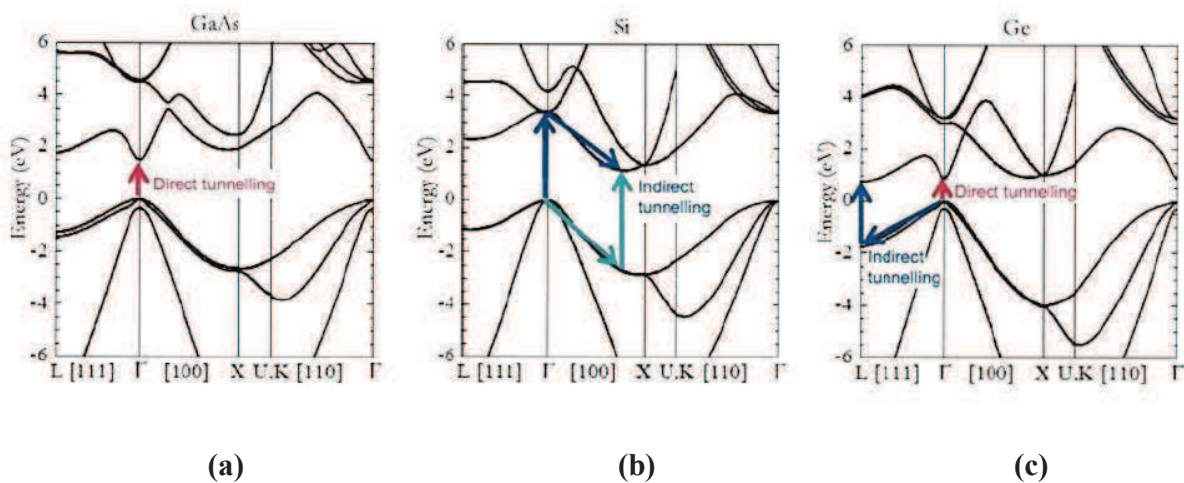


Figure C.23: Effet tunnel direct dans des semiconducteurs à gap direct tels que le GaAs (a) ; dans des semiconducteurs à gap indirect tels que le Si (b) ; effet tunnel direct ou indirect dans des semiconducteurs à gap indirect tels que le Ge (c) [Richard04]

Toutes les mesures qui suivent sont faites sur du Ge-pMOSFET, les valeurs de la tension de grille sont prises en valeur absolue pour faciliter l'extraction.

Comme montré sur la [figure C.24a](#), le courant de bulk I_B est faible dans le régime linéaire. Cependant dans le régime de saturation ([figure C.24b](#)), le BTBT devient important et joue un rôle non négligeable dans la courant I_{OFF} . Dans ce régime, le courant I_D doit être corrigé avant de

l'utiliser pour l'extraction de paramètres. Par exemple, pour une extraction claire du DIBL sous le seuil à courant constant, le courant de saturation doit être bien corrigé du BTBT.

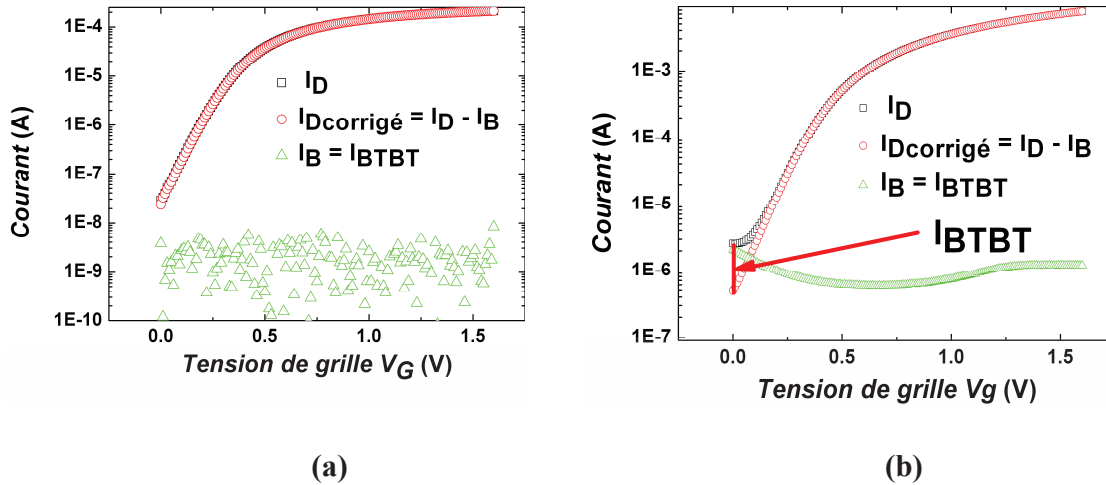


Figure C.24: Courant de drain, courant de drain corrigé et courant de bulk en régime linéaire ($V_D=10\text{mV}$) (a) et en saturation ($V_D=1\text{V}$) (b) ($L=105\text{ nm}$)

C.II.2) Vérification des propriétés de transport sur nos dispositifs Ge

C.II.2.a) Mobilité effective

Des mesures de mobilité split CV sont faites sur des longs transistors. Comme attendu pour le germanium, grâce à sa plus faible masse effective pour les trous, de meilleures propriétés de transport sont à noter comparé au silicium. Ceci est confirmé sur la [figure C.25](#) où de plus grandes valeurs de mobilité de transistors longs Ge-pMOSFET sont enregistrées comparées à la mobilité universelle sur Si. Des mesures statistiques sont faites sur plusieurs puces pour vérifier la fiabilité du résultat. L'erreur relative liée à la variation puce à puce du pic de mobilité est de l'ordre de 4%. Conséquemment, la question de savoir si ces grandes valeurs de mobilité peuvent engendrer de la quasi-balisticité sur des transistors Ge ultra-courts se pose.

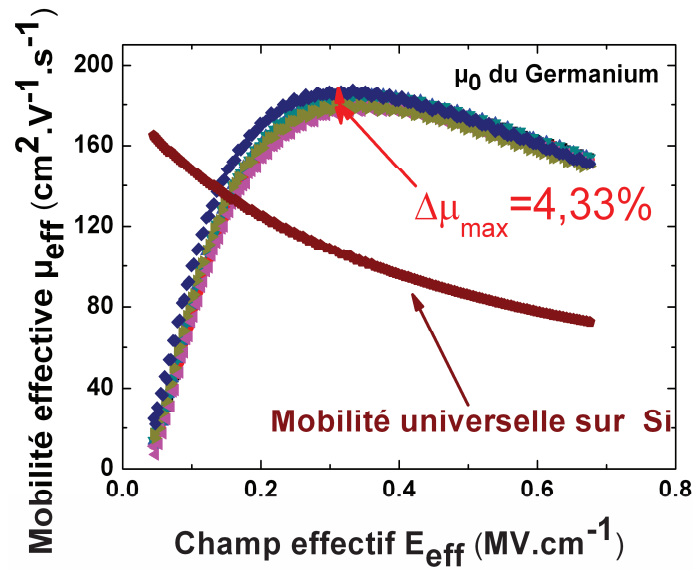


Figure C.25: Mobilité split CV sur Ge comparée à la mobilité universelle sur Si

C.II.2.b) Estimation de la balisticité dans le Ge en régime linéaire

La détermination du taux de balisticité en régime saturé étant très difficile, nous le ferons uniquement en régime linéaire. Cela donnera une idée quant à la balisticité attendue dans nos dispositifs germanium. Pour cela, on détermine le coefficient de rétrodiffusion r compris entre 0 et 1. Dans le cas où r tend vers 1, La majorité des porteurs qui franchit la barrière de potentiel à la source rebrousse chemin et n'atteignent pas le drain : c'est le pur régime dérive diffusion. Dans le cas opposé où $r = 0$, tous les porteurs qui franchissent la barrière de potentiel à la source atteignent le drain : c'est le régime balistique. Le régime quasi-balistique est compris entre ces deux cas ($0 < r < 1$). Dans [Pappas09], le calcul de r est donné par :

$$r = 1 - \frac{\mu_{\text{exp}}}{\mu_{\text{bal}}} \quad (\text{C.17})$$

Dans cette expression, μ_{exp} (figure C.26) est la mobilité expérimentale donnée par :

$$\frac{1}{\mu_{\text{exp}}} = \frac{1}{\mu_{\text{bal}}} + \frac{1}{\mu_{\text{dd}}} \quad (\text{C.18})$$

La mobilité balistique μ_{bal} (figure C.26) est donnée par [Shur02-Cros06a]:

$$\mu_{bal} = \frac{q \cdot v_T \cdot L}{2 \cdot k \cdot T} \quad (C.19)$$

T est la température, k la constante de Boltzmann, q la charge de l'électron et v_T la vitesse thermique évaluée à $3.2 \cdot 10^7 \text{ cm.s}^{-1}$ pour les trous dans le germanium [<http://nanohub.org/>].

La mobilité de dérive diffusion μ_{dd} , corrigée des effets balistiques, peut donc être obtenue à partir des relations (C.18) et (C.19) et permet d'extraire la composante purement dérive diffusion du transport en fonction de la longueur de canal (figure C.26).

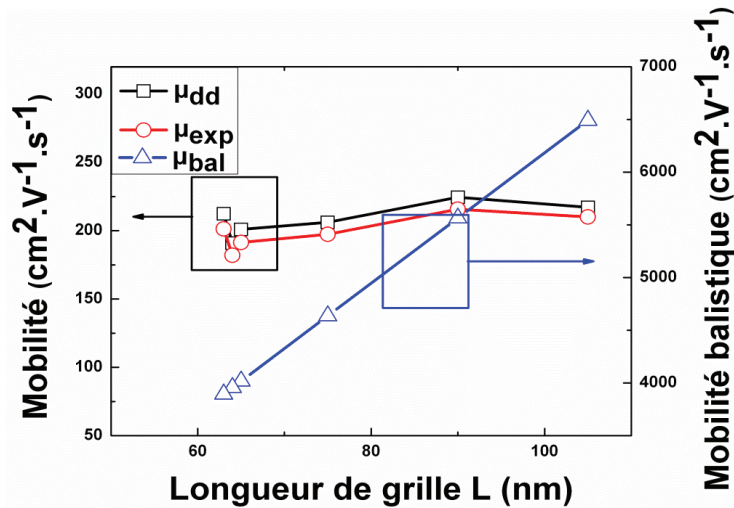


Figure C.26: Illustration de la mobilité dérive diffusion, expérimentale et balistique

Comme illustré sur la figure C.27 le coefficient de rétrodiffusion r et le taux de balisticité $1-r$ sont déterminés en utilisant (C.17). On peut remarquer que pour des longueurs allant de 105nm à 63nm, le coefficient de rétrodiffusion est légèrement inférieur à 1. Le meilleur taux de balisticité est obtenu sans surprise pour $L = 63\text{nm}$ et est autour de 5.5%, ce qui est donc très faible pour ces dispositifs. Nous allons maintenant caractériser le transport en régime de saturation pour confirmer ou infirmer ces prédictions en régime linéaire.

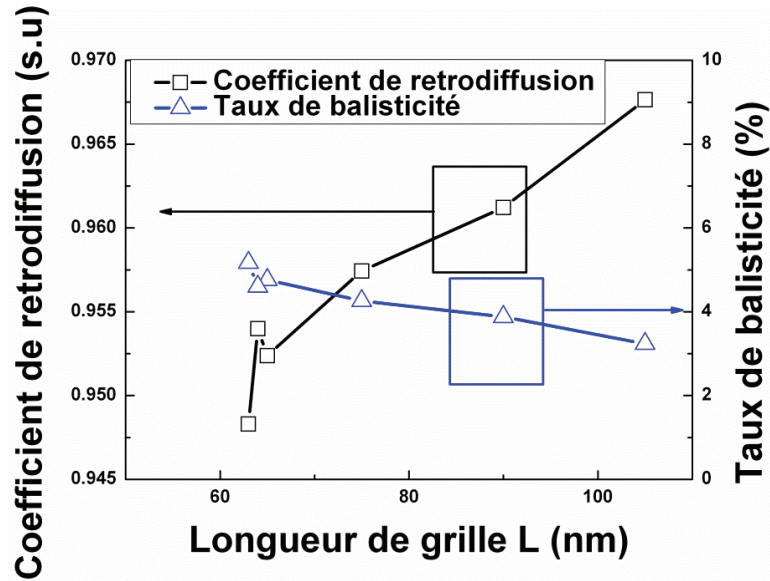


Figure C.27: Coefficient de rétrodiffusion et taux de balisticité en fonction de la longueur de grille

C.II.3) La méthode de la vitesse limitante: le v_{lim}

Le concept appelé mobilité apparente est considéré pour définir un paramètre de même dimension que la mobilité en linéaire mais en régime saturé. Ce concept aide aussi à unifier le régime linéaire et saturé comme dans [Fleury09b].

Considérant le courant en linéaire (I_{lin}), le courant en saturation (I_{sat}), la tension de drain (V_{DS}), la tension de grille de drive (V_{GTsat}) et la charge d'inversion Q_s , la mobilité linéaire (μ_{lin}) et la mobilité apparente en saturation (μ_{sat}) peuvent être exprimées comme suit :

$$\frac{\mu_{lin}}{L} = \frac{I_{lin}}{W \cdot Q_s \cdot V_{DS}} \quad (C.20) \quad \frac{\mu_{on}}{L} = \frac{2 \cdot I_{on}}{W \cdot Q_s \cdot V_{GTsat}} \quad (C.21) \quad \text{avec } V_{GTsat} = V_G - V_{THsat}$$

(C.22)

En utilisant les lois de Mathiessen évoquées dans le chapitre II de la partie A, et en combinant la mobilité linéaire (C.20) et la mobilité apparente en saturation (C.21), on obtient une expression de la vitesse limitante suivante :

$$v_{\text{lim}} = \frac{1}{2} \left[\frac{\partial}{\partial V_{\text{GTsat}}} \left(\frac{L}{\mu_{\text{app}}^{\text{sat}}} - \frac{L}{\mu_{\text{app}}^{\text{lin}}} \right) \right]^{-1} \quad (\text{C.23})$$

Ainsi le tracé de $\Delta = \left(\frac{L}{\mu_{\text{app}}^{\text{sat}}} - \frac{L}{\mu_{\text{app}}^{\text{lin}}} \right)$ en fonction de V_{GTsat} permet d'accéder assez aisément à la vitesse limitante v_{lim} .

C.II.4) Amélioration de la méthode de « v_{lim} »

Une difficulté apparaît sur l'extraction de Q_s et de V_{GTsat} dans l'expression (C.19). En effet, Q_s est approximé par $Q_s = C_{\text{ox}} \cdot V_{\text{GTsat}}$, l'expression (C.21) devient donc $\frac{\mu_{\text{on}}}{L} = \frac{2 \cdot I_{\text{on}}}{W \cdot C_{\text{ox}} \cdot V_{\text{GTsat}}^2}$ où le calcul du paramètre V_{GTsat} est problématique à cause de l'incertitude sur le paramètre V_{THsat} qui est la tension de seuil en saturation. Dans [Fleury09b], V_{THsat} est extrait en utilisant la méthode du courant constant comme illustré sur la figure C.28. Grâce à la fonction Y décrite au chapitre I de cette partie, on extrait les paramètres en régime linéaire comme V_{THlin} , R_{SD} ,...etc. R_{SD} servira pour la correction de V_G dans l'expression (C.22) : $V_{\text{G,cor}} = V_G - R_{\text{SD}} \cdot I_{\text{Dsat}}$.

C.II.4.a) Extraction de la tension de seuil en saturation

Avec le V_{THlin} , on remonte à V_{THsat} grâce au DIBL calculé comme illustré sur la figure C.28. Cette manière de calculer le V_{THsat} en utilisant le DIBL sous le seuil n'est pas précise et peut induire des incertitudes [Giusi10] sur l'extraction.

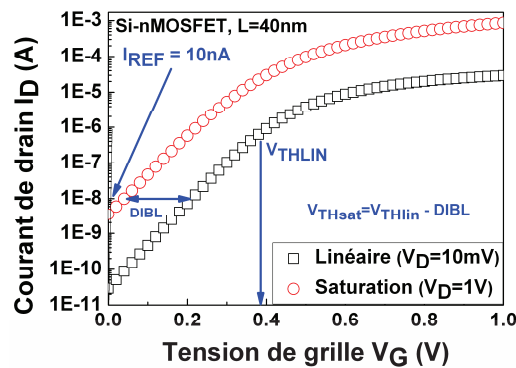


Figure C.28: Procédure d'extraction de V_{THsat} utilisant le DIBL sous le seuil à courant constant

Par conséquent, pour minimiser l'erreur commise sur l'extraction de V_{THsat} , et donc de Q_s et conséquemment sur la mobilité apparente en saturation, nous avons décidé d'extraire le V_{THsat} directement en utilisant la méthode de la dérivée seconde du courant de drain (figure C.29). Il faut noter que les deux dérivées successives amènent aussi des erreurs sur le résultat.

