

50376  
1994  
291

26 102 556  
50376  
1994  
291

N° d'ordre: 1406

# THESE

présentée à

**l'Université des Sciences et Technologies de Lille**

pour obtenir le grade de

**Docteur de l'Université**

**Spécialité: Electronique**

par

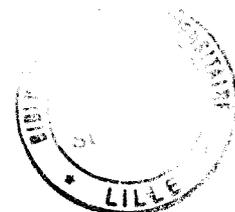
**Paul-Henri BRICOUT**

**Ingénieur ISEN**

## **CONTRIBUTION A L'ETUDE DE DISPOSITIFS NMOS SUBMICRONIQUES PAR LES METHODES DE MONTE CARLO ET DE DERIVE-DIFFUSION**

**Soutenue le 29 novembre 1994, devant la commission d'examen:**

**Président : M. G. SALMER**  
**Rapporteurs : M. T. SKOTNICKI**  
**M. J. C. VAISSIERE**  
**Examineurs : M. J. N. DECARPIGNY**  
**M. E. DUBOIS**  
**M. R. FAUQUEMBERGUE**  
**M. J. SLOTBOOM**



*à Hélène,*

*à notre fils Rémi,*

*à toute ma famille*

Cette thèse a été préparée à  
l'Institut Supérieur d'Electronique du Nord  
au sein de l'Institut d'Electronique et de Microélectronique du Nord,  
Unité Mixte de Recherche CNRS 9929

*J'exprime ma reconnaissance à Monsieur le Professeur G. SALMER qui a accepté la présidence de ce jury.*

*Je remercie vivement Monsieur T. SKOTNICKI et Monsieur le Professeur J.-C. VAISSIERE pour l'intérêt qu'ils ont porté à ce travail en acceptant la tâche de rapporteur.*

*Que Monsieur R. FAUQUEMBERGUE soit ici remercié pour le soutien apporté lors de la phase initiale de cette thèse et pour sa participation au jury.*

*Je suis reconnaissant à Monsieur J. SLOTBOOM de l'honneur qu'il me fait en ayant accepté de juger ce travail.*

*J'exprime toute ma gratitude à Monsieur J.-N. DECARPIGNY pour l'intérêt qu'il a témoigné à cette thèse et pour le soutien dont j'ai bénéficié à l'Institut Supérieur d'Electronique du Nord.*

*Je tiens à remercier également Messieurs M. LANNOO et D. COLLARD pour l'accueil qui m'a été réservé au sein du laboratoire et de l'équipe de microélectronique. Ma formation doit beaucoup à leur enthousiasme communicatif.*

*Je suis particulièrement reconnaissant à Monsieur E. DUBOIS, instigateur de ces travaux de recherche, pour les discussions fructueuses que nous avons eues au cours de la préparation de cette thèse. Ses conseils avisés m'ont toujours été précieux. J'espère que nous aurons encore longtemps l'occasion de travailler ensemble.*

## SOMMAIRE

	Page
<b>CHAPITRE I : INTRODUCTION</b>	1
<b>REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE I</b>	7
<b>CHAPITRE II :</b>	
<b>MODELISATION DU TRANSPORT ELECTRONIQUE</b>	11
II-1 INTRODUCTION	11
II-2 EQUATION DE TRANSPORT DE BOLTZMANN	11
II-3 RESOLUTIONS EXACTES DE L'EQUATION DE BOLTZMANN	15
II-3.1 Procédures itératives	15
II-3.2 Développement en série	16
II-3.3 Principe de résolution de l'équation de transport de Boltzmann par la méthode de Monte Carlo	16
II-4 METHODES DERIVEES DE L'EQUATION DE BOLTZMANN	17
II-4.1 Moments successifs	17
II-4.2 Formalisme hydrodynamique	19
II-4.3 Formalisme des équations de dérive-diffusion	20
II-5 CHOIX D'UNE STRATEGIE DE RESOLUTION	22
II-5.1 Monte Carlo	23
II-5.2 Méthodes dérivées, hydrodynamique ou dérive-diffusion	24
II-5.3 Couplage	24
II-6 MODELES UTILISES POUR LES SIMULATIONS DE DERIVE-DIFFUSION	26
II-6.1 Génération-recombinaison	26
II-6.2 Mobilité et diffusivité	27

<b>II-7 MODELES UTILISES POUR LES SIMULATIONS MONTE CARLO</b>	<b>29</b>
II-7.1 Bande d'énergie	29
II-7.2 Modélisation des interactions subies par les électrons	32
II-7.2.1 Collisions intervallées	32
II-7.2.2 Collisions intravallées	33
II-7.2.3 Collisions sur impuretés ionisées	36
II-7.2.4 Interactions entre porteurs	41
<b>REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE II</b>	<b>43</b>
<b>CHAPITRE III :</b>	
<b>MISE EN OEUVRE DES METHODES DE SIMULATION</b>	<b>49</b>
III-1 CHAINE DE SIMULATION	49
III-2 DERIVE-DIFFUSION	50
III-3 SIMULATION MONTE CARLO	52
III-3.1 Vol libre	53
III-3.2 Calcul des probabilités	54
III-3.3 Traitement des interactions: choix du vecteur d'onde après collision	56
III-3.4 Choix de la valeur de $\Delta t$	57
III-3.5 Phénomène de survitesse	58
III-4 SIMULATION MONTE CARLO DE DISPOSITIF	60
III-4.1 Localisation des particules	61
III-4.2 Interpolation des grandeurs macroscopiques	61
III-4.3 Evaluation de grandeurs macroscopiques	63
III-4.4 Conditions aux frontières	64
III-4.5 Gestion des particules: technique de multiplication	70

III-5 COUPLAGE	74
III-5.1 Evaluation des termes du tenseur thermique	75
III-5.2 Evaluation de la mobilité	78
REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE III	87

## CHAPITRE IV :

<b>ETUDE DE TRANSISTORS NMOS FORTEMENT SUBMICRONIQUES</b>	<b>93</b>
IV-1 INTRODUCTION	93
IV-2 STRATEGIE DE MINIATURISATION	94
IV-3 PRINCIPALES ETAPES TECHNOLOGIQUES DE FABRICATION	97
IV-4 PHENOMENE DE SURVITESSE	99
IV-4.1 Dépendance par rapport à la longueur	99
IV-4.2 Influence de la polarisation	104
IV-5 COUPLAGE MONTE CARLO/DERIVE-DIFFUSION: APPLICATION A L'ETUDE DE TRANSISTORS NMOS FORTEMENT SUBMICRONIQUES	106
IV-5.1 Extraction des résultats Monte Carlo	106
IV-5.2 Interprétation dans le cadre du formalisme de dérive-diffusion étendu	109
IV-5.2.1 Vitesse des porteurs	109
IV-5.2.2 Tenseur de tension thermique	111
IV-5.2.3 Mobilité et diffusivité	115
IV-5.3 Influence du modèle d'interactions sur impuretés ionisées	120
IV-5.3.1 Impact sur les paramètres de transport	120
IV-5.3.2 Dépendance des caractéristiques électriques	124
IV-5.4 Convergence du couplage Monte Carlo/dérive-diffusion	126
IV-5.5 Caractéristiques électriques obtenues par la méthode de couplage	133
IV-5.5.1 Réseaux de caractéristiques $I_D(V_{DS})$	134
IV-5.5.2 Transconductance	136

IV-5.5.3 Influence des zones résistives	137
IV-5.6 Conclusion	139
REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE IV	141
<b>CHAPITRE V :</b>	
<b>TRANSISTORS SUBMICRONIQUES NON PLANS</b>	<b>145</b>
V-1 TRANSISTORS NON PLANS	146
V-2 ETUDE A FAIBLE POLARISATION	147
V-2.1 Effet de coin	147
V-2.2 Réduction de la tension de seuil et pouvoir de coupure	148
V-2.3 Influence de la forme du canal	150
V-2.4 Influence de la concentration du substrat	152
V-3 ETUDE A FORTE POLARISATION	156
V-3.1 Simulations classiques de dérive-diffusion	156
V-3.2 Effets balistiques: étude par la méthode de Monte Carlo	158
V-3.3 Influence de la longueur du canal	163
V-3.4 Influence de la polarisation	166
V-3.4.1 Influence de la tension drain-source $V_{DS}$	166
V-3.4.2 Influence de la tension grille-source $V_{GS}$	170
V-3.5 Performances en courant des dispositifs non plans	174
V-4 CONCLUSION	176
REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE V	177
<b>CHAPITRE VI : CONCLUSIONS</b>	<b>179</b>
REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE VI	183

## Liste des symboles utilisés

Les grandeurs vectorielles sont représentées en gras.