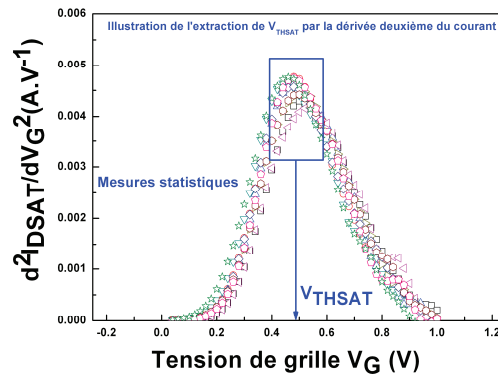


Figure C.29: Illustration de la procédure d'extraction de V_{THsat} en utilisant la dérivée seconde du courant de drain en saturation

Cette méthode d'obtention du DIBL ($V_{THlin} - V_{THsat}$) donne des valeurs plus faibles comparées à celles de la méthode du courant constant. On peut voir sur la figure C.30 que cette différence entre les deux méthodes d'extraction n'est pas négligeable et peut engendrer une erreur significative sur la détermination de la charge d'inversion.

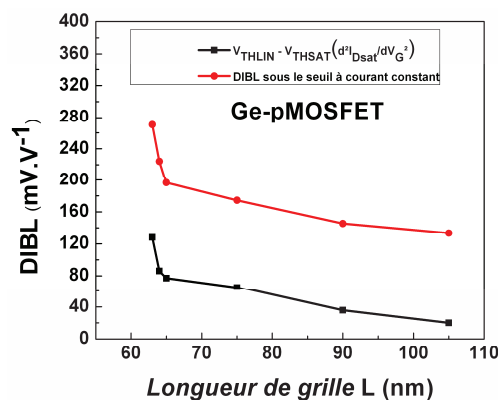


Figure C.30: Comparaison du DIBL extrait par la méthode du courant constant et la méthode de la dérivée seconde du courant de saturation

C.II.4.b) Confirmation de la méthode d'extraction de V_{THsat} par des simulations TCAD

Le DIBL qui correspond à la variation $\frac{\partial V_G}{\partial V_D}$ peut être simulé en utilisant le minimum de potentiel de surface ψ_{min} (figure C. 31a, 31b & 31c). Le minimum de potentiel simulé peut être écrit comme suit:

$$\frac{\partial \psi_{min}}{\partial V_D} = \frac{\partial \psi_{min}}{\partial V_G} \cdot \frac{\partial V_G}{\partial V_D} \quad (C.24)$$

La pente du tracé ψ_{min} en fonction de V_G et ψ_{min} en fonction de V_D permettra de calculer le DIBL (figure C.32a, C.32b & C.33a). Les simulations permettent aussi d'extraire le DIBL sous le seuil à courant constant et le DIBL obtenu en considérant le maximum de la dérivée seconde du courant. La comparaison à la figure C.32b montre clairement un accord entre la méthode qui utilise le minimum de potentiel et celle utilisant la dérivée seconde du courant de drain en saturation. Cela valide notre choix d'utiliser la dérivée seconde du courant de drain en saturation pour calculer le V_{THsat} servant lui-même à calculer la charge d'inversion.

CHAPITRE C.II : A LA RECHERCHE DE BALISTICITE DANS LE GERMANIUM : LA METHODE DE LA « VITESSE LIMITE »

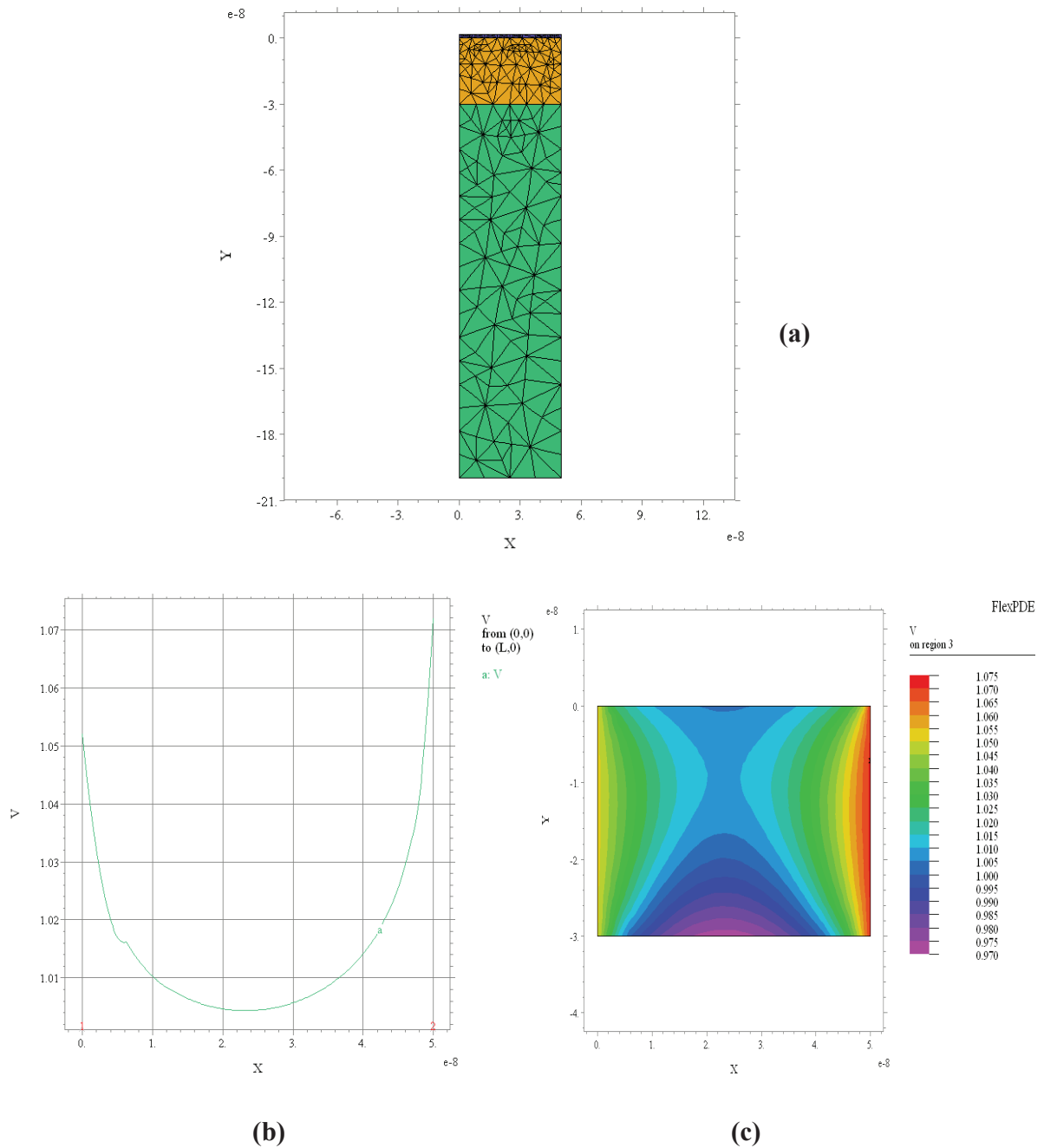
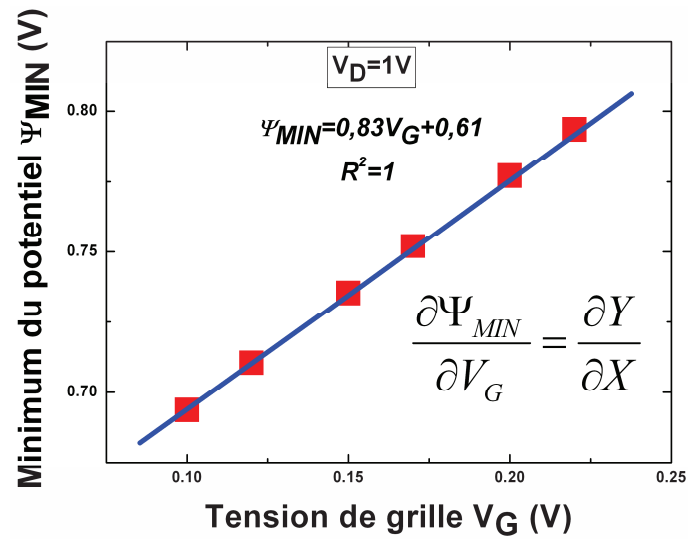
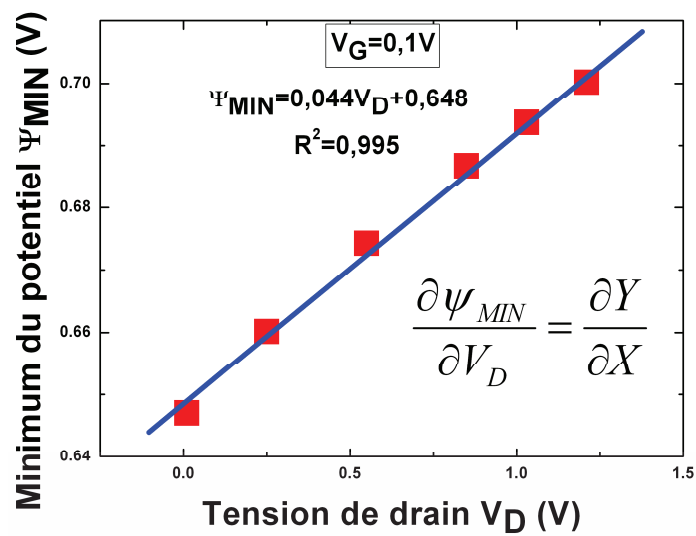


Figure C.31: maillage du transistor (a) ; potentiel à la surface du canal (b), potentiel dans le volume du canal (c)

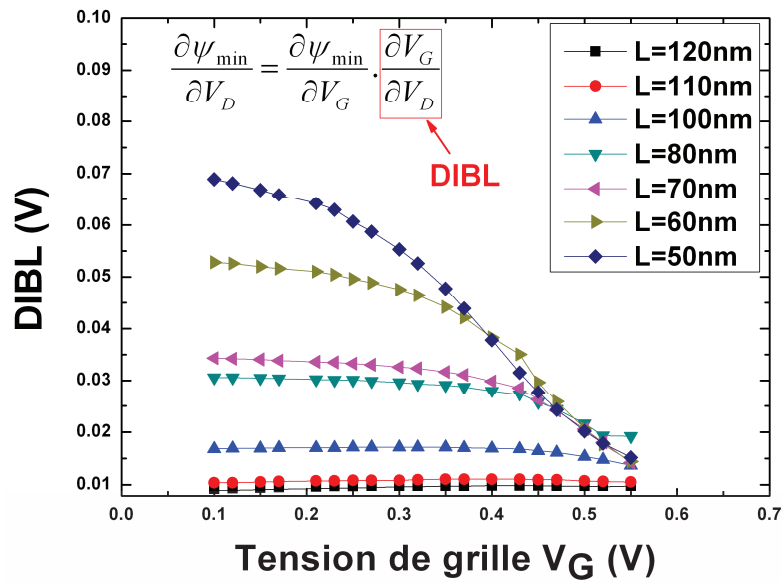


(a)

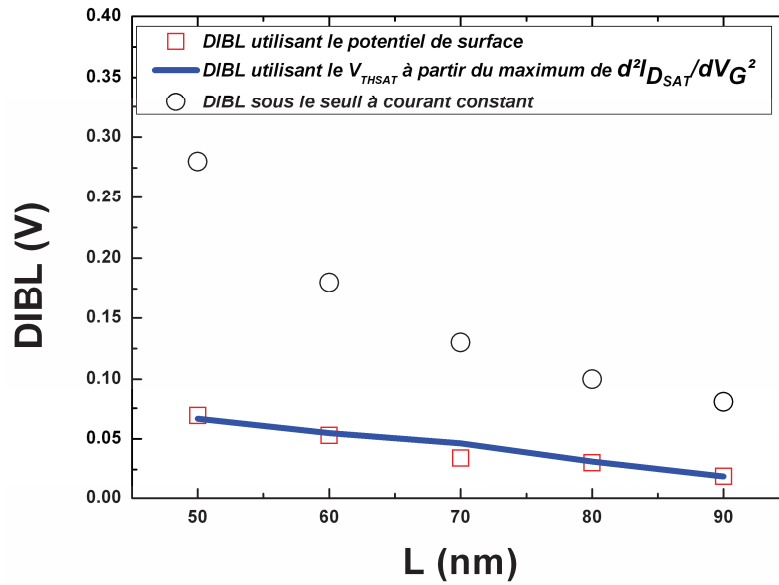


(b)

Figure C.32: Minimum du potentiel de surface en fonction de la tension de Grille (a) et de la tension de drain (b)



(a)



(b)

Figure C.33: DIBL extrait via le minimum de potentiel de surface (a) et comparaison des trois méthodes de calcul du DIBL (b)

C.II.5) Application de la méthode de « v_{lim} » sur Ge-pMOSFET

Pour déterminer le régime de transport dans nos dispositifs germanium, nous étudions la dépendance en température de la vitesse limitante. Son comportement nous permettra de discriminer la limitation de vitesse entre la vitesse d'injection (régime quasi-balistique) et la vitesse de saturation (régime de saturation).

Pour le régime de saturation usuel, la vitesse limitante est la vitesse de saturation qui est principalement affectée par les interactions avec les phonons (phonon scattering). Les collisions avec les phonons augmentant avec la température, dans ce cas la vitesse limitante diminuera avec la température.

Pour ce qui est du transport quasi-balistique, la vitesse limitante est la vitesse d'injection. Cette vitesse augmente avec la température car étant liée à la vitesse thermique, elle-même liée à l'énergie cinétique [Jacoboni77-Zilli07].

C.II.5.a) Détermination de la nature du transport

Les mesures de courant I_D - V_G sont faites avec un prober automatique 200 mm (Accretech APM90) équipé d'une matrice de connexion type Agilent 4073B. Cela permet une combinaison de plusieurs plots de mesures. Pour le traitement, un programme scilab est utilisé pour permettre un traitement statistique (36 puces), évitant l'incertitude due à la dispersion stochastique. Les barres d'erreur représentent la déviation standard. La longueur de grille varie de 415 nm jusqu'à 63 nm. Les études sont faites dans une gamme de température qui varie de 398 K à 243 K.

La **figure C.34** montre la procédure d'extraction de v_{lim} . L'inverse de la pente de

$\Delta = \left(\frac{L}{\mu_{app}^{sat}} - \frac{L}{\mu_{app}^{lin}} \right)$ en fonction de V_{GTsat} donne la vitesse limitante v_{lim} (cf. **équation C.23**).

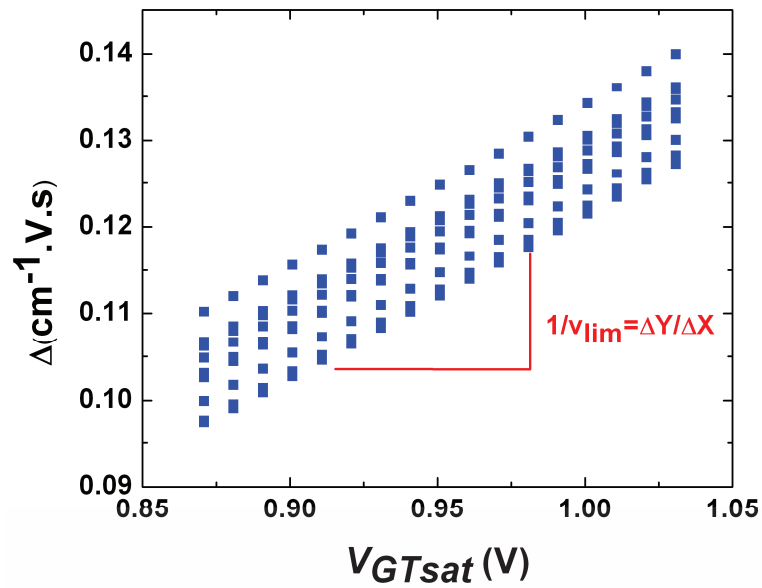


Figure C.34: Procédure d'extraction de la vitesse limitante

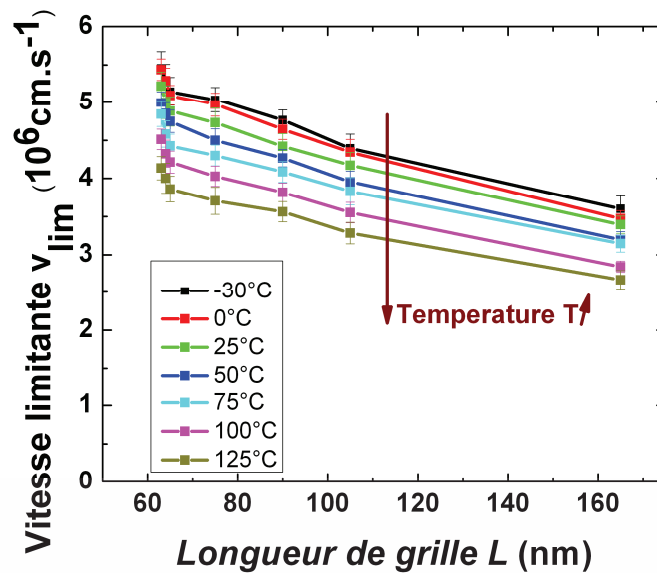


Figure C.35: Vitesse limitante en fonction de la longueur de grille pour différentes températures

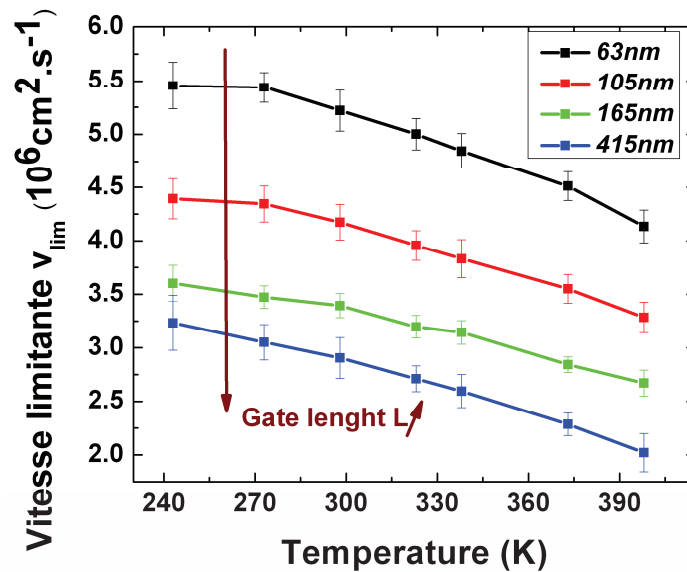


Figure C.36: Vitesse limitante en fonction de la température pour différentes longueurs de grille

Le comportement observé sur les figures C.35 & C.36 est caractéristique d'une saturation du transport en champ électrique même pour les courts transistors.

Comme la vitesse limitante diminue lorsqu'on augmente la température, nous concluons que le transport à fort V_D est limité par les phonons. Ceci est en adéquation avec les prédictions faites à la partie II.2.b en régime linéaire.

Cependant plus d'attention à la figure C.36 montre une petite différence mais non négligeable. En effet, le v_{lim} des transistors plus courts tend à se stabiliser lorsqu'on diminue la température, tandis que les transistors longs montrent une dépendance plus claire.

En définitive, même pour les transistors Ge-pMOSFETs aussi courts que 60 nm, il n'existe pas de quasi-balisticité évidente. Cependant la stabilisation de la tendance à basse température pour les plus courts transistors laisse présager que du transport quasi-balistique apparaîtrait en dessous de 200K. On serait cependant loin des températures d'utilisation des transistors qui avoisinent les

300K. L'autre solution pour espérer une inversion de la tendance de v_{lim} en fonction de T serait de fabriquer des transistors Ge-pMOSFETs plus courts que 60 nm.

C.II.6) Bilan et conclusion

Le transport dans des transistors germanium avec un procédé de fabrication optimisé a été caractérisé par la méthode de la vitesse limitante. Des améliorations telles que l'extraction de la tension de seuil sont apportées à la méthode. Des simulations TCAD ont permis de valider la manière d'extraire V_{THsat} par la dérivée seconde du courant.

Malgré les bonnes propriétés de transport du germanium, nous avons montré que la limite quasi-balistique n'était pas encore atteinte sur des transistors aussi courts que 63 nm et pour une gamme de température allant de 415 K à 243 K. Le transport reste donc gouverné par les collisions sur les phonons. Il faut cependant noter que la limite pourrait être atteinte pour des températures inférieures à 200K ou pour des transistors beaucoup plus courts que 60 nm dans le germanium.

Dans cette étude, on a utilisé le V_{THsat} pour approximer la charge d'inversion Q_s en saturation. En régime linéaire, la charge d'inversion est obtenue en intégrant la capacité. En régime de saturation, les mesures de capacité ne sont pas possibles en DC ou à basse fréquence. Nous verrons au chapitre suivant comment nous pouvons utiliser des mesures hautes fréquences (HF) pour mesurer directement la charge d'inversion en saturation, sans passer par l'extraction « pénible » de V_{THsat} .

CHAPITRE C.III : MESURES HF POUR LE TRANSPORT ELECTRONIQUE

C.III.1) Contexte

Avec la miniaturisation, l'architecture du transistor ne cesse d'évoluer. L'architecture standard qui constitue à réaliser un transistor sur du silicium massif ne suffit plus pour contrecarrer les effets néfastes de la miniaturisation. On assiste ainsi à l'émergence de nouveaux types de MOSFETs comme le FDSOI, le FINFET ou les composants à base de germanium. Ainsi l'obtention d'informations dans les conditions d'opération ($V_{dd} = 1V$) du MOSFETs est un sujet plus qu'alléchant. La mobilité des porteurs (μ_{eff} , μ_0) dans le canal est généralement extraite en régime linéaire. Dans le régime de saturation du transistor, la mobilité perd tout son sens. La vitesse moyenne v_{av} peut donner des informations dans ce régime mais son extraction y est imprécise à cause de la difficulté (voire l'impossibilité) liée au calcul de la charge d'inversion Q_{inv} . En effet, à faible V_D (régime linéaire) Q_{inv} est obtenu en intégrant par rapport à V_G la capacité d'inversion C_{GC} comme illustré à la **figure C.37**. La mesure de la capacité à fort V_D est difficilement réalisable, entraînant une impossibilité d'obtenir la valeur exacte de Q_{inv} .

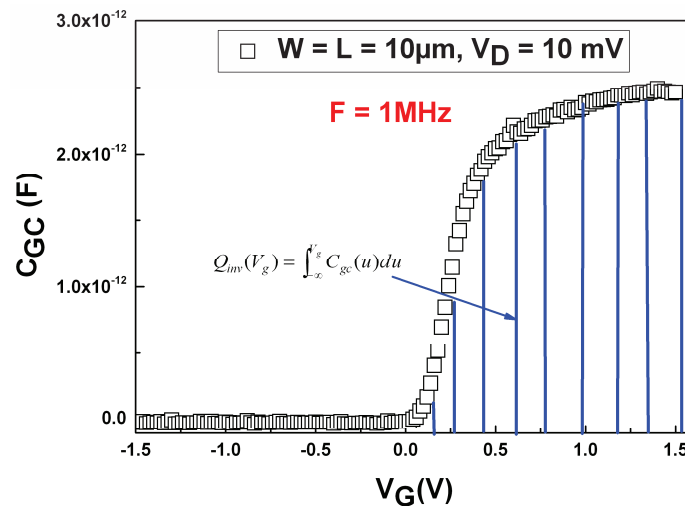


Figure C.37: Méthodologie d'obtention de Q_{inv} en régime linéaire

Ce chapitre présente une procédure qui a permis de réaliser des mesures de capacité en régime de saturation jusqu'à 10 GHz sur des transistors ultra-courts pour extraire d'une manière fiable la charge d'inversion et la vitesse moyenne des porteurs [Diouf13c]. Les taux de balisticité obtenus

en linéaire et en saturation ne sont pas négligeables et mettent en évidence le transport non stationnaire.

C.III.2) Problématique

La caractérisation en régime de saturation de transistors avec des matériaux à haute mobilité est redevenue un sujet d'actualité dans le monde de la microélectronique. Le régime de transport balistique est en effet plus que jamais attendu dans ces types de matériaux. Etant donné que la mobilité n'a pas de sens en régime de saturation, la solution est d'améliorer l'extraction expérimentale de la vitesse moyenne v_{av} (C.25) en améliorant l'extraction de la charge d'inversion Q_{inv} . Au chapitre II de cette partie, une amélioration de l'extraction de Q_{inv} via l'extraction de la tension de seuil en saturation V_{THSAT} (C.26) a été présentée (en utilisant la dérivée deuxième du courant de saturation). A défaut d'une extraction directe comme en régime linéaire, (26) est utilisée pour approximer Q_{inv} . Il faut rajouter à cette imprécision le besoin d'extraire V_{THSAT} d'une manière fiable. Nous avons vu que malgré tous ces efforts, le calcul de Q_{inv} reste imprécis.

$$v_{av} = \frac{I_{DSAT}}{W \cdot Q_{inv}} \quad (C.25) \quad Q_{inv} = C_{ox} \cdot (V_G - V_{THSAT}) \quad (C.26)$$

Une manière de mesurer Q_{inv} sans ambiguïté est de mesurer la capacité d'inversion directement à fort V_D .

C.III.3) Solutions techniques

C.III.3.a) Mesures basses fréquences

La relation menant au facteur de qualité Q d'une mesure de capacité est donnée en (C.27).

$$Q = 2 * \Pi * f * C_p * R \quad (C.27)$$

Comme C_p diminue au fur et à mesure que la longueur de grille diminue, le facteur de qualité diminue et la mesure de la capacité est négativement affectée. Lors de l'application d'une forte tension V_D , C_p est plus faible et la partie DC du signal rend difficile la mesure de la capacité.

Une manière de mesurer facilement la capacité sur des transistors courts à faible V_D est d'en connecter plusieurs en parallèle [Fleury08b]. Une augmentation très importante du nombre de transistors connectés, et donc de la surface de grille, est souhaitable pour des mesures à V_{DS} nul.

Toutefois, pour des mesures à V_{DS} fort, le courant de drain dépasse rapidement la gamme des appareils de mesure. Une autre procédure est suivie dans [Romanjek04] pour des mesures à basse fréquence ($<1\text{MHz}$) en linéaire sur des transistors individuels, mais celle-ci reste longue et à la limite de la précision des appareils, ce qui la rend difficile à utiliser en saturation.

Lors de l'application d'une forte tension V_D la méthode RFCV peut être utilisée avec un analyseur de réseau pour mesurer la capacité comme dans [Pantisano08]. La mesure faite à 100 MHz est cependant bruitée et difficile à exploiter pour les transistors courts en particulier.

Pour éviter la perturbation du signal RF par la partie DC à fort V_D , nous avons inséré deux « Bias-T » entre la source, le drain et la sortie Low (application d'une faible tension) du capacimètre comme illustré sur la figure C.38. Ce montage permet de ne récupérer que la partie RF du signal au niveau du capacimètre.

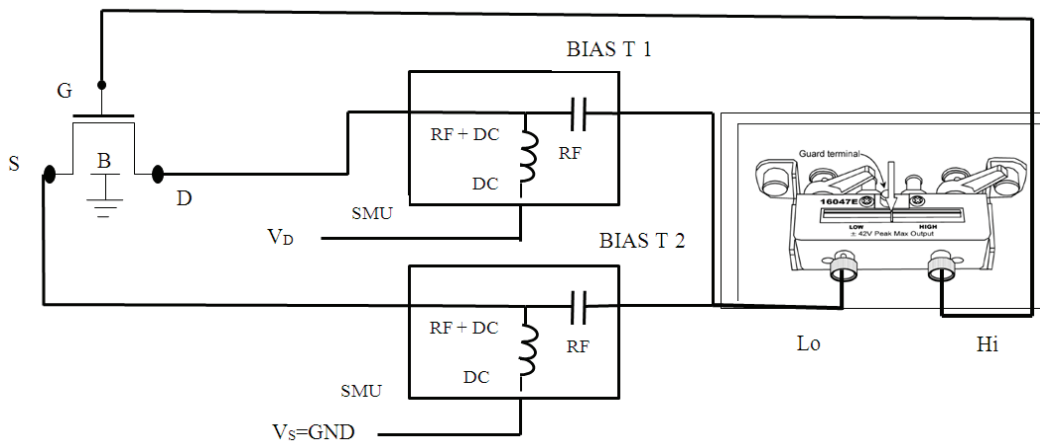


Figure C.38: Procédure de mesure de capacité à fort V_D sur des transistors longs

De cette manière, la capacité grille-canal C_{GC} et conséquemment la charge d'inversion Q_{inv} ont pu être mesurées en fonction de V_G à différents V_D sur des transistors longs comme montré sur les figures C.39a & C.39b.

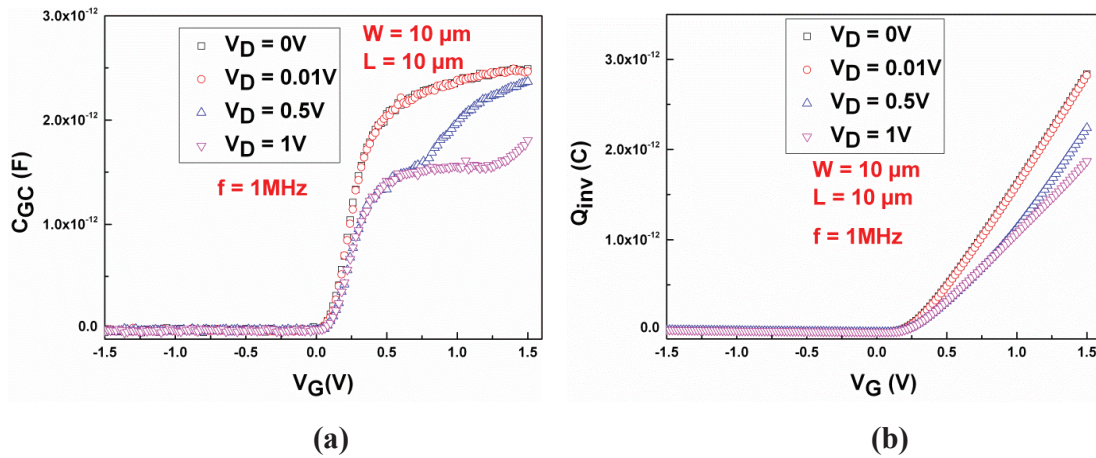


Figure C.39: Capacité grille-canal (a) et charge d’inversion (b) en fonction de V_G à différents V_D pour des transistors longs

Si l’insertion des « Bias-T » a suffi pour mesurer C_{GC} en régime de saturation sur des transistors longs, elle n’a pas suffi pour mesurer C_{GC} sur des transistors courts à fort V_D à cause de la faible valeur de C_p dans (C.27).

C.III.3.b) Mesures hautes fréquences (HF)

Une manière naturelle d’améliorer le facteur de qualité Q est d’augmenter la fréquence f dans (C.27). Les mesures de C_{GC} à fort V_D sur un transistor long de 250 nm sont identiques à 100 MHz et à 10 GHz (figure C.40a). Cependant, sur un transistor très court, comme précédemment fait dans la littérature [Pantisano08], la mesure à 100 MHz est trop bruitée pour être exploitable tandis que celle à 10 GHz reste très précise et plus facilement exploitable (figure C.40b). Ces observations justifient le besoin d’augmenter la fréquence pour obtenir des mesures fiables de capacité à fort V_D .

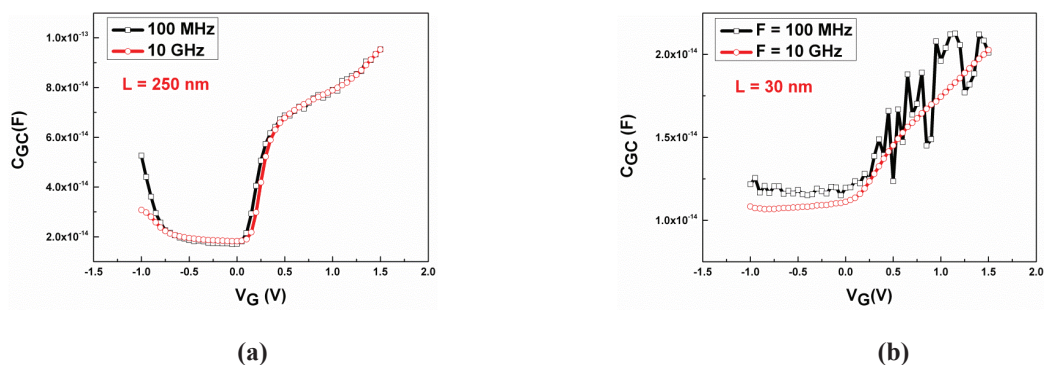


Figure C.40: Mesures de C_{GC} à fort V_D à 100 MHz et 10 GHz pour $L=250$ nm (a) et $L=30$ nm (b)

Nous allons maintenant décrire la procédure qui a été développée pour obtenir ce résultat à partir d'une structure deux ports. Le schéma équivalent est donnée en [figure C.41a](#). Dans cette structure, source et substrat sont connectés ensemble puis connectés à la masse, la grille en port 1 et le drain en port 2 comme on peut le voir sur les [figures C.41a & C.41b](#). Ce type de schéma est aussi appelé structure à source commune.