Les tenseurs sont surmontés du signe  $\sim$

$b$  : distance moyenne entre impuretés ionisées

$C_D$  : capacité de déplétion d'un transistor MOS

$C_{ox}$  : capacité d'oxyde d'un transistor MOS

$C_n$  : taux de génération recombinaison

$C_P$  : taux de changement du moment

$C_W$  : taux de changement d'énergie cinétique

$d^3k$  : volume élémentaire de l'espace des vecteurs d'onde

$d^3r$  : volume élémentaire de l'espace réel

$D$  : diffusivité des électrons

$\tilde{D}$  : tenseur de diffusivité

$D$  : vecteur déplacement électrique ( $D = \epsilon E$ )

$d_{max}$  : distance maximale parcourue par une particule lors d'un pas de temps (Monte Carlo)

$E$  : vecteur champ électrique

$E_x$  : composante du champ électrique selon x

$E_{\perp}$  : composante du champ électrique orthogonale au vecteur densité de courant

$E_{\parallel}$  : composante du champ électrique parallèle au vecteur densité de courant

$\mathcal{E}$  : énergie d'un électron

$\mathcal{E}_{ac}$  : constante du modèle d'interactions intravallées

$\mathcal{E}_a$  : énergie après absorption d'un phonon

$\mathcal{E}_e$  : énergie après émission d'un phonon

$\mathcal{E}_i$  : niveau d'énergie des pièges (génération-recombinaison de Shockley-Read-Hall)

$\mathcal{E}_p$  : énergie utilisée dans les modèles de collisions sur impuretés ionisées

$f(\mathbf{r}, \mathbf{k}, t)$  : fonction de distribution des électrons

$\left[ \frac{\partial f}{\partial t} \right]_{coll}$  : taux de variation de  $f$  dû aux interactions

**F** : résultante des forces exercées sur les électrons  
**F<sub>D</sub>** : forces correspondant aux différentes interactions  
**F<sub>E</sub>** : force due au champ électrique  
*g<sub>m</sub>* : transconductance  
**G-R** : génération recombinaison  
 $\hbar$  : constante de Planck réduite ( $\hbar/2\pi = 1.05458 \cdot 10^{-34}$  Js)  
 $\hbar\mathbf{k}$  : moment  
**I<sub>D</sub>** : courant de drain  
**J<sub>n</sub>** : vecteur densité de courant d'électrons  
**J<sub>p</sub>** : vecteur densité de courant de trous  
**k** : vecteur d'onde  
**k'** : vecteur d'onde après interaction  
**k\*** : vecteur d'onde par la transformation de Herring-Vogt  
**k<sub>f</sub>** : vecteur d'onde en fin de vol libre  
**k<sub>i</sub>** : vecteur d'onde en début de vol libre  
**k<sub>l</sub>** : composante longitudinale du vecteur d'onde (vallée ellipsoïdale)  
**k<sub>t</sub>** : composante transversale du vecteur d'onde  
**k<sub>0</sub>** : position du centre des vallées en bande de conduction ( $k_0 = 9.836 \cdot 10^7 \text{ cm}^{-1}$ )  
*k<sub>B</sub>* : constante de Boltzmann ( $1.38066 \cdot 10^{-23}$  J/K)  
**L<sub>eff</sub>** : longueur effective d'un transistor MOS  
*m<sub>0</sub>* : masse de l'électron ( $9.1095 \cdot 10^{-31}$  kg)  
*m<sub>d</sub>* : masse effective de densité d'état  
*m<sub>l</sub>* : masse effective selon l'axe principal de la vallée  
*m<sub>t</sub>* : masse effective selon une direction transversale de la vallée  
*n* : concentration en électrons  
*n<sub>i</sub>* : concentration intrinsèque en électrons  
*n<sub>v</sub>* : nombre de vallées de destination possibles pour les interactions intervalles  
**N<sub>A</sub>** : concentration en accepteurs d'électrons  
**N<sub>D</sub>** : concentration en donneurs d'électrons

$N_I$  : concentration en impuretés ionisées  
 $N_{Sub}$  : concentration du substrat  
 $p$  : concentration en trous  
 $P_i(\Delta t)$  : probabilité qu'une interaction de type  $i$  surviennent pendant  $\Delta t$   
 $P$  : moment moyen des électrons  
 $q$  : charge de l'électron ( $1.60218 \cdot 10^{-19}$  C)  
 $Q$  : densité du flux de chaleur  
 $r$  : nombre aléatoire compris entre 0 et 1  
 $\mathbf{r}$  : vecteur position  
 $\mathbf{r}_f$  : vecteur position en fin de vol libre  
 $\mathbf{r}_i$  : vecteur position en début de vol libre  
 $S(\mathbf{k}, \mathbf{k}')$  : probabilité pour un électron de passer de  $\mathbf{k}$  en  $\mathbf{k}'$   
 $S$  : pouvoir de coupure d'un transistor (swing)  
 $S$  : densité moyenne du flux d'énergie  
 $t$  : temps  
 $\tilde{\mathbf{T}}$  : tenseur de température  
 $T_L$  : température du réseau cristallin  
 $T_{Ph}$  : température de phonon  
 $t_{ent}$  : profondeur de gravure pour les dispositifs non plans  
 $t_{ox}$  : épaisseur d'oxyde  
 $\tilde{\mathbf{U}}$  : tenseur de tension thermique  
 $U_{ij}$  : terme du tenseur de tension thermique  
 $\mathbf{v}$  : vitesse d'un électron  
 $\mathbf{v}_f$  : vitesse en fin de vol libre  
 $\mathbf{v}_i$  : vitesse en début de vol libre  
 $\mathbf{v}_l$  : composante de vitesse selon l'axe principal de la vallée  
 $\mathbf{v}_t$  : composante de vitesse selon un axe transversal de la vallée  
 $\mathbf{v}_x$  : composante selon  $x$  de la vitesse d'un électron  
 $v_s$  : vitesse du son dans le silicium

$v_{Sat}$  : vitesse de saturation  
 $V$  : vitesse moyenne des électrons  
 $V_x$  : composante selon x de la vitesse moyenne  
 $V(\mathbf{r})$  : potentiel créé par une impureté ionisée  
 $V_{DS}$  : tension drain source  
 $V_{GS}$  : tension grille source  
 $V_T$  : tension thermique scalaire ( $V_T = \frac{k_B T}{q}$ )  
 $W$  : énergie moyenne des électrons  
 $W_0$  : énergie moyenne des électrons à l'équilibre  
 $W$  : largeur de canal  
 $x_j$  : profondeur de jonction  
 $x$  : direction du canal  
 $y$  : direction orthogonale au canal (profondeur)  
 $Z$  : nombre d'ionisations des impuretés

$\alpha$  : facteur de non parabolicité des vallées en bande de conduction  
 $\alpha$  : pente de l'interface (chapitre V)  
 $\beta_s$  : facteur d'écrantage du potentiel créé par les impuretés ionisées  
 $\gamma(\mathcal{E}) = \mathcal{E}(1 + \alpha\mathcal{E})$   
 $\Delta t$  : pas de temps utilisé lors des simulations Monte Carlo  
 $\epsilon$  : permittivité diélectrique  
 $\epsilon_0$  : permittivité diélectrique du vide ( $8.85418 \cdot 10^{-12}$  F/m)  
 $\epsilon_{Ox}$  : permittivité diélectrique de l'oxyde de silicium  $\text{SiO}_2$  ( $3.9 \epsilon_0$ )  
 $\epsilon_{Si}$  : permittivité diélectrique du silicium ( $11.7 \epsilon_0$ )  
 $\theta$  : angle de déviation du vecteur d'onde lors d'une interaction  
 $\theta_m$  : angle minimal de déviation  
 $\lambda(\mathbf{k})$  : nombre moyen d'interactions par seconde  
 $\lambda_a$  : taux d'interactions (absorptions de phonons)  
 $\lambda_e$  : taux d'interactions (émissions de phonons)

$\lambda_{ac}$  : taux de collisions intravallées

$\lambda_i$  : taux d'interactions de type i

$\lambda_{BH}$  : taux de collisions sur impuretés ionisées (modèle de Brooks-Herring)

$\lambda_{CW}$  : taux de collisions sur impuretés ionisées (modèle de Conwell-Weisskopf)

$\lambda_m(\mathbf{k})$  : taux de perte du moment

$\mu$  : mobilité des électrons

$\tilde{\mu}$  : tenseur de mobilité

$\mu_{ij}$  : terme du tenseur de mobilité

$\rho$  : densité du semiconducteur

$\tau_i$  : temps moyen entre deux interactions de type i

$\tau_n$  : durée de vie des électrons (génération-recombinaison de Shockley-Read-Hall)

$\tau_p$  : durée de vie des trous (génération-recombinaison de Shockley-Read-Hall)

$\tau_{\mathbf{p}}$  : temps de relaxation du moment

$\tau_W$  : temps de relaxation d'énergie

$\varphi$  : angle utilisé en coordonnées sphériques

$\Psi$  : potentiel

$\nabla_{\mathbf{k}}$  : gradient par rapport aux coordonnées du vecteur d'onde

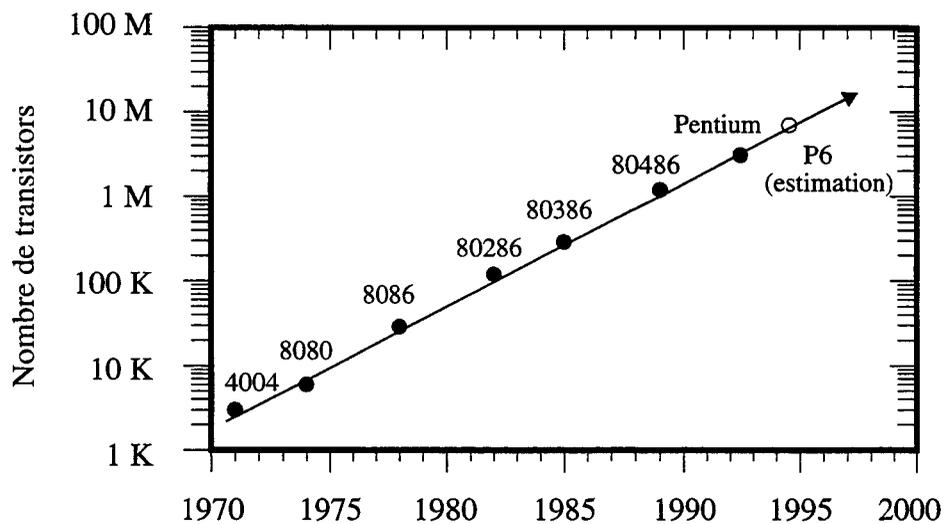
$\nabla_{\mathbf{r}}$  : gradient par rapport aux coordonnées spatiales