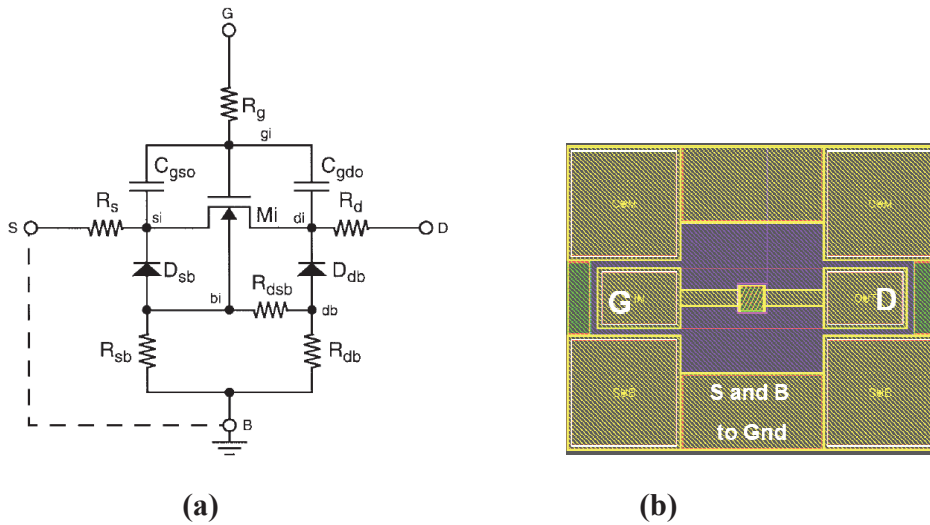


Figure C.41: Schéma équivalent de la structure deux ports utilisée **(a)**; coupe schématique de la structure deux ports avec source et substrat connectés à la masse **(b)**

A partir des paramètres Y (Y_{11} et Y_{12}), la capacité grille-source C_{GS} , la capacité grille-drain C_{GD} et la capacité grille-canal C_{GC} sont obtenues en utilisant [\(C.28\)](#) à [\(C.30\)](#).

$$C_{GS} = \frac{\text{imag}(Y_{11} + Y_{12})}{2\pi \cdot f} \quad (\text{C.28})$$

$$C_{GD} = \frac{-\text{imag}(Y_{12})}{2\pi \cdot f} \quad (\text{C.29})$$

$$C_{GC} = C_{GS} + C_{GD} \quad (\text{C.30})$$

➤ Définition des paramètres Y

Dans le cas d'une configuration deux ports, le transistor en HF est souvent représenté par un quadripôle. Ce quadripôle illustré à la [figure C.42](#) est représenté par des relations matricielles. Des matrices de transfert sont reliées ainsi aux entrées et sorties du quadripôle. Parmi ces matrices de transfert on distingue les paramètres Y notamment.



Figure C.42: Illustration d'un quadripôle avec des entrées-sorties en courant et tension

Les paramètres Y sont reliés aux entrées-sorties par les relations suivantes :

$$\text{Paramètres } Y \left\{ \begin{array}{l} i_1 = Y_{11} \cdot v_1 + Y_{12} \cdot v_2 ; \left\{ \begin{array}{l} Y_{11} = \frac{i_1}{v_1} \Big|_{v_2=0} ; Y_{12} = \frac{i_1}{v_2} \Big|_{v_1=0} \\ Y_{21} = \frac{i_2}{v_1} \Big|_{v_2=0} ; Y_{22} = \frac{i_2}{v_2} \Big|_{v_1=0} \end{array} \right. \\ i_2 = Y_{21} \cdot v_1 + Y_{22} \cdot v_2 ; \end{array} \right. \quad (C.31)$$

La mise en circuit ouvert ou fermé de l'entrée ou de la sortie du quadripôle est indispensable pour mesurer les éléments de ces matrices avec précision.

Cette précision devient difficile à atteindre au-delà de 100MHz pour des raisons telles que :

- ✓ les circuits ouverts ou fermés ne sont pas idéaux en haute fréquence.
- ✓ des effets de propagation peuvent intervenir dès que les dimensions en jeu sont grandes devant la longueur d'onde du signal.

Il est donc nécessaire d'introduire une nouvelle matrice utilisant des paramètres mesurables et en remplaçant les circuits ouverts et fermés par des ports adaptés généralement à 50Ω [Gilmour86].

Cette matrice est appelée la matrice des paramètres S.

➤ Définition des paramètres S

Ces paramètres sont reliés aux ondes de puissance transmises ou réfléchies lorsqu'un dispositif tel que le transistor est inséré sur une ligne de transmission [Negre11]. On peut relier les paramètres S aux paramètres Y nécessaires à l'obtention des capacités (Y_{11} et Y_{12}) par les relations C.32 et C.33.

$$Y_{11} = \frac{(1 - S_{11}) \cdot (1 + S_{22}) + S_{12} \cdot S_{21}}{(1 + S_{11}) \cdot (1 + S_{22}) - S_{12} \cdot S_{21}} \quad (C.32) \qquad Y_{12} = \frac{-2 \cdot S_{12}}{(1 + S_{11}) \cdot (1 + S_{22}) - S_{12} \cdot S_{21}} \quad (C.33)$$

Il faut mentionner qu'une première calibration est faite afin de ramener l'analyseur de réseau au niveau des pointes RF de mesure en corrigeant les effets des câbles, des sondes...etc. Il s'en suit une deuxième correction pour corriger les capacités parasites dues aux contacts hyperfréquence : c'est le de-embedding.

C.III.3.c) Obtention des capacités et de la charge d'inversion

La capacité grille-substrat C_{GB} doit d'abord être déterminée. Lorsque $V_D=0V$, la symétrie du circuit est conservée et la capacité grille-drain C_{GD} est égale à la capacité grille-source C_{GS} . Dans ces conditions $C_{GB}=C_{GS}-C_{GD}$ comme source et bulk sont connectés ensemble (figure C.43).

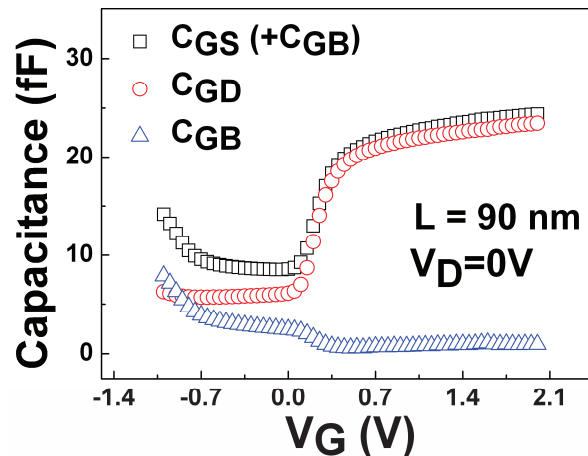


Figure C.43: Méthodologie d'extraction des capacités avec une structure deux ports

Sachant qu'au fur et à mesure que V_D augmente le circuit perd sa symétrie, la valeur de C_{GB} à $V_D=0V$ sera utilisée pour corriger C_{GS} à fort V_D . Ce pendant une incertitude existe sur la variation (ou non) de C_{GB} en fonction de V_D .

Les capacités parasites ne sont pas négligeables et doivent être corrigées à faible V_D (figure C.44a, et son inset) et à fort V_D (figure C.44b et son inset).

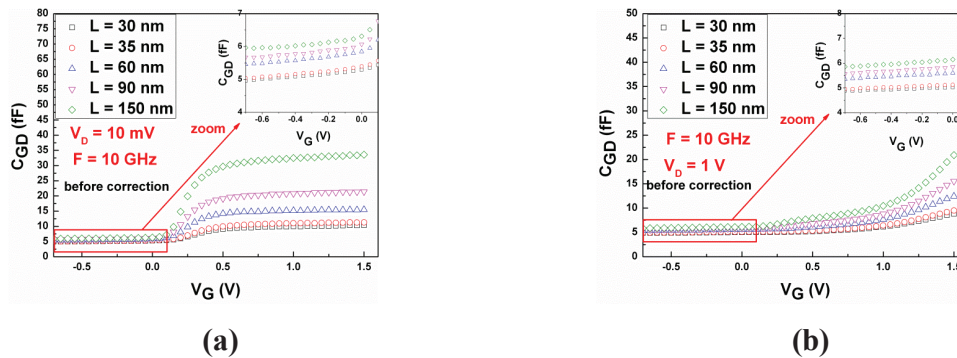


Figure C.44: Illustration des capacités parasites à V_D faible **(a)** et V_D fort **(b)** : avant correction. Un zoom des capacités parasites totales est présenté dans les insets

Les différentes contributions de capacités parasites dans un MOSFET sont la capacité d'inner fringe C_{if} , la capacité d'outer fringe C_{of} et la capacité d'overlap C_{ov} comme illustré sur la [figure C.45](#).

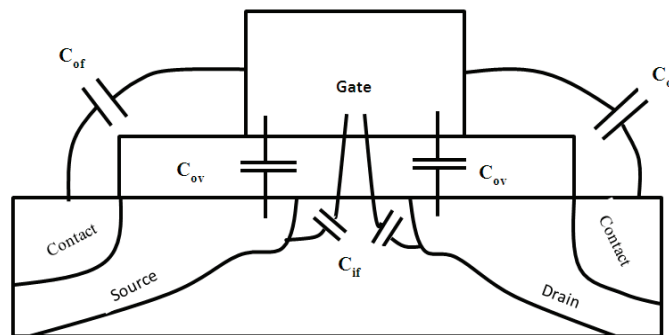


Figure C.45: Localisation des différentes contributions de capacités parasites dans un MOSFET

C_{ov} et C_{of} ne dépendent quasiment que de la longueur de grille, et non des polarisations du transistor, et elles peuvent être corrigées en ramenant à zéro la capacité totale en début de zone de déplétion ([figures C.46a & C.46b](#)). En effet dans ce régime, ces capacités sont les seules mesurables. Les inserts de ces figures montrent que la capacité parasite restante C_{if} dépend à la fois de L et de V_G . En inversion ou en accumulation, la capacité C_{if} est écrantée par les charges d'inversion ou d'accumulation. En plus de cela, sa valeur est très faible et affecte peu l'extraction ($C_{if} < 0.5\text{fF}$).

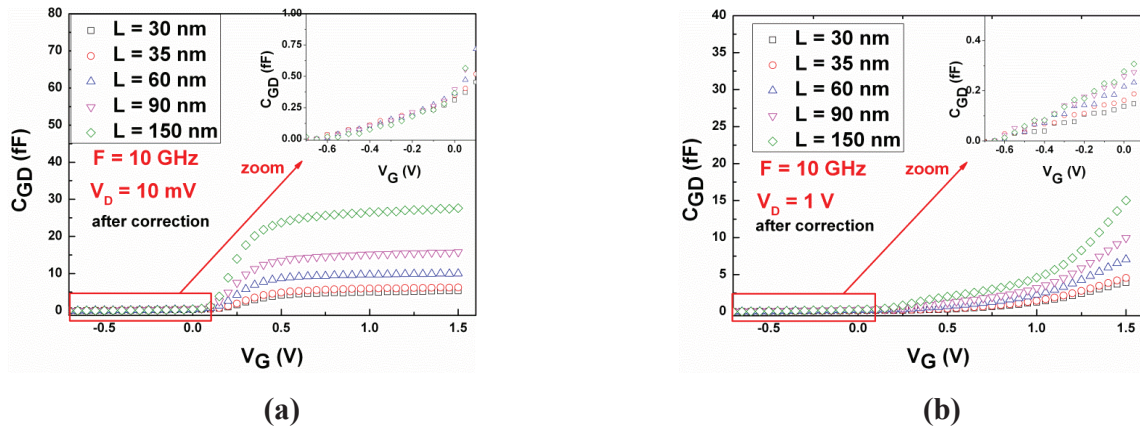


Figure C.46: Illustration des capacités parasites à V_D faible (a) et V_D fort (b) : après correction. Un zoom des capacités parasites C_{if} est présenté dans les insets

De cette manière, les figures C.47a & C.47b montrent des C_{GD} et des C_{GC} en fonction de V_G extraites d'une manière fiable pour des tensions de drain variant de 10mV à 1V.

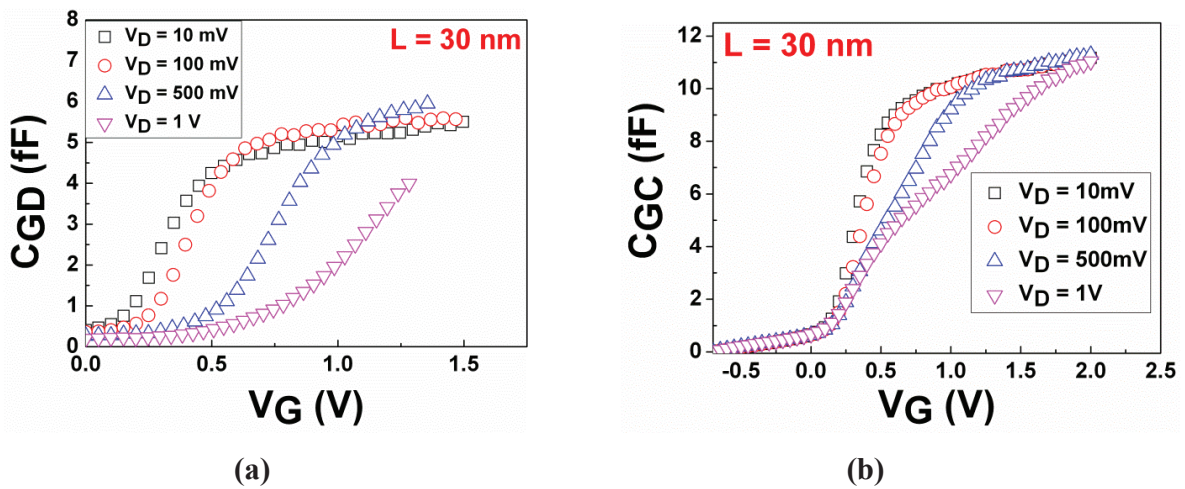


Figure C.47: C_{GD} (a) et C_{GC} (b) en fonction de V_G pour différent V_D

La charge d'inversion correspondante peut être obtenue pour différentes longueurs en fonction de la tension de grille en saturation et en linéaire comme illustré sur la figure C.48 et son insert.

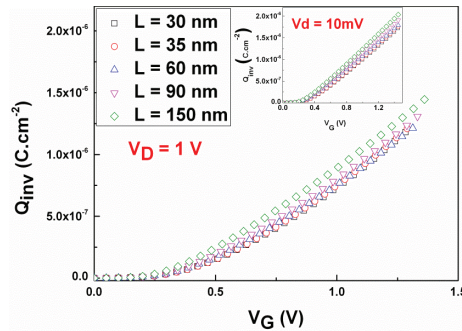


Figure C.48: Extraction de la charge d'inversion pour différentes longueurs en fonction de V_G en saturation. La charge d'inversion en linéaire est représentée dans l'inset

C.III.4) Extraction de mobilité et de vitesse

La technique du split-CV est le plus souvent appliquée sur des transistors longs. En effet comme nous l'avons évoqué à la [partie III.2](#) de ce chapitre, sur un transistor court, les mesures de capacité mises en jeu sont beaucoup plus difficiles à mettre en œuvre. Le facteur de qualité diminue lorsqu'il faut mesurer de faibles valeurs de capacité. Dans [\[Romanjek04\]](#), la méthode du split-CV améliorée est appliquée sur des dispositifs jusqu'à 50nm. Pour arriver à cela, il a fallu maximiser le temps d'intégration et effectuer des corrections périlleuses et imprécises des capacités parasites. Tout cela a fait que la mesure est longue et difficile à mettre en place. Une première application des mesures capacitives à 10GHz est donc l'application de la technique du split-CV sur des transistors aussi courts que 30nm pour extraire la mobilité effective μ_{eff} ([figure C.49](#)).

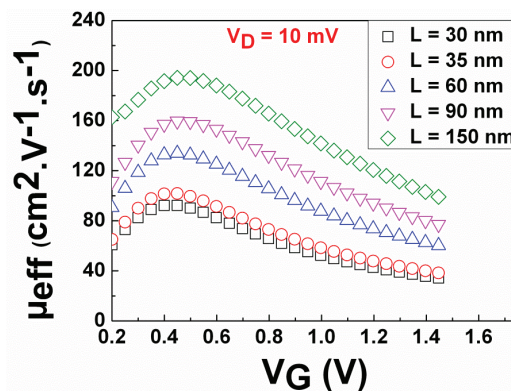


Figure C.49: Extraction de la mobilité effective par la technique du split-CV sur des transistors ultra-courts en utilisant des mesures capacitives à 10 GHz

A partir de ces valeurs de mobilité, le taux de rétrodiffusion r_{lin} (ou le taux de balisticité $1-r_{lin}$) peut être obtenu en régime linéaire en utilisant [Pappas09].

$$r_{lin} = 1 - \frac{\mu_{exp}}{\mu_{bal}} \quad (C.34)$$

$$\frac{1}{\mu_{exp}} = \frac{1}{\mu_{bal}} + \frac{1}{\mu_{dd}} \quad (C.35)$$

$$\mu_{bal} = \frac{q \cdot v_T \cdot L}{2 \cdot k \cdot T} \quad (C.36)$$

La mobilité dérive-diffusion μ_{dd} qui équivaut à la mobilité expérimentale μ_{exp} corrigée des effets balistiques, la mobilité balistique μ_{bal} , le coefficient de rétrodiffusion r_{lin} et le taux de balisticité $1-r_{lin}$ sont représentés sur les figures C.50a & C.50b (cf. chapitre II.2.b de cette partie pour plus de détails).

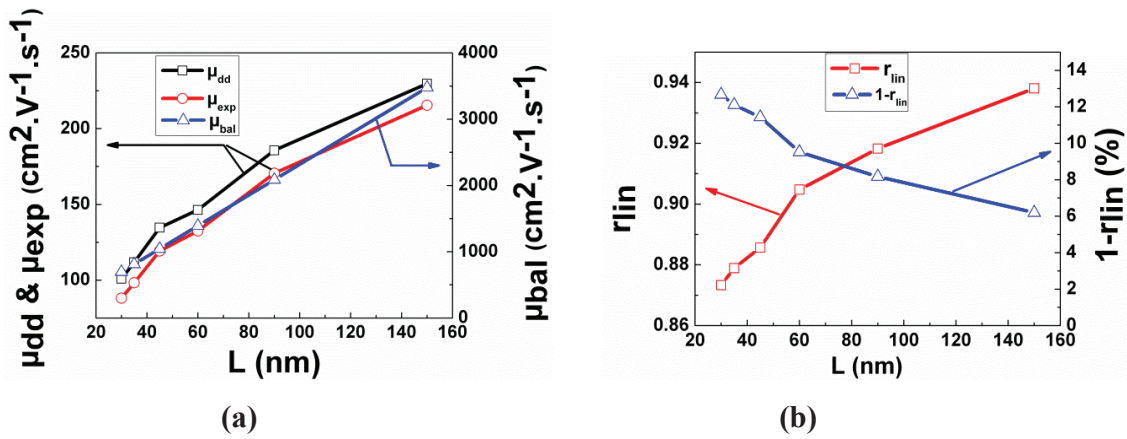


Figure C.50: μ_{dd} , μ_{exp} et μ_{bal} en fonction de la longueur (a) ; r_{lin} et $1-r_{lin}$ (%) en fonction de la longueur (b)

Un coefficient de rétrodiffusion de 0.87 en régime linéaire est à noter pour la plus courte dimension, étant en phase avec les grandes valeurs de mobilité balistique obtenues.

Une deuxième application de ces mesures capacitives à 10 GHz est l'extraction de la vitesse moyenne expérimentale en régime de saturation en utilisant la charge d'inversion mesurée Q_{inv} (figures C.51a & C.51b).

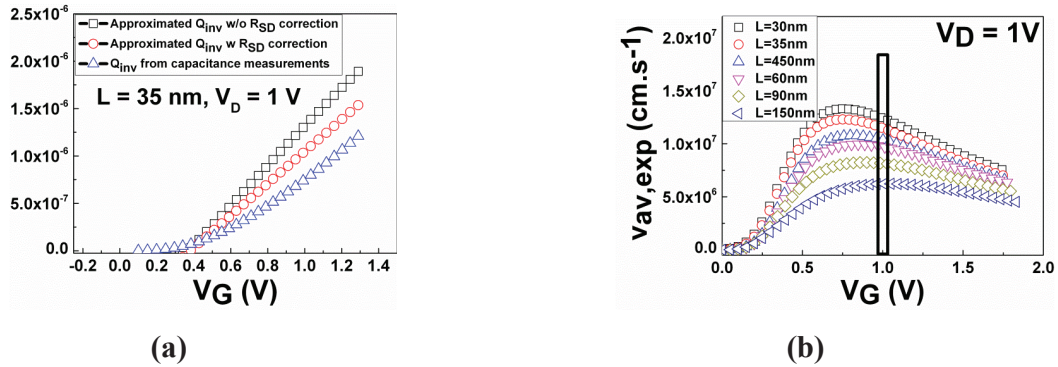


Figure C.51: Q_{inv} en fonction de V_G utilisant (C.26) sans correction R_{SD} , avec correction R_{SD} sur V_G et vraie Q_{inv} provenant des mesures capacitives (a) ; vitesse moyenne expérimentale utilisant la vraie Q_{inv} en fonction de V_G pour plusieurs longueurs (b)

La valeur de la vitesse moyenne correspondant à $V_G=V_D=1\text{V}$ sera prise pour la suite. Elle n'est en effet pas loin de la vitesse maximale comme on peut s'en apercevoir sur la figure C.51b.

En comparant la vitesse moyenne issue de la vraie charge d'inversion et celle issue de la charge d'inversion approximée, on remarque que l'approximation de la charge d'inversion sous-estime la vitesse moyenne comme illustré sur la figure C.52. Cette sous-estimation est plus importante pour les petits dispositifs.

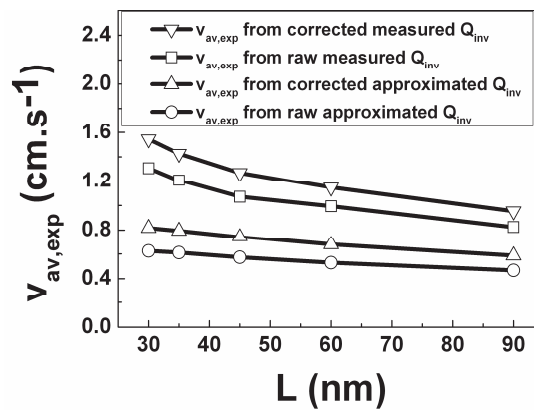


Figure C.52: Vitesse moyenne expérimentale issue de la vraie charge d'inversion avec et sans correction R_{SD} sur le courant et vitesse moyenne expérimentale issue de la charge d'inversion approximée avec et sans correction R_{SD} sur V_G

En s'inspirant de ce qui est fait plus haut en linéaire, on peut calculer le coefficient de rétrodiffusion en saturation r_{sat} comme suit :

$$r_{sat} = 1 - \frac{v_{av,exp}}{v_{av,bal}} \quad (C.37)$$

Il faudra au préalable déterminer la vitesse moyenne balistique $v_{av,bal}$ dans le cas d'un régime purement balistique, on a :

$$I_{bal} = Q_{is} \cdot v_T = Q_i(x) \cdot v(x) \quad (C.38)$$

Q_{is} est la charge à la source, v_T est la vitesse thermique à la source (égale à la vitesse d'injection), $Q_i(x)$ et $v(x)$ sont la charge et la vitesse à un point x donné dans le canal respectivement.

En utilisant la conservation de l'énergie cinétique, $v(x)$ est donné par :

$$v(x) = v_T \cdot \sqrt{1 + \frac{2 \cdot \frac{q}{m} \cdot V_D \cdot \frac{x}{L}}{v_T^2}} \quad (C.39)$$

Sur le segment $[0, L]$, la valeur moyenne de la vitesse est donnée en fonction de V_D par :

$$v_{av}(V_D) = \frac{1}{L} \int_0^L v(x) \cdot dx \quad (C.40)$$

En utilisant (C.38) et (C.40), on montre que :

$$v_{av}(V_D) = \frac{Q_{is} \cdot v_T}{\frac{1}{L} \int_0^L \frac{Q_{is} \cdot v_T}{v(x)} \cdot dx} \quad (C.41)$$

Finalement on obtient:

$$v_{av,bal}(V_D) = \frac{v_T}{\int_0^{V_D} \frac{1}{\sqrt{1 + \frac{2 \cdot \frac{q}{m} \cdot V_D \cdot u}{v_T^2}}} \cdot du} = \frac{v_T}{m \cdot v_T^2 \frac{\left(\frac{m \cdot v_T^2 + 2 \cdot q \cdot V_D}{m \cdot v_T^2} \right)^{0.5} - 1}{q \cdot V_D}} \quad (C.42)$$

Lorsque $V_D=1V$, $v_{av,bal}(1) = 6.048 \cdot v_T$ et $v_T = 1.2 \times 10^7 \text{ cm} \cdot \text{s}^{-1}$ comme dans [Lundstrom02].

Sur la figure C.53, une comparaison de la vitesse moyenne expérimentale $v_{av,exp}$ obtenue et celle théorique dans le cas d'un régime purement balistique $v_{av,bal}$ est faite.

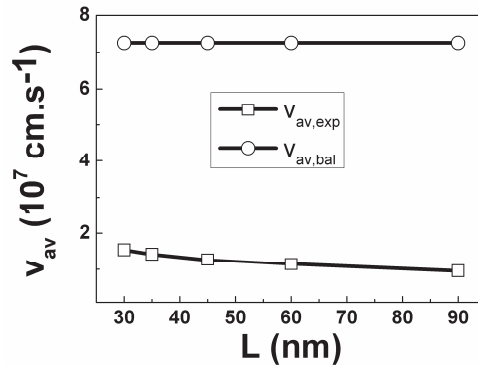


Figure C.53: Comparaison de la vitesse moyenne expérimentale $v_{av,exp}$ et la vitesse moyenne dans le cas d'un régime purement balistique $v_{av,bal}$

On peut constater que $v_{av,bal}$ est quasiment cinq fois plus grand que $v_{av,exp}$ à $V_D=V_G=1V$. Le coefficient de rétrodiffusion r_{sat} ainsi que le taux de balisticité $1-r_{sat}$ qui en découle est montré en [figure C.54](#).

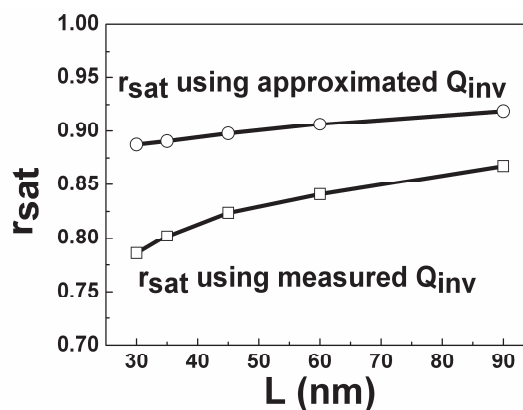


Figure C.54: r_{sat} en fonction de la longueur

Le taux de balisticité est plus important en saturation qu'en linéaire. Cependant Il n'excède pas 25% même pour les longueurs les plus petites.

C.III.5) Application à l'extraction du v_{lim}

L'extraction de la charge d'inversion intervient aussi dans le processus d'extraction de la vitesse limite [Fleury09]. Pour améliorer la procédure d'extraction de la vitesse limite, nous proposons (au [chapitre II de cette partie](#)) d'utiliser V_{THSAT} obtenu par la dérivée deuxième du courant de saturation pour l'approximation de la charge d'inversion. Avec les mesures

capacitives à 10 GHz, l'obtention directe de la charge d'inversion donnent des valeurs de vitesse limite plus grandes qu'auparavant (figure C.55) mais plus proches de la réalité. La démarche d'évaluation de la balisticité suivie dans ce chapitre reste plus concrète et plus cohérente que celle de la vitesse limite. En effet la technique de la vitesse limite fait appel à la mobilité en régime de saturation qui n'a plus vraiment de sens.

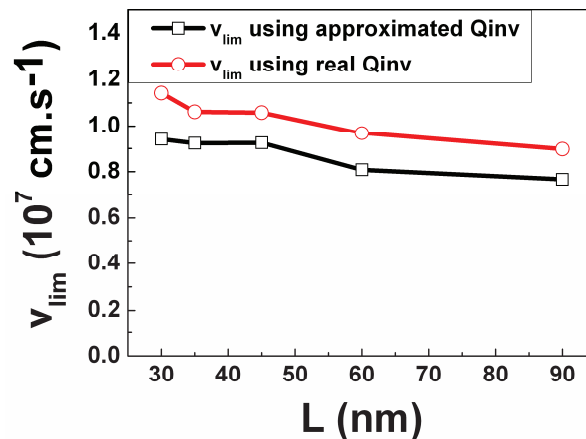


Figure C.55: calcul de v_{lim} comme dans [Fleury09] en utilisant Q_{inv} approximé et VTHSAT via la dérivé deuxième du courant et v_{lim} en utilisant la vraie charge d'inversion

C.III.6) Bilan et conclusion

Des mesures de capacité à 10 GHz ont été menées sur des transistors ultra-courts et offrent un vrai avantage pour l'investigation du transport électronique. En régime linéaire, elles donnent des valeurs fiables de mobilité effective et en régime de saturation elles permettent une extraction directe de la charge d'inversion et conséquemment, de la vitesse moyenne des porteurs pour des dispositifs jusqu'à 30 nm. Une extraction correcte du coefficient de rétrodiffusion est donc possible aussi bien en linéaire qu'en saturation. Si le coefficient de rétrodiffusion tourne autour de 0.87 en régime linéaire sur un dispositif mesurant 30 nm de long, il n'est pas supérieur à 0.78 en régime de saturation pour le même dispositif.

Ce genre de mesures pourraient être fait sur des transistors encore plus courts que 30 nm (FDSOI ou pas) vu que la fréquence peut encore être augmentée jusqu'à 110 GHz sans grosse difficulté au plan expérimental.

Table des figures de la partie C

Figure C.1: Exemple de courant de drain en fonction de la tension de grille de nMOSFETs. $V_D = 10\text{mV}$	106
Figure C.2: Figure d'illustration de la fonction Y en fonction de la tension de grille.....	108
Figure C.3: Illustration de la tension de seuil et de la mobilité en fonction de la longueur de grille extraits par la « fonction Y ».....	108
Figure C.4: Illustration du courant de drain mesuré et du courant de drain calculé avec la formule 30.....	110
Figure C.5: Méthode d'extraction de la résistance série $R_{TOT} (\frac{1}{\beta})$	111
Figure C.6: Mobilité en fonction de la tension de grille: extraction avec les paramètres extraits via la fonction Y.....	112
Figure C.7: Illustration de l'extraction de V_{TH}^* avec la méthode d'extrapolation.....	114
Figure C.8: Illustration de la fonction Y en fonction de VG en saturation.....	115
Figure C.9: Courant mesuré et calculé avec l'équation C.30 en saturation.....	115
Figure C.10: Illustration de la mobilité apparente en saturation en fonction du champ longitudinal. (a) nMOSFET & (b) pMOSFET.	117
Figure C.11: Procédure d'extraction de la vitesse de saturation en utilisant la fonction Y en saturation.....	117
Figure C.12: Vitesse de saturation extraite avec la fonction Y et la méthode standard, avec et sans correction des R_{SD} (faibles valeurs de R_{SD})	117
Figure C.13: Vitesse de saturation extraite avec la fonction Y avec de faibles valeurs de R_{SD} .	118
Figure C.14: Vitesse de saturation extraite avec la fonction Y avec de grandes valeurs de R_{SD}	119
Figure C.15: Extraction de la mobilité faible charge et de la mobilité apparente en saturation sur du $\text{Si}_{0.7}\text{Ge}_{0.3}$ et Ge	120