## CHAPITRE I

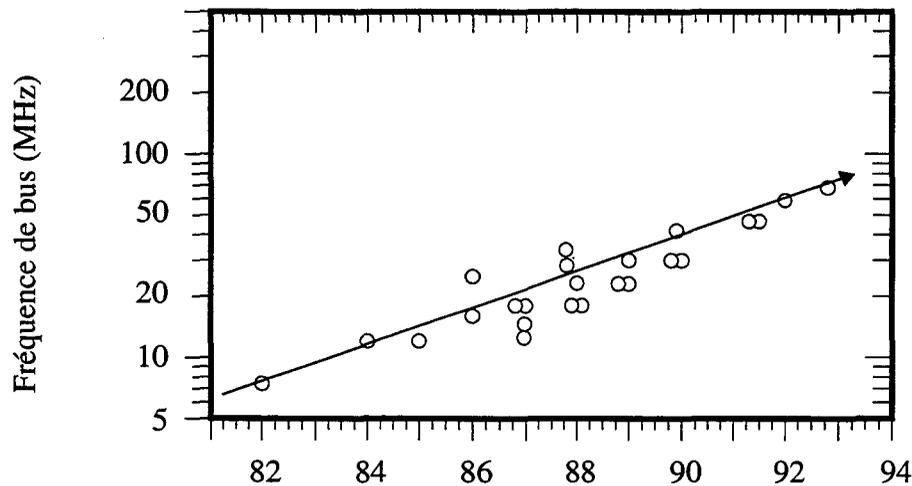
### INTRODUCTION

Les progrès réalisés en microélectronique ont permis depuis quelques décennies la fabrication de circuits dont la densité est régulièrement accrue. Ces développements sont liés à l'évolution des besoins de l'industrie électronique, qui se traduit par la création de produits nouveaux et par l'amélioration de ceux qui existent déjà. Ceci est particulièrement bien illustré par le domaine de l'informatique où les systèmes sont de plus en plus rapides et puissants, offrant des utilisations toujours plus variées. Cette évolution est rendue possible par la fabrication de circuits intégrés dont la densité, la complexité et la vitesse sont constamment augmentées. Les circuits produits ces dernières années montrent une progression exponentielle en termes de miniaturisation des composants élémentaires et de densité d'intégration. La figure ci-dessous donne à titre d'exemple le nombre de composants intégrés sur les micro-processeurs développés par Intel.



**Figure I-1** : Nombre de transistors intégrés sur les microprocesseurs fabriqués par Intel [1].

Pendant les trente dernières années, le nombre de transistors intégrés sur une puce a doublé environ tous les 18 mois, atteignant aujourd'hui plusieurs millions. La capacité des mémoires dynamiques a suivi la même progression. La phase de mise en production des DRAMs de 64 Mbits est bien engagée en 1994, alors que les circuits de 256 Mbits sont en développement [2]. La miniaturisation permet également une progression régulière de la vitesse des systèmes par l'utilisation de connexions plus courtes, ainsi que l'illustre la figure I-2.



**Figure I-2** : Evolution de la fréquence de bus des processeurs [3].

Ces progrès dans la réalisation de composants complexes dépendent du développement rapide des techniques de fabrication, et notamment de l'évolution de la lithographie haute résolution à faisceau d'électrons ou à rayons X. Ceci permet de réaliser de façon reproductible des motifs de plus en plus fins, dont les dimensions géométriques peuvent descendre en dessous de  $0.1 \mu\text{m}$ . [4]

La mise au point de la fabrication d'un dispositif utilise pour une bonne partie l'extrapolation des règles de miniaturisation utilisées pour les générations précédentes. Cependant, cette approche ne suffit pas à faire face à l'émergence de phénomènes qui étaient jusque-là négligeables, comme cela s'est produit avec le perçage des transistors MOS à forte tension de drain et avec les problèmes liés aux porteurs chauds. La miniaturisation des composants est donc également liée à une connaissance approfondie de leur fonc-

tionnement. La compréhension du comportement électrique des dispositifs s'appuie sur le développement de modèles plus sophistiqués et sur l'utilisation d'outils de simulation qui permettent de réduire le recours à des procédures expérimentales coûteuses. Même si la réalisation de certains composants précède parfois la compréhension détaillée de leur fonctionnement, le domaine de l'étude des dispositifs et celui de leur fabrication progressent ensemble en s'enrichissant mutuellement.

Les équations de dérive-diffusion sont à la base d'une méthode de simulation qui a été et reste l'une des plus populaires pour l'étude et la simulation des dispositifs MOS et bipolaires. Elles fournissent une description du transport électronique prenant en compte les deux principaux mécanismes de conduction, qui sont le mouvement d'ensemble dû au champ électrique et la diffusion due au gradient de concentration. D'abord utilisé pour une analyse unidimensionnelle [5], cette modélisation a été exploitée comme un outil standard de simulation bidimensionnel au cours des années 80 [6] [7]. L'augmentation de la puissance de calcul des systèmes informatiques a permis une utilisation intensive des méthodes de simulation afin d'accélérer la mise au point des nouvelles générations de composants bipolaires ou MOS. Afin d'en estimer quantitativement les performances et de déterminer les meilleurs compromis, il est en effet indispensable de connaître l'influence de la géométrie des différentes zones des dispositifs: profils de base et d'émetteur, ou de source, de drain et de canal. L'étude et la caractérisation de certaines structures fait même appel à des simulateurs tridimensionnels, pour étudier par exemple les effets de canal étroit, ou l'influence de structures d'isolation [7]. La topologie et les profils de dopage doivent alors être décrits en volume le plus précisément possible, de manière analytique ou en faisant appel à des simulateurs tridimensionnels de procédé de fabrication [8].

Suite aux travaux de Bløtekjaer en 1970 [9] [10], les modèles hydrodynamiques ont été largement développés dans les années 85 lorsqu'il est devenu nécessaire de considérer le flux d'énergie des porteurs de charge pour décrire le fonctionnement des dispositifs [11]. Cette méthode, dérivée de l'équation de transport de Boltzmann par le calcul de ses trois premiers moments, utilise une équation de conservation de l'énergie en plus des relations déjà pré-

sentes dans le formalisme de dérive-diffusion.

Une autre approche, amorcée dès la fin des années 60 [13], mais qui a pris son essor au cours des dernières années grâce aux progrès de l'informatique, est de chercher à résoudre l'équation de Boltzmann elle-même et non plus un jeu d'équations en dérivant. En effet, la miniaturisation faisant apparaître des champs électriques intenses, présentant de brusques variations temporelles et spatiales, il est nécessaire de décrire plus finement la population électronique dans les dispositifs ainsi que son évolution, afin de prendre en compte les phénomènes balistiques. La technique la plus utilisée dans ce but est actuellement la méthode dite de Monte Carlo, simulant l'évolution d'une population d'électrons par un procédé stochastique [12] [13] [14] [15] [16].

Le travail présenté ci-après porte en partie sur l'élaboration d'un programme de simulation de dispositif par la méthode de Monte Carlo, plus spécialement dédié aux transistors nMOS. Afin de tirer profit des simulateurs de procédé de fabrication disponibles IMPACT2 et IMPACT4 [17] [18] [19] qui fournissent une description bidimensionnelle des structures, la méthode de Monte Carlo qui est par essence tridimensionnelle est exploitée uniquement en deux dimensions. Ceci se justifie par le fait que les principaux phénomènes de transport se produisent dans le sens de la longueur des dispositifs et en profondeur, les effets latéraux étant dans la plupart des cas secondaires. La restriction à deux dimensions présente en outre l'avantage de permettre l'utilisation du programme de simulation de dispositif IMPACT3.3 [20] pour déterminer les conditions initiales utilisées par le simulateur Monte Carlo et diminuer ainsi les temps de calcul. Le programme IMPACT3.3 qui utilise les équations de dérive-diffusion a d'autre part été étendu afin d'intégrer les résultats des simulations Monte Carlo concernant le transport hors d'équilibre thermique<sup>1</sup>. Les programmes Monte Carlo et IMPACT3.3 étendu s'inscrivent ainsi dans une chaîne de simulation bidimensionnelle permettant la description des dispositifs nMOS et de leur fonctionnement électrique, y compris en termes d'effets balistiques.

---

1. Le terme hors d'équilibre thermique signifie ici que la population des électrons ne peut plus être décrite de façon valide par une distribution de type Maxwellienne caractérisée par la température du réseau cristallin.

Les programmes IMPACT peuvent a priori traiter tous les dispositifs MOS ou bipolaires réalisés en silicium. Or la simulation de transistors bipolaires nécessite une description précise des mécanismes de génération et de recombinaison de paires électrons-trous. Ces phénomènes mettent en jeu des constantes de temps de l'ordre de la microseconde qui sont hors de portée des simulations Monte Carlo pour lesquelles l'échelle de temps est plutôt de quelques femtosecondes ( $10^{-15}$ s). En outre, les différences de concentrations de porteurs inhérentes aux dispositifs bipolaires sont très importantes, rendant extrêmement difficile la définition d'une population d'électrons ou de trous à simuler par la méthode de Monte Carlo. Malgré quelques travaux réalisés dans ce domaine [21] [22], l'optimisation de transistors bipolaires par la méthode de Monte Carlo n'est pas actuellement aussi accessible que peut l'être celle des dispositifs MOS. Les études de composants réalisées dans le cadre de ce travail concernent uniquement les transistors MOS à canal n, mais il serait tout à fait possible d'effectuer un travail similaire pour des dispositifs pMOS. Il faudrait alors utiliser pour la méthode de Monte Carlo une description de la bande de valence et les probabilités d'interaction appropriées [13] [23] [24].

Les éléments de physique du transport électronique dans le silicium, décrit notamment par l'équation de Boltzmann, font l'objet du chapitre II. La justification d'une méthode de simulation originale, qui consiste en un couplage de la technique de Monte Carlo et d'une version étendue du formalisme de dérive-diffusion, y est également présentée, ainsi que les modèles utilisés par les différents programmes.

Le chapitre III donne les éléments essentiels utilisés lors de l'élaboration du simulateur Monte Carlo. Il présente de même les principaux points nécessaires à la mise en oeuvre de la méthode de couplage et à la résolution des équations étendues de dérive-diffusion.

Le chapitre IV est une application de cette méthode à l'étude du phénomène de survitesse qui accompagne la miniaturisation de transistors nMOS sous l'échelle du micron. A ces dimensions, le transport balistique prend en effet une importance grandissante dans le fonctionnement des dispositifs.