Table des figures de la partie C

Figure C.16: Comparaison de vitesse de saturation (extraite par la méthode de la fonction Y) entre $\text{Si}_{0.7}\text{Ge}_{0.3}$ et Ge	120
Figure C.17: Courant de drain en fonction de la tension de grille: illustration de l'amélioration entre $\text{Si}_{0.7}\text{Ge}_{0.3}$ et Ge	121
Figure C.18: Coupe TEM illustrant un transistor Ge de longueur de grille 65 nm [Mitard08]..	123
Figure C.19: Valeurs record en corrélation $I_{\text{OFF}}/I_{\text{ON}}$ de transistor Ge étudiés dans ce chapitre [Mitard08].....	124
Figure C.20: Extraction de la résistance source-drain dans les dispositifs Ge-pMOSFETs.....	124
Figure C.21: Illustration des courants de fuite dans un nMOSFET bulk: vue schématique (à gauche) et structure de bande (à droite) [Batail09].....	125
Figure C.22: Coupe d'une structure de bande d'une jonction PN à l'équilibre : pas de courant de tunnel possible (a) ; Coupe d'une structure de bande d'une jonction PN en polarisation inverse modérée : la largeur de tunnel est assez importante pour empêcher les porteurs de passer par effet tunnel (b) ; Coupe d'une structure de bande d'une jonction PN en forte polarisation inverse : Les porteurs ont de forte chance de passer d'une zone à l'autre (c) [Batail09].....	126
Figure C.23: Effet tunnel direct dans des semiconducteurs à gap direct tels que le GaAs (a) ; dans des semiconducteurs à gap indirect tels que le Si (b) ; effet tunnel direct ou indirect dans des semiconducteurs à gap indirect tels que le Ge (c) [Richard04]	127
Figure C.24: Courant de drain, courant de drain corrigé et courant de bulk en régime linéaire ($V_{\text{D}}=10\text{mV}$) (a) et en saturation ($V_{\text{D}}=1\text{V}$) (b) ($L=105\text{ nm}$)	128
Figure C.25: Mobilité split CV sur Ge comparée à la mobilité universelle sur Si	129
Figure C.26: Illustration de la mobilité dérive diffusion, expérimentale et balistique	130
Figure C.27: Coefficient de rétrodiffusion et taux de balisticité en fonction de la longueur de grille.....	131
Figure C.28: Procédure d'extraction de V_{THsat} utilisant le DIBL sous le seuil à courant constant	132
Figure C.29: Illustration de la procédure d'extraction de V_{THsat} en utilisant la dérivée seconde du courant de drain en saturation.....	133
Figure C.30: Comparaison du DIBL extrait par la méthode du courant constant et la méthode de la dérivée seconde du courant de saturation	133

Table des figures de la partie C

Figure C.31: maillage du transistor (a) ; potentiel à la surface du canal (b), potentiel dans le volume du canal (c).....	135
Figure C.32: Minimum du potentiel de surface en fonction de la tension de Grille (a) et de la tension de drain (b)	136
Figure C.33: DIBL extrait via le minimum de potentiel de surface (a) et comparaison des trois méthodes de calcul du DIBL (b).....	137
Figure C.34: Procédure d'extraction de la vitesse limitante	139
Figure C.35: Vitesse limitante en fonction de la longueur de grille pour différentes températures	139
Figure C.36: Vitesse limitante en fonction de la température pour différentes longueurs de grille	140
Figure C.37: Méthodologie d'obtention de Q_{inv} en régime linéaire	142
Figure C.38: Procédure de mesure de capacité à fort V_D sur des transistors longs	144
Figure C.39: Capacité grille-canal (a) et charge d'inversion (b) en fonction de V_G à différents V_D pour des transistors longs.....	145
Figure C.40: Mesures de C_{GC} à fort V_D à 100 MHz et 10 GHz pour $L=250$ nm (a) et $L=30$ nm (b)	145
Figure C.41: Schéma équivalent de la structure deux ports utilisée (a); coupe schématique de la structure deux ports avec source et substrat connectés à la masse (b).....	146
Figure C.42: Illustration d'un quadripôle avec des entrées-sorties en courant et tension	147
Figure C.43: Méthodologie d'extraction des capacités avec une structure deux ports.....	148
Figure C.44: Illustration des capacités parasites à V_D faible (a) et V_D fort (b) : avant correction. Un zoom des capacités parasites totales est présenté dans les insets.....	149
Figure C.45: Localisation des différentes contributions de capacités parasites dans un MOSFET	149
Figure C.46: Illustration des capacités parasites à V_D faible (a) et V_D fort (b) : après correction. Un zoom des capacités parasites C_{if} est présenté dans les insets.....	150
Figure C.47: C_{GD} (a) et C_{GC} (b) en fonction de V_G pour différent V_D	150

Table des figures de la partie C

Figure C.48: Extraction de la charge d'inversion pour différentes longueurs en fonction de V_G en saturation. La charge d'inversion en linéaire est représentée dans l'inset	151
Figure C.49: Extraction de la mobilité effective par la technique du split-CV sur des transistors ultra-courts en utilisant des mesures capacitives à 10 GHz.....	151
Figure C.50: μ_{dd} , μ_{exp} et μ_{bal} en fonction de la longueur (a) ; r_{lin} et $1-r_{lin}$ (%) en fonction de la longueur (b).....	152
Figure C.51: Q_{inv} en fonction de V_G utilisant (C.26) sans correction R_{SD} , avec correction R_{SD} sur V_G et vraie Q_{inv} provenant des mesures capacitives (a) ; vitesse moyenne expérimentale utilisant la vraie Q_{inv} en fonction de V_G pour plusieurs longueurs (b)	153
Figure C.52: Vitesse moyenne expérimentale issue de la vraie charge d'inversion avec et sans correction R_{SD} sur le courant et vitesse moyenne expérimentale issue de la charge d'inversion approximée avec et sans correction R_{SD} sur V_G	153
Figure C.53: Comparaison de la vitesse moyenne expérimentale $v_{av,exp}$ et la vitesse moyenne dans le cas d'un régime purement balistique $v_{av,bal}$	155
Figure C.54: r_{sat} en fonction de la longueur	155
Figure C.55: calcul de v_{lim} comme dans [Fleury09] en utilisant Q_{inv} approximé et V_{THSAT} via la dérivé deuxième du courant et v_{lim} en utilisant la vraie charge d'inversion	156

Liste des tableaux de la partie C

Tableau C.1: Illustration de la faible fluctuation de la résistance série avec la tension de grille dans nos dispositifs 113

Tableau C.2: Processus de fabrication des dispositifs Ge étudiés dans ce chapitre [Mitard08]. 123

Conclusion générale et perspectives

Les résultats obtenus durant cette thèse ont été sub-divisés en deux parties de trois chapitres chacune.

La première partie décrit l'influence du procédé de fabrication sur les mécanismes de transports électronique dans des dispositifs déca-nanométriques. Quant à la deuxième partie, elle est consacrée à l'introduction et à l'amélioration de techniques d'extractions de paramètres dans les transistors MOSFETs ultra-courts.

Première partie :

Dans cette partie, un premier chapitre a permis de montrer que le recuit à haute pression permet d'avoir de meilleures performances électriques avec une réduction de la densité d'états d'interface. Le résultat varie selon que l'on considère les NMOS ou les PMOS, les transistors longs ou courts. Même si le recuit sous H₂ donne de meilleures performances électriques, l'étude en fiabilité a permis de montrer que le recuit sous deutérium permet un meilleur compromis entre performances électriques et durée de vie du composant.

Nous avons pu voir que le PMOS semblait présenter davantage de remote Coulomb scattering que le NMOS. Ceci n'est pas très étonnant au vu des différences qui existent entre l'empilement de grille du NMOS et du PMOS. La couche de SiGe épitaxiée dans le cas du PMOS vient s'ajouter à cette différence. Le deuxième chapitre a été consacré à l'étude de l'impact de cette couche de SiGe sur les performances du transistor.

L'origine d'un décalage supplémentaire inattendu de 150 mV dans des dispositifs SiGe a été étudiée dans ce deuxième chapitre de la première partie. Des simulations Poisson-Schrödinger ont montré la nécessité d'un ajustement de 150 mV du travail de sortie effectif du métal du transistor Si_{0.7}Ge_{0.3} pour reproduire la capacité totale mesurée expérimentalement. Les mesures de mobilité (faibles charges et effectives) montrent une plus grande amélioration comparée à la littérature pour les dispositifs SiGe. Ce phénomène n'a pu être expliqué ni par le dopage ni par les états d'interface. Les mesures à basses températures ont permis de révéler un

taux important de remote Coulomb scattering dans les dispositifs Si dû à une plus grande quantité de charge dans son empilement de grille. Les mesures de bruit ont aussi révélé une plus grande contribution de fluctuations de mobilité corrélées et une plus grande densité de piège dans le pMOS/Si.

Toutes ses conclusions laissent penser que le PMOS étudié dans ce chapitre présente un high- κ avec des charges et/ou des pièges qui agissent négativement sur la mobilité. On suspecte de plus une interaction physico-chimique entre le Ge et l'empilement de grille, lors de sa formation. Cette interaction tend à améliorer la mobilité en plus de l'amélioration due à la plus faible masse effective du Ge.

En utilisant des oxydes interfaciaux d'épaisseurs différentes, on modifie la position du high- κ dans le stack de grille. Dans ce cas, l'impact du high- κ sur les porteurs du canal peut être étudié. Dans le chapitre suivant, des transistors GO1 (oxyde standard) et GO2 (oxyde plus épais) ont été comparés afin de discriminer le rôle du high- κ sur les performances du PMOS 28 nm bulk.

Un gain en mobilité est constaté sur les transistors de type GO2 avec un oxyde de grille plus épais. Cela veut dire que le high- κ , qui dans un tel cas est repoussé plus loin du canal, a un impact négatif sur la mobilité. L'analyse de la mobilité additionnelle a permis de montrer que cet impact sur le transport est majoritairement de nature coulombienne. Il en découle avec certitude que même si le high- κ présente un vrai avantage en terme de miniaturisation, son impact sur le transport électronique est non négligeable.

Deuxième partie :

Le premier chapitre de cette deuxième partie montre à quel point il est utile de maîtriser l'extraction de paramètres pour comprendre le fonctionnement du transistor. La fonction Y est une technique qui permet d'extraire des paramètres en s'affranchissant de l'influence des résistances source-drain. Nous l'avons étendue en régime de saturation pour obtenir les paramètres en régime de fonctionnement du transistor.

Une mobilité apparente ainsi que la vitesse de saturation ont été calculées. Il en a découlé grâce à l'application faite sur du SiGe et du Ge, que la mobilité en linéaire ne donnait pas une vraie

image du courant à l'état ON du transistor lorsqu'il s'agit de prévoir des améliorations de performances.

La mauvaise corrélation entre la mobilité linéaire et le courant I_{ON} a ainsi été surmontée grâce à la mobilité apparente et à la vitesse de saturation extraite d'une manière plus précise.

Ces paramètres que sont la mobilité linéaire, la mobilité apparente et la vitesse de saturation sont très intéressants lorsqu'il faut faire un diagnostic rapide du transport. Pour aller plus loin et étudier la nature du transport de façon plus précise, une extraction plus sophistiquée a été nécessaire, notamment à la recherche de la balisticité dans les transistors ultra-courts. C'est pourquoi, nous avons utilisé dans le chapitre suivant la méthode dite de v_{lim} pour caractériser le transport dans des transistors à base de germanium où une meilleure mobilité est attendue.

Le chapitre II de la deuxième partie a porté sur le transport dans des transistors germanium avec un procédé de fabrication optimisé grâce à la méthode de la « vitesse limitante ». Des améliorations telles que l'extraction de la tension de seuil ont été apportées à cette méthode. Des simulations TCAD ont permis de valider la manière d'extraire V_{THsat} par la dérivée seconde du courant.

Malgré les bonnes propriétés de transport du germanium, nous avons montré que la limite quasi-balistique n'était pas encore atteinte sur des transistors aussi courts que 63 nm et pour une gamme de température allant de 415 K à 243 K. Le transport reste donc gouverné par les collisions sur les phonons et sur les impuretés. Il faut cependant noter que d'après la méthode de v_{lim} , la limite pourrait être atteinte pour des températures inférieures à 200K ou pour des transistors beaucoup plus courts que 60 nm dans le germanium.

Dans cette étude, on a utilisé le V_{THsat} pour approximer la charge d'inversion Q_s en saturation. En régime linéaire, la charge d'inversion est obtenue en intégrant la capacité. En régime de saturation, les mesures de capacité ne sont pas possibles en DC ou à basse fréquence. Au chapitre suivant, des mesures hautes fréquences (HF) ont été utilisées pour mesurer directement la charge d'inversion en saturation, sans passer par V_{THsat} .

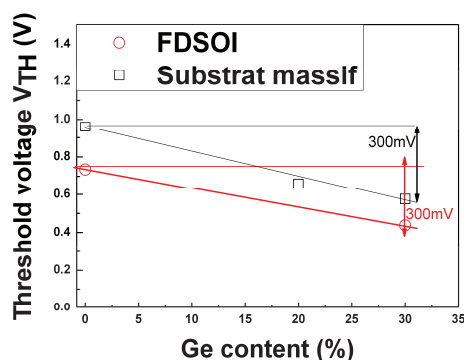
Dans ce dernier chapitre, des mesures de capacité à 10 GHz ont été effectuées sur des transistors ultra-courts et offrent un vrai avantage pour l'investigation du transport électronique. L'introduction de deux « bias-T » entre le capacimètre, le drain et la source n'a pas suffi pour mesurer des capacités à fort V_D sur des transistors ultra-courts.

En régime linéaire, les mesures à 10 GHz donnent des valeurs fiables de mobilité effective et en régime de saturation, elles permettent une extraction directe de la charge d'inversion et conséquemment, de la vitesse moyenne des porteurs pour des dispositifs aussi courts que 30 nm. Une extraction correcte du coefficient de rétrodiffusion a donc été rendue possible aussi bien en linéaire qu'en saturation. Si le coefficient de rétrodiffusion est autour de 0.87 en linéaire sur un dispositif mesurant 30 nm de long, il atteint 0.78 en régime de saturation.

Ce type de mesures pourrait donc être réalisé sur des transistors encore plus courts que 30 nm (FDSOI ou sur substrat massif) car la fréquence peut encore être augmentée jusqu'à 110 GHz sans grosse difficulté au plan expérimental, autre que la disponibilité de l'équipement.

Perspectives

Sur substrat massif, on s'est aperçu que le germanium permet d'augmenter les performances en réduisant l'impact négatif de l'empilement de grille sur le transport. Une étude similaire est nécessaire pour comprendre l'effet qu'aurait le germanium sur le PMOS du FDSOI. On peut d'ores et déjà remarquer que le décalage de V_{TH} obtenu en FDSOI est du même ordre de grandeur que celui obtenu en substrat massif (cf. figure ci-dessous).



Décalage de V_{TH} en FDSOI comparé à celui substrat massif après introduction du Ge

Conclusion générale et perspectives

Pour le FDSOI, les techniques d'analyse du transport devront être adaptées. Par exemple, les mesures HF faites à 10 GHz en substrat massif pourront être étendues jusqu'à 110 GHz afin de faciliter l'extraction des capacités.

Egalement, ce travail a permis de mettre en lumière l'intérêt de l'utilisation de matériaux de canal alternatif. Pour dépasser les limitations rencontrées pour les PMOS à canal Ge, l'élaboration de transistors à canal de type III-V est à l'étude. L'analyse des nouveaux phénomènes physiques mis en jeu devra être menée.

Références bibliographiques

- [Akiyama07]** Akiyama. K, W Wang, W Mizubayashi, M Ikeda, H Ota, T Nabatame, and A Toriumi. “Roles of oxygen vacancy in HfO₂ / ultra-thin SiO₂ gate stacks.” In International Symposium of VLSI-TSA, 172906:72-73, 2007.
- [Alatise09]** Alatise Olayiwola M, Sarah H Olsen, Nick E B Cowern, Anthony G O Neill, and Prashant Majhi. “Performance Enhancements in Scaled Strained-SiGe pMOSFETs With HfSiO_x / TiSiN Gate Stacks.” TED 56, no. 10 (2009): 2277-2284.
- [Andrieu06]** Andrieu. F, O Faynot, X Garros, D Lafond, L Tosti, S Minoret, V Vidal, et al. “Comparative Scalability of PVD and CVD TiN on HfO₂ as a Metal Gate Stack for FDSOI cMOSFETs down to 25nm Gate Length and Width.” In IEDM, 2006.
- [Ang04]** Ang. King Jien Chui, Vladimir Blimetsov, Anyan JJu, N. Balasubramanian, Ming Fu Li, Ganesh Samudra, and Yee-Chia Ye0. “Enhanced Performance in 50 nm N-MOSFETs with Silicon-Carbon Source/Drain Regions.” In IEDM, 1069-1071, 2004.
- [Arnaud09]** Arnaud. F et al. “Competitive and Cost Effective high-k based 28nm CMOS Technology for Low Power Applications.” In IEDM proceedings, 651-654, 2009.
- [Auth12]** Auth. C, C Allen, A Blattner, D Bergstrom, M Brazier, M Bost, M Buehler, et al. “A 22nm High Performance and Low-Power CMOS Technology Featuring Fully-Depleted Tri-Gate Transistors, Self-Aligned Contacts and High Density MIM Capacitors IDSAT (A / μm) VGS (V).” In VLSIT, m: 131-132, 2012.
- [Ayhan00]** Ayhan. A. M and Mahmud Rahman. “Two-Dimensional Analytical Model for Drain Induced Barrier Lowering (DIBL) in Short Channel MOSFETs.” IEEE proceedings (2000): 340-344.
- [Baccarini85]** Baccarini et al. “An Investigation of Steady-State Velocity Overshoot in Silicon.” SSE 28, no. 4 (1985): 407.
- [Baudot11]** Baudot, S, C Leroux, F Chave, R Boujamaa, E Martinez, P Caubet, M Silly, and F Sirotti. “Microelectronic Engineering Understanding reversal effects of metallic aluminum introduced in HfSiON / TiN PMOSFETs.” Microelectronic Engineering 88, no. 7 (2011): 1305-1308.
- [Bennamane13]** Bennamane, K, I Ben Akkez, A Cros, F Balestra, and G Ghibaudo. “Mobility coupling effects due to remote Coulomb scattering in thin-film FD-SOI CMOS devices.” Electronics Letters 49, no. 7 (2013): 21-22.