Les problèmes rencontrés lors de la miniaturisation des MOS, notamment en terme de résis-

tance aux effets canaux courts et de sensibilité aux paramètres technologiques de fabrication, poussent à envisager de nouvelles structures pour les transistors à effet de champ [25]. Le chapitre V présente les résultats obtenus lors de l'étude de dispositifs nMOS non plans, pour lesquels la grille est enterrée ou semi-enterrée, ce qui permet de réduire les effets canaux courts par rapport aux structures MOS classiques.

Enfin, la dernière partie regroupe les principales conclusions de ce travail ainsi que quelques perspectives de recherche.

## REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE I

- [1] : C. R. Barrett, *Microprocessor evolution and technology impact*, 1993 symposium on VLSI technomogy, Kyoto, Digest of technical papers p7, 1993
- [2] : L. Geppert, IEEE SPECTRUM, janvier 1994
- [3] : M. Kinugawa, M. Kakumu, *Cell technology directions for advanced MPUs and other memory-embedded logic devices*, IEDM technical digest, p37, 1993
- [4] : S. Y. Chou, H. I. Smith, D. A. Antoniadis, *Sub-100-nm-channel-length transistors fabricated using x-ray lithography*, J. Vac. Sci. Technol. B 4 (1 , janvier/février 1986
- [5] : H. K. Gummel, IEEE Trans. Elec. Dev., vol.11, p445, 1964, cité dans *The use of simulation in semiconductor technology development*, Solid State Elec., vol.33, No.6, p591-623, 1990
- [6] : S. Selberherr, *Analysis and simulation of semiconductor devices*, Ed. Spinger Verlag, 1984.
- [7] : D.C. Cole, E.M. Buturla, S.S. Furkay, K. Varahramyan, J. Slinkman, J.A. Mandelman, D.P. Foty, O. Bula, A.W. Strong, J.W. Park, T.D. Linton Jr, J.B. Johnson, M.V. Fischetti, S.E. Laux, P.E. Cottrell, H.G. Lustig, F. Pileggi, D.Katcoff, *The use of simulation in semiconductor technology development*, Solid State Elec., vol.33, No.6, p591-623, 1990
- [8] : M. R. Pinto, C. S. Rafferty, R. K. Smith, J. Bude, *ULSI technology development by predictive simulation*, IEDM technical digest, p701, 1993
- [9] : K. Bløtekjaer, *High-frequency conductivity, carrier waves and acoustic amplification in drifted semiconductor plasmas*, Ericsson Technics, vol. 2, p 125, octobre 1966
- [10] : K. Bløtekjaer, *Transport equations for electrons in two-valley semiconductors*, IEEE Trans. Elec. Dev., vol.17, No.1, p38, janvier 1970
- [11] : B. Meinerzhagen, W. L. Engl, *The influence of the thermal equilibrium approximation on the accuracy of classical two-dimensional numerical modeling of silicon*

- submicrometer MOS transistors*, IEEE Trans. Elec. Dev., vol.35, No.5, p689, mai 1988
- [12] : C. Jacoboni, L. Reggiani, *The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials*, Rev. of Modern Phys., vol.55, No.3, p645, juillet 1983
- [13] : C. Jacoboni, P. Lugli, *The Monte Carlo method for semiconductor device simulation*, Ed. S. Selberherr, Springer-Verlag, 1989
- [14] : E. Constant, *Non steady state carrier transport in semiconductor in perspective with submicron devices*. Topics in Applied Physics : Hot electron transport in semiconductor, vol.58, Ed. by L. Reggiani, Springer-Verlag.
- [15] : M. V. Fischetti, S. E. Laux, *Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects*, Phys. Rev. B, vol.38, No.14, p9721, novembre 1988
- [16] : *Monte Carlo device simulation: full band and beyond*, Ed. by K. Hess Kluwer Academic Publishers, 1991
- [17] : D. Collard, B. Baccus, E. Dubois, D. Morel, *Impact 1-2-3: an integrated 2D process/device simulator for MOS technology*, in lecture notes of the software forum, NASECODE VI conf., p16-31, 1989
- [18] : B. Baccus, D. Collard, E. Dubois, D. Morel, *IMPACT 4 - a general two-dimensional multilayer process simulator*, proc. of SISDEP III conf., p255-266, Ed. by G. Baccharani and M. Rudan, Bologna, 1988
- [19] : B. Baccus, D. Collard, E. Dubois, *Adaptative mesh refinement for multilayer process simulation using the finite element method*, IEEE Trans. on CAD, vol.11, No.3, p396, mars 1992
- [20] : E. Dubois, Thèse de doctorat en sciences des matériaux, Université des Sciences et Techniques de Lille Flandres-Artois, 1990
- [21] : Y.J. Park, T.W. Tang, D. H. Navon, *On the Monte Carlo simulation of bipolar devices*, IEDM 82 technical digest, p688, 1982

- [22] : Y.J. Park, D. H. Navon, T.W. Tang, *Monte Carlo simulation of bipolar transistors*, IEEE Trans. Elec. Dev., vol.31, No.12, p1724, décembre 1984
- [23] : G. Ottaviani, L. Reggiani, C. Canali, F. Nava, A. Alberigi-Quaranta, *Hole drift velocity in silicon*, Phys. Rev. B, vol.12, No.8, p3318, octobre 1975
- [24] : L. Reggiani, J.C. Vaissière, J.P. Nougier, D. Gasquet, *Transient regimes of hot carriers in p-type silicon*, Journal de Physique, colloque C7, supplément au No.10, tome 42, octobre 1981
- [25] : C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, B. Riccò, *A new scaling methodology for the 0.1-0.025  $\mu\text{m}$  MOSFET*, 1993 Symposium on VLSI Technology Digest of technical papers, p33, 1993



## CHAPITRE II

### MODELISATION DU TRANSPORT ELECTRONIQUE

#### II-1 INTRODUCTION

Ce chapitre est consacré à la formulation des équations et des modèles qui décrivent le transport électronique dans les semiconducteurs, et plus précisément dans le silicium. Diverses méthodes sont utilisées pour étudier les semiconducteurs en tant que matériaux et simuler les dispositifs électroniques qui les utilisent. De plus, les méthodes de simulations diffèrent également selon les dispositifs étudiés, leur géométrie, leurs dimensions caractéristiques et leur comportement électrique. Même si les méthodes particulières de type Monte Carlo, les équations hydrodynamiques et de dérive-diffusion sont toutes basées sur la formulation semi-classique de l'équation de Boltzmann, ces techniques ont chacune des spécificités d'application et des limites liées aux approximations qu'elles contiennent. La théorie du transport électronique qu'elles utilisent fait appel à certains éléments de physique théorique dont l'élaboration et la démonstration dépassent largement le cadre de ce travail. La suite du chapitre est consacrée à la description de la conduction par électrons, en vue d'étudier plus particulièrement des dispositifs nMOS, mais il est tout-à-fait possible d'effectuer la même démarche pour la conduction par trous.

#### II-2 EQUATION DE TRANSPORT DE BOLTZMANN

La théorie des phénomènes électroniques dans les semiconducteurs s'inspire de la théorie physique élaborée au XIX<sup>e</sup> siècle pour décrire les métaux. Cette théorie classique considère que l'ensemble des électrons en équilibre avec le réseau cristallin se comporte

comme un gaz, assimilé à un gaz parfait. Les théories développées se sont enrichies de résultats de mécanique quantique décrivant les interactions des électrons avec le cristal. Cependant, l'assimilation à un gaz des électrons participant à la conduction étant conservée, les phénomènes de transport électronique dans un matériau semiconducteur peuvent être étudiés avec les outils développés dans le cadre de la théorie cinétique des fluides [1]. Cela revient à considérer un gaz dilué d'électrons, soumis à différentes forces et interactions que l'on décrit en s'intéressant à l'évolution de la fonction de distribution statistique des électrons [2]. Cette fonction est notée  $f(\mathbf{r}, \mathbf{k}, t)$ , où  $\mathbf{r}$  représente le vecteur position, et  $\mathbf{k}$  le vecteur d'onde des électrons. Si l'on considère un volume élémentaire  $d^3 r \cdot d^3 k$  de l'espace des phases autour du point de coordonnées  $(\mathbf{r}, \mathbf{k})$ , ce qui correspond à un volume géométrique  $d^3 r$  et à un volume de l'espace réciproque  $d^3 k$  alors  $f(\mathbf{r}, \mathbf{k}, t) d^3 r \cdot d^3 k$  est le nombre d'électrons compris dans ce volume à l'instant  $t$ . En utilisant le théorème de Liouville, on obtient la relation:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \times \frac{d\mathbf{r}}{dt} + \nabla_{\mathbf{k}} f \times \frac{d\mathbf{k}}{dt} = 0 \quad (\text{II-1})$$

où  $\nabla_{\mathbf{r}}$  est le gradient par rapport aux coordonnées géométriques et  $\nabla_{\mathbf{k}}$  par rapport aux coordonnées de l'espace des vecteurs d'onde. Cette équation signifie que le volume des phases reste invariable, ou encore que le nombre d'états permis pour les électrons se conserve.

Dans l'équation (II-1),  $\frac{d\mathbf{r}}{dt} = \mathbf{v}$  est la vitesse d'un électron, et  $\frac{d\hbar\mathbf{k}}{dt} = \mathbf{F}$  est la somme des forces appliquées à un électron de quasi-impulsion  $\hbar\mathbf{k}$ . Il en résulte la formulation suivante:

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \times \mathbf{v} + \nabla_{\mathbf{k}} f \times \frac{\mathbf{F}}{\hbar} = 0 \quad (\text{II-2})$$

La variation temporelle de la distribution (premier terme) est due pour une part au mouvement des particules (deuxième terme), et d'autre part aux forces qui leur sont appliquées (troisième terme). Pour l'étude des semiconducteurs, les forces appliquées peuvent être séparées en deux composantes agissant sur des échelles de distance et de temps différentes: les forces extérieures, macroscopiques, et les forces internes, microscopiques. Les premiè-

res sont les forces électromagnétiques macroscopiques et les secondes correspondent à la diffusion des particules suite aux phénomènes de collision.

$$\mathbf{F} = \mathbf{F}_E + \mathbf{F}_D \quad (\text{II-3})$$

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \times \mathbf{v} + \nabla_{\mathbf{k}} f \times \frac{\mathbf{F}_E}{\hbar} + \nabla_{\mathbf{k}} f \times \frac{\mathbf{F}_D}{\hbar} = 0 \quad (\text{II-4})$$

$\mathbf{F}_E$  est la force due au seul champ électrique  $\mathbf{E}$ , dans la mesure où l'on néglige ici l'influence du champ magnétique:  $\mathbf{F}_E = -q\mathbf{E}$  pour un électron. Le dernier terme de l'équation (II-4), que nous pouvons également noter  $-\left[\frac{\partial f}{\partial t}\right]_{coll}$ , et qui sera traité de façon plus détaillée ultérieurement lors de l'étude des modèles de collision, peut être exprimé par la relation suivante [3] [4]:

$$\nabla_{\mathbf{k}} f \times \frac{\mathbf{F}_D}{\hbar} = -\left[\frac{\partial f}{\partial t}\right]_{coll} = -\int (f(\mathbf{r}, \mathbf{k}', t) S(\mathbf{k}', \mathbf{k}) (1 - f(\mathbf{r}, \mathbf{k}, t)) - f(\mathbf{r}, \mathbf{k}, t) S(\mathbf{k}, \mathbf{k}') (1 - f(\mathbf{r}, \mathbf{k}', t))) d^3 k' \quad (\text{II-5})$$

où  $S(\mathbf{k}, \mathbf{k}') d^3 k'$  représente la probabilité pour un électron de passer de l'état  $\mathbf{k}$  à un état contenu dans  $d^3 k'$  autour de  $\mathbf{k}'$ , par unité de temps. On obtient alors l'équation de transport de Boltzmann sous la forme:

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \times \mathbf{v} + \frac{1}{\hbar} \nabla_{\mathbf{k}} f \times \mathbf{F}_E = \left[\frac{\partial f}{\partial t}\right]_{coll} \quad (\text{II-6})$$

Le terme de droite de cette équation montre que celle-ci est du type intégral-différentielle. La résolution d'une telle équation à 7 variables indépendantes est extrêmement difficile, car il n'est le plus souvent pas possible de trouver une solution analytique, à moins d'imposer des hypothèses simplificatrices, notamment sur les termes de collision, afin d'en simplifier l'expression. Cette équation est cependant le point de départ de toutes les méthodes de simulation de matériaux et de composants semiconducteurs. Notons que pour résoudre l'équation de Boltzmann, il faut connaître la structure de bande d'énergie du semiconducteur. En effet, la vitesse et l'énergie sont reliées par l'équation:

$$\mathbf{v} = \frac{1}{\hbar} \left( \nabla_{\mathbf{k}} \mathcal{E} \right) \quad (\text{II-7})$$

où  $\mathcal{E}$  est l'énergie d'un électron. La connaissance de la structure de bande est également nécessaire à la modélisation des phénomènes regroupés sous l'intégrale de collision  $\left[ \frac{\partial f}{\partial t} \right]_{coll}$ .

L'ensemble du développement précédent est sujet à certaines approximations:

- \* La théorie physique sous-jacente traite le mouvement des particules de façon semi-classique, en ce sens que les trajectoires des porteurs de charge sont décrites par la mécanique classique alors que les phénomènes de diffusion utilisent le formalisme quantique et les procédés de la physique statistique.
- \* L'assimilation d'un ensemble de particules à un gaz parfait est valable si le volume propre des particules est très faible par rapport au volume occupé par le gaz et si les interactions entre particules sont négligeables. Le premier point ne pose pas de problème dans le cas d'un gaz électronique. La deuxième approximation correspond à l'approximation de Hartree: les effets sur un électron des autres électrons peuvent être décrits par le potentiel de Hartree, ce qui permet de considérer le mouvement de chaque électron séparément.
- \* Les concepts de bande d'énergie et de masse effective demandent une dimension minimale du cristal pour être valides. Il pourraient être remis en cause pour des grandeurs caractéristiques inférieures à 0.01 micron. C'est notamment le cas pour de très petits dispositifs MOS où le confinement des électrons à l'interface silicium/oxyde introduit une quantification des niveaux d'énergie.
- \* La séparation entre les forces macroscopiques et les interactions microscopiques reste valable dans la mesure où les forces extérieures ne modifient pas la structure énergétique du semiconducteur et où les phénomènes de collision ne dépendent pas directement des forces macroscopiques. Sous ces conditions, les probabilités d'interaction sont indépendantes du champ électrique.
- \* De même, pour que cette séparation soit valable, il faut que les forces macroscopiques ne subissent pas de changement important sur des distances de l'ordre de la longueur d'onde de de Broglie d'un électron.

\* Assimiler les phénomènes de diffusion à des chocs suppose que la durée de l'interaction soit brève par rapport au temps entre collisions, et que la distance parcourue entre deux chocs soit grande devant la distance pendant laquelle la particule subit l'interaction. Cette distance peut être de l'ordre du nanomètre. La durée d'interaction est alors inférieure à  $10^{-14}$ s pour une particule dont la vitesse est de l'ordre de  $10^7$ cm/s. Pour des dispositifs dont une dimension caractéristique se rapproche de l'échelle nanométrique, les collisions ne peuvent plus être considérées comme instantanées.

## II-3 RESOLUTIONS EXACTES DE L'EQUATION DE BOLTZMANN

Jusqu'à des longueurs caractéristiques de composant de l'ordre du micron, les méthodes de simulation utilisant les équations de dérive-diffusion étaient suffisantes pour décrire avec précision le comportement des dispositifs électroniques, même si des phénomènes non stationnaires se produisaient, dans la zone de pincement des canaux MOS par exemple [5]. Suite à la miniaturisation des composants sous l'échelle du micron, cette approche est devenue insuffisante, rendant nécessaire la connaissance de la distribution en vitesse des porteurs de charge. Plusieurs façons de procéder ont été développées durant les deux dernières décennies pour résoudre l'équation de Boltzmann: d'une part des méthodes itératives, d'autre part l'utilisation de développements de la fonction de répartition  $f$  en harmoniques sphériques ou à l'aide de polynômes de Legendre, et enfin des méthodes particulières. Les principaux avantages et inconvénients de ces approches sont présentés ci-dessous.

### II-3.1 Procédures itératives

La technique itérative est liée au caractère intégral-différentiel de l'équation de Boltzmann. La fonction  $f$  est exprimée sous forme d'une intégrale faisant intervenir les probabilités de collision. La méthode de résolution consiste à définir a priori une fonction test  $f_0$ , et à l'utiliser dans les estimations d'intégrales pour calculer une nouvelle solution  $f_1$ . L'opération est ainsi répétée jusqu'à convergence du résultat. Cette procédure est intéressante pour certaines études où une connaissance détaillée de la fonction de distribution est nécessaire [6], mais les calculs d'intégrales sont complexes, et la convergence est difficile à

contrôler. L'extension à une simulation de dispositif n'est pas triviale, puisqu'il faut alors évaluer la fonction de répartition en tous points du composant, et que  $f$  n'est pas une fonction locale du champ électrique et des concentrations, mais dépend aussi de la dynamique électronique dans le reste du dispositif.

### II-3.2 Développement en série

Cette méthode analytique consiste à développer la fonction de répartition en polynômes de Legendre [7] [8] [9] [10] ou en harmoniques sphériques [12]. L'équation de Boltzmann permet alors d'écrire un système d'équations reliant les coefficients des différents termes du développement, qui est ensuite résolu de façon classique. Cette méthode est cependant limitée, car elle fait le plus souvent appel à une hypothèse sur la direction du champ électrique, pris dans une direction cristallographique particulière [8]. De plus, le développement de la fonction, mathématiquement infini, doit être tronqué pour permettre une résolution numérique [11]. Enfin, l'extension à des simulations bidimensionnelles pose difficulté, ne serait-ce que par la taille des systèmes algébriques obtenus après discrétisation des équations [6] [12].

### II-3.3 Principe de résolution de l'équation de transport de Boltzmann par la méthode de Monte Carlo

Cette méthode stochastique utilise des nombres aléatoires pour simuler le comportement de particules. L'équation de Boltzmann est en fait traitée de façon semi-classique, dans le sens où les termes du premier membre  $\nabla_{\mathbf{r}} f \times \mathbf{v}$  et  $\frac{1}{\hbar} \nabla_{\mathbf{k}} f \times \mathbf{F}_E$  sont traités dans le cadre de la mécanique classique, et où le terme  $\left[ \frac{\partial f}{\partial t} \right]_{coll}$  est déterminé par un formalisme quantique. Une particule simulée par la technique de Monte Carlo subit une série de vols libres et de chocs. Les vols libres sont calculés de façon classique comme des trajectoires balistiques. La simulation des chocs, dont les probabilités sont déterminés par des calculs de mécanique quantique, utilise des tirages au sort pour le choix du type d'interaction d'une part, pour la direction prise par la particule suite au choc d'autre part. La méthode de Monte Carlo résout donc de façon implicite l'équation de transport de Boltzmann permet-

tant de simuler la dynamique du transport électronique. Sa programmation est relativement simple, une fois déterminés les modèles et les probabilités d'interactions. Elle reste de plus flexible, car il est assez facile de modifier la gestion des interactions, ou d'introduire de nouveaux modèles. La précision de cette dernière méthode est liée à la représentation statistique des porteurs de charge. Ces éléments font de la méthode de Monte Carlo l'une des plus utilisées, même si le caractère statistique impose des temps de simulation relativement longs dépendant de la précision souhaitée. Les simulations de type Monte Carlo servent également de référence pour établir d'autres modèles et outils, de par leur capacité à prendre en compte le détail de l'évolution des phénomènes électroniques.