Références bibliographiques

- [Beranger07]** Beranger et al. “Fully-Depleted SOI Technology using High-K and Single-Metal Gate for 32nm Node LSTP Applications featuring 0.179tm2 6T-SRAM bitcell.” In IEDM, 6-9, 2007.
- [Beranger12]** Beranger et al. “Enhancement of Devices Performance of hybrid FDSOI / Bulk Technology by using UTBOX sSOI substrates.” In VLSI T, 4-5, 2012.
- [Betail09]** Batail Estelle. Etude et intégration de dispositifs MOS a canal de conduction haute mobilité obtenus en technologie Silicon-On-Nothing/ germanium-On-Nothing technology, PhD thesis, 2009.
- [Boutchacha99]** Boutchacha, G. Ghibaudo’ and B. Belmekki. “STUDY OF LOW FREQUENCY NOISE IN THE 0.18 μ m SILICON CMOS TRANSISTORS.” In ICMTS, 12:84-88, 1999.
- [Brugler69]** Brugler et al. “Charge Pumping in MOS Devices.” TED 16, no. 06 (1969).
- [Cassé06]** Cassé Mikaël, Laurent Thevenod, Bernard Guillaumot, Lucie Tosti, François Martin, Jérôme Mitard, Olivier Weber, et al. “Carrier Transport in HfO₂ / Metal Gate MOSFETs: Physical Insight Into Critical Parameters.” TED 53, no. 4 (2006): 759-768.
- [Chang10]** Chang, Yang-hua, and Yao-jen Liu. “A New Extraction Method for Source / Drain Resistance in MOSFETs.” In ICSICT, 4-6, 2010.
- [Charbonnier10a]** Charbonnier. M, C. Leroux, G. Reimbold, F. Allain, A. Toffoli, G. Ghibaudo, F. Martin, H. Grampeix, F. Boulanger. “Investigation of mechanisms shifting metal effective workfunction towards P+ for various Al incorporation scenarii.” In International Symposium of VLSI-TSA, 84–85, 2010.
- [Charbonnier10b]** Charbonnier. M, C. Leroux, V. Cosnier, P. Besson, E. Martinez, N. Benedetto, C. Licitra, N. Rochat, C. Gaumer, K. Kaja, G. Ghibaudo, F. Martin, G. Reimbold. “Measurement of Dipoles/Roll-Off /Work Functions by Coupling CV and IPE and Study of Their Dependence on Fabrication Process.” IEEE TED 58 (2010): 1809–1819.
- [Chiarella10]** Chiarella, T, L Witters, A Mercha, C Kerner, M Rakowski, C Ortolland, L Ragnarsson, et al. “Benchmarking SOI and bulk FinFET alternatives for PLANAR CMOS scaling succession.” Solid State Electronics 54, no. 9 (2010): 855-860.
- [Cho06]** Cho, Sung-man, Jeong-hyn Lee, M- Chang, M Jo, H Hwang, J Lee, S Hwang, and Jong-ho Lee. “High Pressure Deuterium Annealing Effect on Nano- Scale CMOS Devices with Different Channel Width.” In NMDC, 6:98-99, 2006.
- [Choi02]** Choi, Rino, Katsunori Onishi, Chang Seok Kang, Sundar Gopalan, Renee Nieh, Y H Kim, Jeong H Han, et al. “Fabrication of High Quality Ultra-thin HfO₂ Gate Dielectric MOSFETs Using Deuterium Anneal.” In IEDM proceedings, 613-616, 2002.

Références bibliographiques

- [Chui04]** Chui and Saraswat. “Advanced Germanium MOSFET Technologies with High-K Gate Dielectrics and Shallow Junctions.” In ICICDT, 245-251, 2004.
- [Cros06a]** Cros Antoine. Caractérisation électrique des transistors MOS à grille enrobante pour les technologies CMOS sub-45nm. PhD thesis, 2006.
- [Cros06b]** Cros, Antoine, Krunoslav Romanjek, Dominique Fleury, Samuel Harrison, Robin Cerutti et al. “Unexpected mobility degradation for very short devices: A new challenge for CMOS scaling.” In IEDM proceedings, 15-18, 2006.
- [Deen92]** Deen. J and Z. X. Yan. “DIBL in Short-Channel NMOS Devices.” TED 39, no. 4 (1992).
- [Diouf12a]** Diouf et al. “On the Understanding of the Effects of High Pressure Deuterium and Hydrogen Final Anneal.” In ULIS, 10-13, 2012.
- [Diouf12b]** Diouf, C, A Cros, S Monfray, J Mitard, J Rosa, D Gloria, G Ghibaudo, Jean Monnet, and Minatec Inpg. “AN EXTENDED ‘Y FUNCTION’ METHOD FOR SATURATION REGIME CHARACTERIZATION: APPLICATION TO BULK Si AND Ge TECHNOLOGIES.” In SSDM, 3-4, 2012.
- [Diouf11]** Diouf, C, A Cros, S Monfray, J Mitard, J Rosa, F Boeuf, G Ghibaudo, “TRANSPORT CHARACTERISATION OF Ge pMOSFETS IN SATURATION REGIME.” In ESSDERC, 223-226, 2011.
- [Diouf13a]** Diouf, C, A Cros, S Monfray, J Mitard, J Rosa, D Gloria, and G Ghibaudo. “Y function” method applied to saturation regime: Apparent saturation mobility and saturation velocity extraction.” SOLID STATE ELECTRONICS 85, no. 12 (2013): 12-14.
- [Diouf13b]** Diouf, C, Cros, A, Rideau, D et al. “Unexpected impact of germanium content in SiGe bulk PMOSFETs.” Solid-State Electronics 86 (2013): 45-50.
- [Diouf13c]** Diouf, C, A Cros, D Gloria, J Rosa, M Buczko, and G Ghibaudo. “High Field Transport Characterization in Nano MOSFETs using 10GHz capacitance measurements.” IEDM proceedings, 2013
- [Eminente07]** Eminente, Simone, Nicola Barin, Pierpaolo Palestri, Claudio Fiegna, Enrico Sangiorgi, and Abstract A Monte. “Small-Signal Analysis of Decanometer Bulk and SOI MOSFETs for Analog / Mixed-Signal and RF Applications Using the Time-Dependent Monte Carlo Approach.” TED 54, no. 9 (2007): 2283-2292.
- [Erginsoy50]** Erginsoy Cavid. “Neutral Impurity scattering in semiconductors.pdf.” Letter to Editor (1950).

Références bibliographiques

- [Fikri94]** Fikri et al. "Temperature dependence of drain-induced barrier lowering in deep submicrometre MOSFETs." *Electronics Letters* 30, no. 11 (1994).
- [Fischetti01]** Fischetti, Massimo V, Deborah A Neumayer, and Eduard A Cartier. "Effective electron mobility in Si inversion layers in metal-oxide-semiconductor systems with a high- κ insulator: The role of remote phonon scattering." *Journal of Applied Physics* 90, no. 4587 (2001).
- [Fleury09a]** Fleury, Dominique. contribution à l'étude expérimentale du transport dans les transistors de dimension deca-nanométriques des technologies sub-45 nm, M, PhD thesis. *Traitement du Signal*, 2009.
- [Fleury09b]** Fleury. D, G Bidal, A Cros, F Boeuf, T Skotnicki, and G Ghibaudo. "New Experimental Insight into Ballisticity of Transport in Strained Bulk MOSFETs." In *VLSIT*, 16-17, 2009.
- [Fleury09c]** Fleury Dominique, Student Member, Antoine Cros, and Julien Rosa. "A New Technique to Extract the Source/Drain Series Resistance of MOSFETs." *EDL*, no. 2 (2009): 1-3.
- [Fleury08a]** Fleury, Dominique, Antoine Cros, Hugues Brut, Jean Monnet, F-Crolles, and Parvis Louis Neel. "New Y -Function-Based Methodology for Accurate Extraction of Electrical Parameters on Nano-Scaled MOSFETs." In *ICMTS*, 160-165, 2008.
- [Fleury08b]** Fleury, Dominique, Student Member, Antoine Cros, Krunoslav Romanjek, David Roy, Franck Perrier, Benjamin Dumont, Hugues Brut and Gérard Ghibaudo "Automatic Extraction Methodology for Accurate Measurements of Effective Channel Length on 65-nm MOSFET Technology and Below." *Technology* 21, no. 4 (2008): 504-512.
- [Ge01]** Ge Lixin, Jerry G Fossum, and Bin Liu. "Physical Compact Modeling and Analysis of Velocity Overshoot in Extremely Scaled CMOS Devices and Circuits." *TED* 48, no. 9 (2001): 2074-2080.
- [Ghani03]** Ghani et al. "A 90nm High Volume Manufacturing Logic Technology Featuring Novel 45nm Gate Length Strained Silicon CMOS Transistors." In *IEDM*, M:978-980, 2003.
- [Ghibaudo09]** Ghibaudo. G, M Mouis, K Bennamane, I Pappas, A Cros, G Bidal, A Claverie, et al. "Electrical Transport characterization of nano CMOS devices with ultra-thin silicon film." In *Invited paper, IWJT*, 2009.
- [Ghibaudo88]** Ghibaudo Gérard. "NEW METHOD FOR THE EXTRACTION OF MOSFET PARAMETERS." *Electronics Letters* 2, no. 9 (1988): 543-545.
- [Ghibaudo97]** Ghibaudo. G. "Microelectronic engineering." *Microelectronic Engineering* 39 (1997): 31-57.
- [Glimour86]** Gilmour. A. S, *Microwave Tubes*. Edited by Artech House Publishers, 1986.

Références bibliographiques

- [Giusi10]** Giusi, Gino, Giuseppe Iannaccone, Senior Member, Debabrata Maji, and Felice Crupi. "Barrier Lowering and Backscattering Extraction in Short-Channel MOSFETs." *TED* 57, no. 9 (2010): 2132-2137.
- [Grove67]** Grove-John Wiley & Sons. Grove.pdf, 1967.
- [Haartman06]** Haartman, Martin Von, Student Member, B Gunnar Malm, Mikael Östling, and Al O Hfo Al O. "Comprehensive Study on Low-Frequency Noise and Mobility in Si and SiGe pMOSFETs With High- κ Gate Dielectrics and TiN Gate" 53, no. 4 (2006): 836-843.
- [Hamer86]** Hamer. "First order parameter extraction on enhancement silicon MOS transistors." *IEE proceedings* 133 I, no. 2 (1986).
- [Hellings09]** Hellings, Geert, Jerome Mitard, Geert Eneman, Brice De Jaeger, David P Brunco, Denis Shamiryan, Tom Vandeweyer, Marc Meuris, Marc M Heyns, and Kristin De Meyer. "High Performance 70-nm Germanium pMOSFETs With Boron LDD Implants." *EDL* 30, no. 1 (2009): 88-90.
- [Hobbs01]** Hobbs, C, H Tseng, K Reid, B Taylor, L Dip, L Hebert, R Garcia, et al. "80 nm Poly-Si Gate." In *IEDM*, 4-7, 2001.
- [Hu87]** Hu et al. "Gate-Voltage-Dependent Effective Channel Length and Series Resistance of LDD MOSFET's." *TED* 34, no. 12 (1987).
- [ITRS03]** "International Technology Roadmap for Semiconductors (ITRS), 2003 [Online]. Available: <http://public.itrs.net/>." Online (2003): 2003-2003.
- [Maserjian94]** J. Maserjian et al. "Saturation capacitance of thin oxide MOS Structures and the effective surface density of states of silicon." *Solid-State Electronics* 17, no. 335 (1994).
- [Jacoboni77]** Jacoboni et al. "C. Jacoboni, C. Canali, G. Ottaviani et A. Alberigi Quaranta. A review of some charge transport properties of silicon." *SSE* 20, no. 2 (1977): 1977-1977.
- [Jaeger07]** Jaeger, G. Nicholas¹, D. P. Brunco, G. Eneman¹, M. Meuris, and M. M. Heyns. "High Performance High-k/Metal Gate Ge pMOSFETs with Gate Lengths down to 125 nm and Halo Implant." In *ESSDERC*, 465-465, 2007.
- [Kesan91]** Kesan et al. "HIGH PERFORMANCE 0.25 μ m p-MOSFETs WITH SILICON-GERMANIUM CHANNELS FOR 300K AND 77K OPERATIONS." In *IEDM*, 25-28, 1991.
- [Kita09]** Kita, Koji, and Akira Toriumi. "Origin of electric dipoles formed at high-k / SiO₂ interface." *Applied Physics Letters* 94, no. 132902 (2009): 1-4.
- [Koomen73]** Koomen. "Investigation of the MOST channel conductance in weak inversion." *Solid State Electron* 16 (1973): 801-810.

Références bibliographiques

- [Krishnan11]** Krishnan, Mestay. K. “A Manufacturable Dual Channel (Si and SiGe) High-K Metal Gate CMOS Technology with Multiple Oxides for High Performance and Low Power Applications.” In IEDM proceedings, 4-6, 2011.
- [Lee83]** Lee. K and Shur.M. S. “Impedance of thin semiconductor films.” JJAP 54, no. 7 (1983): 4034-4034.
- [Liu09]** Liu et al. “22nm CMOS Approaches by PVD TiN or Ti-Silicide as Metal Gate.” In VLSI-TSA, 63-64, 2009.
- [Lochtefield01]** Lochtefeld, Anthony, Student Member, and Dimitri A Antoniadis. “On Experimental Determination of Carrier Velocity in Deeply Scaled NMOS: How Close to the Thermal Limit?” TED 22, no. 2 (2001): 95-97.
- [Lochtefield02]** Lochtefeld, I. J. Djomehri, G. Samudra and D. A. Antoniadis. “New insight into carrier transport in n-MOSFETs.” IBM Journal of research and development 46, no. 2 (2002): 347-357.
- [Lundstrom97]** Lundstrom, Mark. “Elementary Scattering Theory of the Si MOSFET.” EDL 18, no. 7 (1997): 361-363.
- [Lundstrom01]** Lundstrom, Mark S. “On the Mobility Versus Drain Current Relation for a Nanoscale MOSFET.” EDL 22, no. 6 (2001): 293-295.
- [Lundstrom02]** Lundstrom, Mark, and Zhibin Ren. “Essential Physics of Carrier Transport in Nanoscale MOSFETs.” TED 49, no. 1 (2002): 133-141.
- [Maiti98]** Maiti et al. “PVD TiN Metal Gate MOSFETs on Bulk Silicon and Fully Depleted Silicon-On- Insulator (FDSOI) Substrates for Deep Sub-Quarter Micron CMOS Technology.” In IEDM, 781-784, 1998.
- [Masson99]** “Masson.” IEE EL (1999).
- [McGill75]** McGill. T. C, “neutral impurity scattering in semiconductors.” Physical review B 11, no. 12 (1975): 5208-5210.
- [McLarty95]** McLarty et al. “A simple parameter extraction method for ultra – thin oxide MOSFETs.” Solid State Electronics 38, no. 6 (1995): 1175-1177.
- [Minjoo05]** Minjoo et al. “Strained Si, SiGe, and Ge channels for high-mobility metal-oxide-semiconductor field-effect transistors.” Journal of Applied Physics 97, no. 011101 (2005): 1-27.
- [Misra05]** Misra. V, “Issues in metal gate electrode selection for bulk CMOS devices in High Dielectric Constant Materials—VLSI MOSFET Applications.” H. R. Huff and D. C. Gilmer, Eds. Berlin, Germany: Springer-Verlag, no. ch. 14 (2005): 415–434.
- [Mitard08]** Mitard, J, B De Jaeger, F E Leys, G Hellings, K Martens, G Eneman, R Loo, et al. “Record I ON / I OFF performance for 65nm Ge pMOSFET and novel Si passivation scheme for improved EOT scalability.” In IEDM, 11-14, 2008.
- [Moore65]** Moore. “Cramming More Components onto Integrated Circuits.” Electronics (1965): 114-117.