## II-4 METHODES DERIVEES DE L'EQUATION DE BOLTZMANN

D'autres procédés répandus de simulation de dispositifs semiconducteurs utilisent des approximations de l'équation de Boltzmann. Il s'agit notamment des méthodes hydrodynamiques et de dérive-diffusion, qui font appel aux moments successifs de l'équation de transport de Boltzmann. L'obtention des moments, puis leur utilisation dans les formalismes hydrodynamique et de dérive-diffusion sont présentés ci-dessous.

### II-4.1 Moments successifs

La dérivation de ces moments peut être exprimée de diverses façons, selon les variables que l'on souhaite faire apparaître [13] [14] [15] [16] [17] [18] [19]. Le formalisme retenu ici est inspiré des notations de l'article de Lee et Tang [23] dans lequel différentes méthodes de simulation sont comparées. Les moments successifs s'obtiennent en multipliant l'équation de Boltzmann (II-6) respectivement par 1, par la quasi-impulsion  $\hbar\mathbf{k}$ , par l'énergie  $\mathcal{E}$ , et en intégrant par rapport à  $\mathbf{k}$  dans la première zone de Brillouin.

Le moment d'ordre 0 est l'équation de conservation des particules:

$$\frac{\partial n}{\partial t} + \nabla_{\mathbf{r}}(n\mathbf{V}) = nC_n \quad (\text{II-8})$$

Le moment d'ordre 1 régit l'évolution de la quantité de mouvement:

$$\frac{\partial(n\mathbf{P})}{\partial t} + \nabla_{\mathbf{r}}(nq\tilde{\mathbf{U}}) - n\mathbf{F}_E = n\mathbf{C}_P \quad (\text{II-9})$$

Le moment d'ordre 2 décrit la conservation de l'énergie

$$\frac{\partial (nW)}{\partial t} + \nabla_{\mathbf{r}} (n\mathbf{S}) - (n\mathbf{V} \cdot \mathbf{F}_E) = nC_W \quad (\text{II-10})$$

où:

$$n = n(\mathbf{r}, t) = \int f(\mathbf{r}, \mathbf{k}, t) d^3k \quad \text{est la concentration en électrons}$$

$$\mathbf{V} = \frac{1}{n} \int f(\mathbf{r}, \mathbf{k}, t) \mathbf{v} d^3k \quad \text{la vitesse moyenne}$$

$$C_n = \frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} d^3k \quad \text{taux de génération-recombinaison}$$

$$\mathbf{P} = \frac{1}{n} \int f(\mathbf{r}, \mathbf{k}, t) \hbar \mathbf{k} d^3k \quad \text{le moment moyen}$$

$$\tilde{\mathbf{U}} ; U_{ij} = \frac{1}{qn} \int f(\mathbf{r}, \mathbf{k}, t) \hbar k_i \cdot v_j d^3k \quad \text{tenseur de tension thermique}$$

$\nabla_{\mathbf{r}} \tilde{\mathbf{U}}$  est un vecteur dont la  $i$ ème coordonnée est la divergence de la ligne  $i$  du tenseur considérée comme un vecteur.

$$C_{\mathbf{P}} = \frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} \hbar \mathbf{k} d^3k \quad \text{taux de changement du moment [29] [30]}$$

$$W = \frac{1}{n} \int f(\mathbf{r}, \mathbf{k}, t) \mathcal{E} d^3k \quad \text{énergie cinétique moyenne}$$

$$\mathbf{S} = \frac{1}{n} \int f(\mathbf{r}, \mathbf{k}, t) \mathcal{E} \mathbf{v} d^3k \quad \text{densité moyenne de flux d'énergie}$$

$$C_W = \frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} \mathcal{E} d^3k \quad \text{taux de changement d'énergie cinétique dû aux phénomènes de collision}$$

Ces différentes grandeurs sont des fonctions de la position et du temps  $(\mathbf{r}, t)$ .

## II-4.2 Formalisme hydrodynamique

Les méthodes hydrodynamiques utilisent les moments d'ordre 0, 1 et 2 définis précédemment. La résolution des équations (II-8) (II-9) et (II-10) nécessite l'évaluation de chacun des termes liés aux collisions  $C_n$ ,  $C_p$ ,  $C_w$ , dont la précision dépend des approximations réalisées. De plus, il est nécessaire de fixer des conditions supplémentaires, pour imposer la fermeture du système d'équations à résoudre. En effet, chaque moment d'ordre  $i$  contient un terme d'ordre  $i+1$ :  $\tilde{U}$  dans l'expression du moment d'ordre 2 (II-9) et  $S$ , terme d'ordre 3, dans celle du moment d'ordre 3 (II-10).

Pour le moment d'ordre 0,  $C_n$  est le taux de génération-recombinaison des paires électrons-trous:

$$\frac{\partial n}{\partial t} + \nabla_{\mathbf{r}}(n\mathbf{V}) = G - R \quad (\text{II-11})$$

avec  $G$  terme de génération,  $R$  de recombinaison.

Pour les moments d'ordre 1 et 2,  $C_p$  et  $C_w$  sont de façon habituelle exprimés à l'aide des approximations suivantes:

$$C_p = -\frac{P}{\tau_p} \quad (\text{II-12})$$

$$C_w = -\frac{W - W_0}{\tau_w} \quad (\text{II-13})$$

où  $\tau_p$  et  $\tau_w$  sont les temps de relaxation du moment et de l'énergie, caractérisant le retour à l'équilibre du système quand les forces extérieures sont supprimées, et  $W_0$  l'énergie cinétique moyenne des porteurs en équilibre thermique avec le réseau cristallin. Les temps de relaxation dépendent eux-même de l'énergie moyenne  $W$  [34].

La fermeture du système d'équations dans le cadre du formalisme hydrodynamique se fait classiquement en approximant la densité de flux d'énergie à l'aide d'un terme convectif et d'un terme de flux de chaleur.

$$S \approx \mathbf{Q} + W\mathbf{V} + k_B \tilde{T}\mathbf{V} \quad (\text{II-14})$$

$$\mathbf{Q} = \frac{1}{n} \int \frac{1}{2} f(\mathbf{r}, \mathbf{k}, t) (\mathbf{v} - \mathbf{V})^2 m(\mathbf{k}) (\mathbf{v} - \mathbf{V}) d^3k \quad (\text{II-15})$$

où  $m(\mathbf{k})$  est la masse effective d'un électron de vecteur d'onde  $\mathbf{k}$  et

$$\tilde{\mathbf{T}} ; T_{ij} = \frac{1}{k_B n} \int f(\mathbf{r}, \mathbf{k}, t) m(\mathbf{k}) (v_i - V_i) \cdot (v_j - V_j) d^3k \quad (\text{II-16})$$

est le tenseur de température. En fait, la température est souvent considérée comme scalaire, et on suppose  $\mathbf{Q}$  proportionnel au gradient de température. En faisant ces approximations, on donne une valeur du terme  $\mathbf{S}$  d'ordre 3 en fonction de la vitesse et de l'énergie moyennes, ce qui permet la fermeture du système d'équations. Cependant, ces hypothèses simplificatrices peuvent être mises en défaut [20], notamment la température considérée comme scalaire, l'expression du flux de chaleur proportionnel au gradient de température, et le temps de relaxation  $\tau_p$  fonction uniquement de l'énergie [14] [23] [24], par exemple dans le cas de fortes variations du champ électrique.

### II-4.3 Formalisme des équations de dérive-diffusion

La simulation de dispositifs microélectroniques par la résolution des équations de dérive-diffusion a été très largement utilisée dans le cadre de la compréhension de leur fonctionnement et dans celui de leur optimisation. Ce type de simulation a permis de simuler des dispositifs MOS et bipolaires, tant que leurs dimensions caractéristiques restaient assez grandes pour ne pas mettre en jeu d'effet non-stationnaire [25] [26]. Même lorsque certains effets de sur vitesse apparaissaient, par exemple dans la zone pincée des canaux MOS, l'utilisation du formalisme de dérive-diffusion a continué de fournir des résultats valides sur le fonctionnement global des dispositifs, le transport hors d'équilibre ayant peu d'influence pour les transistors dits 'longs'. Ce modèle classique est cependant mis en défaut lorsque les longueurs caractéristiques sont du même ordre de grandeur que la longueur de relaxation en énergie des électrons [5].

Il est possible de retrouver les équations de dérive-diffusion, dans un formalisme étendu, en utilisant les deux premiers moments de l'équation de Boltzmann [27] [28]. En introduisant la densité de courant  $\mathbf{J}_n = -qn\mathbf{V}$  dans l'expression (II-11) du moment

d'ordre 0, on trouve une nouvelle expression:

$$q \frac{\partial n}{\partial t} - \nabla_{\mathbf{r}} \mathbf{J}_n = q(G - R) \quad (\text{II-17})$$

Pour le moment d'ordre 1,  $\mathbf{C}_P$  est considéré de façon usuelle comme proportionnel à la vitesse moyenne des porteurs:  $\mathbf{C}_P = -\frac{q\mathbf{V}}{\mu}$  où  $\mu$  est la mobilité, dont la valeur peut dépendre des concentrations en dopants mises en jeu et des forces présentes, principalement celles dues au champ électrique. Dans le cas le plus général, cette simple relation ne suffit pas à décrire des situations où l'isotropie est mise en défaut, et l'on a plutôt

$$\mathbf{C}_P = -q\tilde{\mu}^{-1}\mathbf{V} \quad (\text{II-18})$$

où  $\tilde{\mu}^{-1}$  est l'inverse du tenseur de mobilité. Le premier moment s'écrit alors:

$$\frac{\partial(n\mathbf{P})}{\partial t} + \nabla_{\mathbf{r}}(nq\tilde{\mathbf{U}}) + nq\mathbf{E} = -qn\tilde{\mu}^{-1}\mathbf{V} \quad (\text{II-19})$$

Pour un régime stationnaire, la dérivée par rapport au temps est nulle, et il reste

$$\nabla_{\mathbf{r}}(nq\tilde{\mathbf{U}}) + nq\mathbf{E} = -qn\tilde{\mu}^{-1}\mathbf{V} \quad (\text{II-20})$$

soit en introduisant la densité de courant

$$\mathbf{J}_n = -qn\mathbf{V} = nq\tilde{\mu}\mathbf{E} + nq\tilde{\mu}\nabla_{\mathbf{r}}\tilde{\mathbf{U}} + q\tilde{\mu}\tilde{\mathbf{U}}\nabla_{\mathbf{r}}n \quad (\text{II-21})$$

En posant  $\tilde{\mu}\tilde{\mathbf{U}} = \tilde{\mathbf{D}}$  où  $\tilde{\mathbf{D}}$  est le tenseur de diffusivité, on retrouve une expression généralisée de la relation d'Einstein. La densité de courant peut alors être exprimée par l'équation:

$$\mathbf{J}_n = -qn\mathbf{V} = nq\tilde{\mu}\mathbf{E} + nq\tilde{\mu}\nabla_{\mathbf{r}}\tilde{\mathbf{U}} + q\tilde{\mathbf{D}}\nabla_{\mathbf{r}}n \quad (\text{II-22})$$

Si l'on suppose une mobilité scalaire et une tension thermique scalaire constante, on retrouve la formulation usuelle:

$$\mathbf{J}_n = -qn\mathbf{V} = nq\mu\mathbf{E} + qD\nabla_{\mathbf{r}}n \quad (\text{II-23})$$

avec la relation d'Einstein  $D = V_T\mu = \frac{k_B T}{q}\mu$ .