Références bibliographiques

- [nanohub.org]** nanohub. “__nanohub.org”
- [Natori01]** Natori. “Scaling Limit of the MOS Transistor – A Ballistic MOSFET (invited paper). In IEICE Trans. Electron., volume E84-C” (2001): pages 1029–1036.
- [Natori94]** Natori. K. “Current-voltage characteristics of silicon on insulator metal oxide semiconductor field effect transistors in ballistic mode.” Japanese Journal of Applied Physics 33 (1994): 1994-1994.
- [Negre11]** Negre, Laurent. Caractérisation et modélisation de la fiabilité des transistors mos en radio fréquence, 2011.
- [Otal07]** Otal, Hiroyuki, Akito Hirano, Yukimune Watanabe, Naoki Yasuda, Kunihiro Iwamoto, Koji Akiyama, Kenji Okada, Shinji Migita, Toshihide Nabatame, and Akira Toriumi. “Intrinsic Origin of Electron Mobility Reduction in High-k MOSFETs - From Remote Phonon to Bottom Interface Dipole Scattering -.” In IEDM, 6-9, 2007.
- [Palestri05]** Palestri Pierpaolo, David Esseni, Simone Eminent, Claudio Fiegna, Enrico Sangiorgi, Luca Selmi, Abstract In, Part II, and Monte Carlo Mc. “Understanding Quasi-Ballistic Transport in Nano-MOSFETs : Part I — Scattering in the Channel and in the Drain.” TED 52, no. 12 (2005): 2727-2735.
- [Pantisano08]** Pantisano, L, L Trojman, J Mitard, B Dejaeger, S Severi, G Eneman, G Crupi, et al. “in sub-100nm (110) -Si and (100) -Ge.” In Symposium on VLSI Technology, 52-53, 2008.
- [Pappas08]** Pappas, I, G Ghibaudo, C A Dimitriadis, and C Fenouillet-béranger. “Backscattering coefficient and drift-diffusion mobility extraction in short channel MOS devices.” Solid State Electronics 53, no. 1 (2009): 54-56.
- [Park09]** Park, Min Sang, Kyong Taek Lee, Seung Ho Hong, Seung Hyun Song, Gil Bok Choi, Rock Hyun Baek, Hyun Sik Choi, et al. “High Pressure Hydrogen Annealing Effect of CESL Nitride Stressor MOSFETs with Metal Gate / High-k Dielectric on the Performance and Reliability.” In IEE NMDC, 229-232, 2009.
- [Rahhal13]** Rahhal, Lama, Aurélie Bajolet, Antoine Cros, Cheikh Diouf, Flore Kergomard, Julien Rosa, Gregory Bidal, Raul-andres Bianchi, and Gérard Ghibaudo. “Solid-State Electronics Impact of Ge proportion on advanced SiGe bulk P-MOSFET matching performances.” Solid State Electronics 85 (2013): 15-22.
- [Richard04]** Richard. Modélisation physique de la structure électronique, du transport et de l’ionisation par choc dans les matériaux IV-IV massifs, contraintes et dans les puits quantiques, PHD thesis, orsay, 2004.

Références bibliographiques

- [Robertson07]** Robertson. J, Ka Xiong, and Koon-yiu Tse. “Importance of Oxygen Vacancies in Gate Dielectrics.” *Journal of Vacuum Science and Technology A* (2007): 5-8.
- [Romanjek04]** Romanjek, K, F Andrieu, T Ernst, G Ghibaudo, and Senior Member. “Improved Split C – V Method for Effective Mobility Extraction in sub-0.1- μ m Si MOSFETs.” *Test* 25, no. 8 (2004): 583-585.
- [Sabnis79]** Sabnis, Anant G, and James T Clemens. “CHARACTERIZATION OF THE ELECTRON MOBILITY IN THE INVERTED <1 0 0> SI SURFACE.” In *IEDM*, 18-21, 1979.
- [Sah72]** Sah. C. T, T. H. Ning and L. L. Tschopp. “The scattering of electrons by surface oxide charges and by lattice vibrations at the silicon-silicon dioxide interface.” *Surface science* 32 (1972): 561-575.
- [Sahrani00]** Sahrani, S. A. Rahman. “High Temperature Annealing Effects on the C-V and G-V Characteristics of the n-Channel Enhancement Mode Power MOSFET Device.” In *ISCE proceedings*, 76-79, 2000.
- [Saito03]** Saito, S, D Hisamoto, S Kimura, and M Hiratani. “Unified Mobility Model for High-Ic Gate Stacks.” In *IEDM proceedings*, 797-800, 2003.
- [Sarah04]** Sarah. H et al. “Study of Single- and Dual-Channel Designs for High-Performance Strained-Si–SiGe n-MOSFETs.” *TED* 51, no. 8 (2004): 1245-1253.
- [Shur02]** Shur, Michael S. “Low Ballistic Mobility in Submicron HEMTs.” *EDL* 23, no. 9 (2002): 511-513.
- [Skotnicki03]** Skotnicki. T et Boeuf. F. T. Skotnicki et F. Boeuf, chapitre *Physique des dispositifs pour circuits intégrés sur silicium*. EGEM, Encyclopédie Hermès, 2003.
- [Skotnicki08]** Skotnicki. T, C. Fenouillet-Beranger, C. Gallon, F. Boeuf, S. Monfray, F. Payet, A. Pouydebasque, M. Szczap, A. Farcy, F. Arnaud, S. Clerc, M. Sellier, A. Cathignol, J-P. Schoellkopf, E. Perea, Richard Ferrant, and H. Mingam. “Innovative Materials, Devices, and CMOS Technologies for Low-Power Mobile Multimedia.” *TED* 55, no. 1 (2008): 96-130.
- [Sodini82]** Sodini. C, T. Ekstedt, and J. Moll. “Charge accumulation and mobility in thin dielectric MOS transistors.” *Solid State Electron* 25 (1982): 833–841.
- [Song09]** Song, Seung Hyun, Min Sang Park, Kyong Taek Lee, Hyun Sik Choi, Gil Bok Choi, Rock Hyun Baek, Hyun Chul Sagong, et al. “Effects of High Pressure Hydrogen Anneal Process on Performance and Reliability in HfO₂ / SiO₂ Dielectric with Contact Etch Stop Layer Stressor.” In *NMDC proceedings*, 225-228, 2009.
- [Soussou10]** SOUSSOU, D. RIDEAU, C. LEROUX, G. GHIBAUDO, C. TAVERNIER, H. Jaouen. “Modeling study of the SiGe / Si heterostructure in FDSOI pMOSFETs.” In *IEEE ISCDG IEEE ISCDG*, 219-222, 2010.

- [Subramanian10]** Subramanian. N, Ghibaudo. G and Mouis. L. "Parameter Extraction of Nano-Scale MOSFETs Using Modified Y Function Method." In ESSDERC, 309-312, 2010.
- [Sze81]** Sze. S. M et al. physics of semiconductor devices, 1981.
- [Takagi94]** Takagi, Shin-ichi, Akira Toriumi, Masao Iwase, and Hiroyuki Tango. "On the Universality of Inversion Layer Mobility in Si MOSFET's : Part I-Effects of Substrate Impurity Concentration." TED 41, no. 12 (1994): 2357-2362.
- [Takagi07]** Takagi, Shinichi, Toshifumi Irisawa, Tsutomu Tezuka, Toshinori Numata, Shu Nakaharai, Norio Hirashita, Yoshihiko Moriyama, Koji Usuda, Eiji Toyoda, and Sanjeeva Dissanayake. "Carrier-Transport-Enhanced Channel CMOS for Improved Power Consumption and Performance." TED 55, no. 1 (2007): 21-39.
- [Taur97]** Taur, Yuan, Senior Member, Douglas A Buchanan, W E I Chen, David J Frank, Khalid E Ismail, Shih-hsien Lo, et al. "CMOS Scaling into the Nanometer Regime." IEEE proceedings 85, no. 4 (1997).
- [Thompson04]** Thompson et al. "A 90-nm Logic Technology Featuring Strained-Silicon." TED 51, no. 11 (2004): 1790-1797.
- [Toniutti08]** Toniutti, Paolo, Pierpaolo Palestri, David Esseni, and Luca Selmi. "Revised analysis of the mobility and I ON degradation in high- κ gate stacks: surface optical phonons vs. Remote Coulomb scattering." In ESSDERC 2008, 246-249, 2008.
- [Tsuchiya10]** Tsuchiya, Yoshinori, Nathaniel Berliner, Ryosuke Iijima, Frederic Monsieur, Leo Tai, Nicolas Loubet, Lisa F Edge, Mariko Takayanagi, and Vamsi Paruchuri. "Physical Origin of pFET Threshold Voltage Modulation by Ge Channel Ion Implantation (GC-I / I)." In IEDM proceedings, 11.4.1-11.4.4, 2010.
- [Wang10]** Wang, Mu-chun, Hsin-chia Yang, Kuang-hung Lin, and Shuang-yuan Chen. "CESL Deposition Promoting nip MOSFETs under 45-nm-node Process Fabrication." In NMDC, 17-20, 2010.
- [Weber04]** Weber et al. "Experimental Determination of Mobility Scattering Mechanisms in Si/HfO₂/TiN and SiGe:C/HfO₂/TiN Surface Channel n- and p-MOSFETs." In IEDM proceedings, 867-870, 2004.
- [Weber05]** Weber. O, " Etude, Fabrication et Propriétés de Transport de Transistors CMOS associant un Diélectrique Haute Permittivité et un Canal de Conduction Haute Mobilité". PhD thesis, 2005. Transport, 2005.
- [Wesley92]** Welsey et al. "NMOS and PMOS Transistors Fabricated in Strained Silicon/Relaxed Silicon-Germanium Structures." TED 20 (1992): 1000-1002.

Références bibliographiques

- [Xiaoyan01]** Xiaoyan, Liu, Shuzuo Lou, Zhiliang Xia, Dechao Guo, Huiwen Zhu, Jinfeng Kang, and Ruqi H. “Characteristics of Different Structure Sub-100nm MOSFETs With High-k Gate Dielectrics.” *SSICT* 1 (2001): 333-336.
- [Xu11]** Xu, Yong, Takeo Minari, Kazuhito Tsukagoshi, Romain Gwoziecki, Romain Coppard, Francis Balestra, and Gerard Ghibaudo. “Power transfer-length method for full biasing contact resistance evaluation of organic field-effect transistors.” *Organic Electronics* 12, no. 12 (2011): 2019-2024.
- [Yu05]** Yu et M. Cardona. P. Yu et M. Cardona. Springer, 2005
- [Zhu04]** Zhu, Wenjuan, Jin-ping Han, and T P Ma. “Mobility Measurement and Degradation Mechanisms of MOSFETs Made With Ultrathin High-k Dielectrics.” *TED* 51, no. 1 (2004): 98-105.
- [Zilli07]** Zilli, M, P Palestri, D Esseni, L Selmi, Udine Iunet, and Via Scienze. “On the experimental determination of channel back-scattering in nanoMOSFETs.” In *IEDM proceedings*, 105-108, 2007.

Publications de l'auteur

Conférences internationales

--C. Diouf, A. Cros, S. Monfray, J. Mitard, J. Rosa, F. Boeuf and G. Ghibaudo, "TRANSPORT CHARACTERISATION OF Ge pMOSFETS IN SATURATION REGIME", in **ESSDERC proceedings**, 2011, Helsinki

--C. Diouf, A. Cros, P. Morin, S. Renard, X. Federspiel, M. Rafik, A. Bianchi, J. Rosa and G. Ghibaudo, "On the Understanding of the Effects of High Pressure Deuterium and Hydrogen Final Anneal", in **ULIS proceedings**, 2012, Grenoble

--C. Diouf, A. Cros, S. Monfray, J. Mitard, J. Rosa, D. Gloria and G. Ghibaudo, "AN EXTENDED "Y FUNCTION" METHOD FOR SATURATION REGIME CHARACTERIZATION: APPLICATION TO BULK Si AND Ge TECHNOLOGIES", in **SSDM proceedings**, 2012, Kyoto

--C. Diouf, A. Cros, D. Gloria, J. Rosa, M. Buczko and G. Ghibaudo, "High Field Transport Characterization in Nano MOSFETs using 10GHz capacitance measurements", in **IEDM proceedings**, 2013, Washington DC

Journaux scientifiques

--C. Diouf, A. Cros, S. Monfray, J. Mitard, J. Rosa, D. Gloria and G. Ghibaudo, "'Y FUNCTION" METHOD APPLIED TO SATURATION REGIME: APPARENT SATURATION MOBILITY AND SATURATION VELOCITY EXTRACTION", in **Solid Sates Electronics**, vol. 85, issue 12, 2013

--C. Diouf, A. Cros, A. Soussou, D. Rideau, S. Haendler, J. Rosa and G. Ghibaudo, "Unexpected impact of germanium content in SiGe bulk PMOSFETs", in **Solid Sates Electronics**, vol. 86, 2013

Conférences nationales

-C. Diouf, "Extraction de paramètres de transport électronique pour la caractérisation des transistors MOSFETs d'architecture innovante", in **JNRDM proceedings**, 2013, Grenoble

Publications en co-auteur

--P. Toniutti, R. Clerc, P. Palestri, C. Diouf, A. Cros, D. Esseni, F. Boeuf, G. Ghibaudo and L. Selmi, "Improved Procedure to Extract the Limiting Carrier Velocity in Ultra Scaled CMOS devices", in **ICMTS proceedings**, 2012, USA

--I. Ben-Akkez, C. Diouf, A. Cros, C. Fenouillet-Beranger, P. Perreau, F. Balestra, G. Ghibaudo and F. Boeuf, "On the understanding of mobility degradation mechanisms in advanced CMOS devices: FDSOI versus bulk technology", in **SSDM proceedings**, 2012, Kyoto

--L. Rahhal, A. Bajolet, C. Diouf, A. Cros, J. Rosa, N. Planes and G. GHIBAUDO, "New methodology for drain current local variability characterization using Y function method", in **ICMTS proceedings**, 2013, Tokyo.

Résumé:

La taille du transistor MOS ne cesse de diminuer pour des questions de performance et de rentabilité de fabrication. Les procédés de fabrication évoluent, l'architecture se complexifie et les méthodologies d'extraction de paramètres électriques doivent être adaptées.

C'est ainsi que dans un premier temps, les effets d'un recuit haute pression sous atmosphère hydrogène (HPH2) ou deutérium (HPD2) sur le transistor MOS sont étudiés en détail dans cette thèse. La comparaison des performances apportées en termes de transport électronique et de dégradations engendrées en fiabilité a permis de montrer que le HPD2 présente un meilleur compromis. Une étude des effets d'un canal silicium-germanium (SiGe) sur les performances des transistors MOS est ensuite réalisée. L'incorporation du SiGe a permis d'atténuer l'impact négatif sur les performances des charges contenues dans l'empilement de grille. La présence de ces charges est d'ailleurs confirmée par l'analyse des effets néfastes d'un oxyde de grille à haute permittivité, utilisé entre autre pour faciliter la miniaturisation.

Dans un deuxième temps, la « fonction Y » a été étendue en régime de saturation afin d'extraire la vitesse de saturation qui est un indicateur de performance obtenu dans les mêmes conditions que le courant I_{ON} . En outre, la problématique liée à l'extraction de la charge d'inversion sur des transistors courts et à forte tension de drain V_D a été résolue grâce à des mesures à hautes fréquences réalisées sur une structure deux ports. Ceci a rendu possible l'obtention de la mobilité effective, de la vitesse moyenne et de la vitesse limitante sur des transistors déca-nanométriques.

Mots clés : Caractérisation électrique, MOS avancés, recuit haute pression, canal silicium-germanium, oxyde à haute permittivité, charge d'inversion, mobilité, vitesse

Abstract:

The MOS gate length is continuously downscaling because of the need of higher performance and cost-effectiveness. In addition to the fabrication process, the device architecture is being more and more complex and parameters extraction need to be adapted.

First in this thesis, the effects of high pressure final anneal with hydrogen (HPH2) or deuterium (HPD2) on MOSFET properties is investigated. The transport performances and reliability degradation comparison allow to consider HPD2 as a good compromise. The effect of a silicon-germanium (SiGe) channel is also studied. It is demonstrated that SiGe channel decreases defects located in the high- κ gate stack. The presence of these defects is confirmed by the study of the negative effects of a high- κ as a gate oxide.

Secondly, the "Y function" method is extended to the saturation regime to reliably extract saturation velocity, obtained in the same conditions as I_{ON} current. The problematic due to inversion charge estimation in short devices is solved using high frequency measurements with a two ports structure. Then, effective mobility, average velocity and limiting velocity are obtained in ultra-scaled devices.

Key words: electrical characterization, advanced MOS, high pressure final anneal, silicon-germanium channel, high- κ , inversion charge, mobility, velocity