Les équations classiques de dérive-diffusion, qui ont été abondamment utilisées pour l'étude de différents dispositifs semiconducteurs, font apparaître clairement la composante du courant due aux forces extérieures, sous l'effet du champ électrique, et la composante de diffusion liée aux variations de concentration en porteurs de charges. Cependant, ces équations, par les approximations qu'elles contiennent pour modéliser la mobilité et la diffusivité, ne peuvent pas décrire les phénomènes de transport hors d'équilibre caractéristiques des dispositifs submicroniques. Les résultats qu'elles fournissent, même à l'aide de modèles détaillés de mobilité, sont notamment moins performants que ceux obtenus par la méthode hydrodynamique [6] [19] dans la mesure où les simplifications introduites sont plus abruptes.

L'équation (II-22) qui correspond au formalisme de dérive-diffusion que nous appelons étendu ne permet pas en tant que telle la résolution et le calcul d'une solution. En effet, la même remarque que pour les équations des trois premiers moments du formalisme hydrodynamique peut être formulée ici à propos de la fermeture du système d'équation. L'intérêt du formalisme de dérive-diffusion étendu réside dans le fait que cette équation, de forme usuelle, permet de prendre en compte des phénomènes non-stationnaires, hors de l'équilibre thermique. Les valeurs de mobilité et de tension thermique ne sont plus dans ce cas des fonctions locales du champ électrique et des concentrations en dopants, mais dépendent du transport électronique dans l'ensemble du dispositif. L'évaluation des valeurs de ces paramètres ne peut être obtenue que par une méthode de simulation plus précise, qui rend compte de leur caractère non local.

## II-5 CHOIX D'UNE STRATEGIE DE RESOLUTION

Afin d'établir une stratégie de simulation de dispositifs submicroniques, il paraît intéressant de rappeler quelques qualités et inconvénients des différentes méthodes que nous avons présentées précédemment.

## II-5.1 Monte Carlo

### Inconvénients:

- \* Les temps de calcul sont liés aux nombre de particules simulées et à la précision souhaitée, et constituent parfois un obstacle à l'utilisation intensive de la méthode de Monte Carlo. Ce point doit toutefois être relativisé, dans la mesure où l'environnement informatique est de plus en plus performant. Le recours à des calculateurs rapides, éventuellement vectoriels ou parallèles permet de diminuer considérablement les temps de simulation.
- \* L'aspect stochastique de cette méthode de simulation se traduit par un "bruit numérique", d'autant plus faible que la population de particules simulées est grande. Ce bruit numérique rend très délicat le calcul fiable de dérivées spatiales, temporelles ou de courant: gradients, tranconductances, capacités.
- \* Lors des simulations de dispositif, une proportion importante du temps de calcul est consacrée à la simulation des zones à fort dopage, à faible champ, où l'équilibre thermique est réalisé.

### Avantages:

- \* L'échelle temporelle des simulations Monte Carlo permet de prendre en compte des phénomènes dont la durée est de l'ordre de grandeur des temps de relaxation, c'est-à-dire inférieure à la picoseconde.
- \* La mise en oeuvre est relativement simple, une fois définis la structure de bande d'énergie et les phénomènes de collision.
- \* L'équation de transport de Boltzmann est résolue exactement, notamment pour le transport hors d'équilibre thermique.
- \* Il est possible de simuler facilement des phénomènes transitoires dans la mesure où la méthode permet d'obtenir l'historique des porteurs de charge.
- \* Les grandeurs microscopiques et les distributions des particules sont accessibles.

## II-5.2 Méthodes dérivées, hydrodynamique ou dérive-diffusion

### Avantages:

- \* La résolution est relativement rapide, si l'on dispose d'algorithmes non linéaires assez robustes et performants pour résoudre les systèmes d'équations qui en résultent. Là encore, l'évolution des calculateurs permet une diminution significative des temps de simulation, et la simulation de dispositifs plus complexes.
- \* La précision des résultats obtenus, le plus souvent par des méthodes itératives, peut être contrôlée à l'aide des critères de convergence, ce qui permet le calcul de grandeurs dérivées.
- \* Ces méthodes ont fait leurs preuves pour l'étude du transport électronique faiblement hors d'équilibre thermique [5].
- \* La forme des équations permet de prendre en compte des phénomènes liés au transport dont les temps caractéristiques sont longs, par exemple les termes de génération et recombinaison des électrons avec les trous. Les simulations intègrent alors une échelle temporelle de l'ordre de la microseconde.

### Inconvénients:

- \* La précision avec laquelle ces méthodes décrivent la dynamique du transport électronique est limitée par les approximations faites pour former le système d'équations à résoudre. Notamment, les phénomènes de collision ne peuvent être pris en compte que par l'intermédiaire des temps de relaxation.
- \* Les seules grandeurs calculées sont des grandeurs moyennes, macroscopiques, telles que concentrations, courants.

## II-5.3 Couplage

La liste des caractéristiques précédente laisse clairement entrevoir la complémentarité de la technique de Monte Carlo et des méthodes hydrodynamiques ou de dérive-diffusion. Un couplage des deux méthodes est un compromis intéressant, permettant de combiner les avantages propres à chacune [21] [22].

Il est par exemple utile d'appliquer la méthode de Monte Carlo au matériau semiconducteur afin de déterminer les paramètres utilisés dans les équations hydrodynamiques, tels que les

temps de relaxation du moment et de l'énergie [23]. La technique de Monte Carlo, qui donne une résolution plus précise de l'équation de Boltzmann, constitue aussi un moyen de valider les modèles qui sont utilisés lors de simulations hydrodynamiques. Cependant, il reste très difficile de mettre au point des modèles hydrodynamiques suffisamment généraux pour être efficaces en toutes situations, principalement à cause du caractère non local de paramètres tels que la mobilité. Ceci fait que l'on doit recourir aux simulations Monte Carlo pour corriger les modèles selon l'application. De plus, comme cela a été développé précédemment, les conditions imposées lors de la fermeture du système d'équations hydrodynamiques limitent cette approche.

Une autre approche suggérée par Bandhyopadhyay [28], reprise par Kosina et Selberherr [29] [30], et développée également dans le cadre de ce travail [31] [32] [33] consiste à utiliser simultanément la méthode de Monte Carlo et le formalisme de dérive-diffusion étendu. La technique de Monte Carlo est utilisée pour évaluer la distribution spatiale des paramètres influencés par le transport hors d'équilibre (mobilité, tension thermique). Un avantage de cette méthode de couplage réside dans la possibilité d'utiliser le fenêtrage pour limiter les simulations Monte Carlo aux zones de transport non stationnaire, les zones en équilibre thermique pouvant être prises en compte par les équations de dérive-diffusion classiques. Cette approche couplée permet de plus de tenir compte des deux échelles de temps (picoseconde et microseconde) pour l'étude de phénomènes de transport mettant en cause des échelles de temps différentes.

C'est cette dernière approche que nous avons retenue pour l'études de dispositifs MOS submicroniques. Elle permet de conjuguer l'efficacité de la méthode classique de dérive-diffusion lorsque l'équilibre thermique est atteint, pour la simulation des zones de source et drain ainsi que pour les faibles polarisations, et la précision de la méthode Monte Carlo pour la description des phénomènes hors d'équilibre. Cette méthode de couplage permet de palier certains inconvénients de l'une ou l'autre méthode, mais reste limitée par les hypothèses sous-jacentes à l'équation cinétique de Boltzmann. Ces limitations seront abordées lors de l'étude de la mise en oeuvre numérique de la méthode de couplage, qui suit la description des modèles utilisés par les deux types de simulations.

## II-6 MODELES UTILISES POUR LES SIMULATIONS DE DERIVE-DIFFUSION

Les équations de dérive-diffusion correspondant aux deux premiers moments de l'équation de Boltzmann sont résolues simultanément avec l'équation de Poisson permettant de calculer la distribution de potentiel:

$$\nabla_{\mathbf{r}} \left( \varepsilon \nabla_{\mathbf{r}} \Psi \right) = -q (p - n + N_D - N_A) \quad (\text{II-24})$$

où  $N_D$  et  $N_A$  sont respectivement les concentrations en donneurs et en accepteurs. Le champ électrique est déduit simplement:

$$\mathbf{E} = -\nabla_{\mathbf{r}} \Psi \quad (\text{II-25})$$

Les équations de conservation de charge (II-17) et de continuité de courant (II-23) permettent de traiter la conduction par électrons, et des équation similaires décrivent la conduction par trous. Ces équations doivent être complétées par les modèles décrivant les mécanismes de génération et recombinaison d'une part, la mobilité et la diffusivité d'autre part. Dans la mesure où ces modèles sont utilisés dans les cas de transport à l'équilibre ou faiblement hors d'équilibre, les formulations classiques, largement utilisées et éprouvées, déjà incorporées au simulateur de dispositifs IMPACT3.3 [18] [39] sont conservées.

### II-6.1 Génération-recombinaison

Les modèles de génération-recombinaison sont utilisés à la fois dans le formalisme classique de dérive-diffusion et dans le formalisme étendu. Ils décrivent plusieurs mécanismes de génération-recombinaison en volume.

#### Génération-recombinaison Shockley-Read-Hall [35] [36]:

Les phénomènes mettant en oeuvre un piège monovalent situé au niveau d'énergie  $\mathcal{E}_i$ , au milieu de la bande interdite, sont exprimés par le terme suivant [25]:

$$(G - R)_{SRH} = \frac{n_i^2 - np}{\tau_n (p + n_i) + \tau_p (n + n_i)} \quad (\text{II-26})$$

où  $n$ ,  $p$ ,  $n_i$  sont respectivement les concentrations en électrons, en trous, la concentration

intrinsèque et  $\tau_n, \tau_p$  les durées de vie des électrons et des trous. La dépendance de ces durées de vie par rapport à la concentration en impuretés ionisées  $N_I = |N_D| + |N_A|$  peut être exprimée par les équations:

$$\tau_n = \frac{\tau_{n0}}{1 + N_I/N_n^{ref}} ; \tau_p = \frac{\tau_{p0}}{1 + N_I/N_p^{ref}} \quad (\text{II-27})$$

avec les valeurs de paramètres suivantes [18] [25]:

**Table 1**

$\tau_{n0}$	$3.95 \cdot 10^{-5} \text{ s}$
$N_n^{ref}$	$7.1 \cdot 10^{15} \text{ cm}^{-3}$
$\tau_{p0}$	$3.52 \cdot 10^{-5} \text{ s}$
$N_p^{ref}$	$7.1 \cdot 10^{15} \text{ cm}^{-3}$

### Recombinaison Auger [37] [38]

Les différents mécanismes intervenant dans la génération-recombinaison Auger sont regroupés dans l'expression:

$$(G - R)_{AUG} = (n_i^2 - np) \times (C_n n + C_p p) \quad (\text{II-28})$$

avec pour valeurs des paramètres [25] [18]:

**Table 2**

$C_n$	$2.7 \cdot 10^{-31} \text{ cm}^6 \text{ s}^{-1}$
$C_p$	$9.9 \cdot 10^{-32} \text{ cm}^6 \text{ s}^{-1}$

## **II-6.2 Mobilité et diffusivité**

Les termes de mobilité et de diffusivité sont modélisés de façon classique. Dans le cas de transport hors d'équilibre, les valeurs de mobilité et de diffusivité peuvent être évaluées dans le cadre de la méthode de couplage par des simulations Monte Carlo. Dans les cas où le transport à l'équilibre prédomine, l'expression de la mobilité, issue des travaux de Yamaguchi [40] [41], reflète la dépendance envers la concentration en impuretés ionisées

et l'influence du champ électrique tant parallèle que perpendiculaire au courant. Pour un transistor MOS, la réduction de mobilité due au champ électrique vertical est exprimée par le terme lié au champ perpendiculaire au courant, alors que le terme lié au champ parallèle modélise la saturation en vitesse des porteurs.

$$\mu = \mu' (N_p, E_{\perp}) \cdot f(E_{\parallel})$$

$$\mu' (N_p, E_{\perp}) = \mu_0 \left[ 1 + \frac{N_I}{N_I/S + N_r} \right]^{-1/2} (1 + aE_{\perp})^{-1/2}$$

$$f(E_{\parallel}) = \left[ 1 + \left[ \frac{\mu' (N_p, E_{\perp}) E_{\parallel}}{V_c} \right]^2 \cdot \left[ \frac{\mu' (N_p, E_{\perp}) E_{\parallel}}{V_c} + G \right]^{-1} + \left[ \frac{\mu' (N_p, E_{\perp}) E_{\parallel}}{V_s} \right]^2 \right]^{-1/2} \quad (\text{II-29})$$

$E_{\perp}$ ,  $E_{\parallel}$  sont les composantes du champ électrique perpendiculaire et parallèle au vecteur densité de courant. Les valeurs des paramètres utilisés sont les suivantes pour les électrons [39]:

**Table 3**

$\mu_0$	$1127 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$
$S$	350
$N_r$	$3 \cdot 10^{16} \text{ cm}^{-3}$
$a$	$4.156 \cdot 10^{-6} \text{ cmV}^{-1}$
$V_c$	$8.824 \cdot 10^6 \text{ cms}^{-1}$
$V_s$	$1.126 \cdot 10^7 \text{ cms}^{-1}$
$G$	13,18

## II-7 MODELES UTILISES POUR LES SIMULATIONS MONTE CARLO

### II-7.1 Bande d'énergie

L'un des ingrédients essentiels de la résolution des équations de transport électronique par la méthode de Monte Carlo est la description de la structure de bande d'énergie du semiconducteur considéré. Notre étude étant consacrée aux transistors silicium NMOS, il est nécessaire de définir avec la plus grande précision la bande de conduction de ce matériau. La bande de conduction la plus basse du silicium [42] est caractérisée par six vallées principales dans la première zone de Brillouin. Les minima sont situés le long des axes  $\langle 100 \rangle$ , en  $\|\mathbf{k}\| = \mathbf{k}_0 = 9.836 \cdot 10^7 \text{ cm}^{-1}$ , comme indiqué sur la figure II-1. Afin de simplifier les notations,  $\mathbf{k}$  désigne dans la suite le vecteur d'onde avec pour origine le centre de la vallée considérée. Plusieurs descriptions des vallées sont possibles, plus ou moins raffinées. La solution la plus simple est de considérer les vallées comme sphériques (isotropes) et paraboliques (masse effective constante, énergie proportionnelle à  $\mathbf{k}^2$ ).

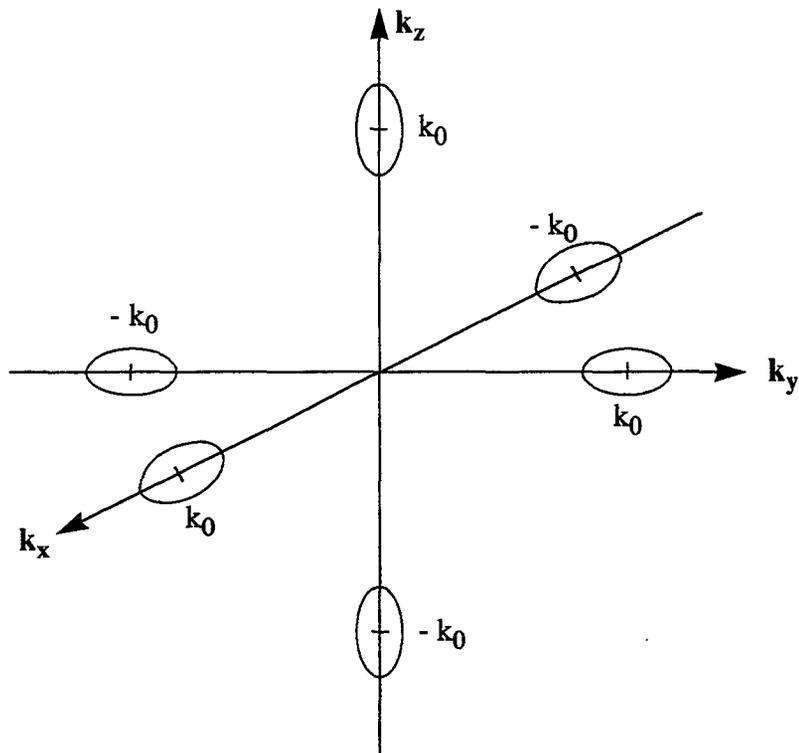
$$\mathcal{E} = \frac{\hbar^2 \mathbf{k}^2}{2m} \quad (\text{II-30})$$

Une meilleure description de ces vallées est obtenue en tenant compte de leur caractère anisotrope: les vallées sont alors ellipsoïdales

$$\mathcal{E} = \frac{\hbar^2 \mathbf{k}_l^2}{2m_l} + \frac{\hbar^2 \mathbf{k}_t^2}{2m_t} \quad (\text{II-31})$$

$l$  indiquant la direction longitudinale, selon l'axe principal de la vallée et  $t$  la direction transversale. On peut améliorer la précision de cette description en considérant que les masses effectives  $m_l$  et  $m_t$  dépendent de l'énergie. Les vallées sont alors ellipsoïdales, mais non paraboliques. Ceci est couramment exprimé en introduisant un facteur correctif de non-parabolicité, noté  $\alpha$  [44].

$$\mathcal{E}(1 + \alpha \mathcal{E}) = \frac{\hbar^2 \mathbf{k}_l^2}{2m_l} + \frac{\hbar^2 \mathbf{k}_t^2}{2m_t} \quad (\text{II-32})$$



**Figure II-1** : représentation des six vallées ellipsoïdales le long des axes  $\langle 100 \rangle$ .

Cette description de la bande d'énergie de conduction n'utilisant que les vallées principales est valable à énergie relativement faible et est suffisante pour des valeurs atteignant environ 1eV [45]. Pour étudier plus spécifiquement les électrons de haute énergie (porteurs chauds), il est nécessaire de compléter la description, soit en ajoutant les vallées supérieures correspondant aux énergies élevées [46] [47] [48] [49], soit en abandonnant la description analytique et en adoptant une structure de bande tabulée [50] [51]. La solution retenue dans ce travail est d'utiliser la description analytique de l'équation (II-32) pour des vallées ellipsoïdales et non paraboliques. En effet, dans la mesure où la miniaturisation des transistors s'accompagne d'une réduction des tensions de polarisation, les valeurs d'énergie cinétique des électrons dépassent rarement 1eV. Les paramètres utilisés pour caractériser les vallées sont donnés dans la table 4.

Table 4

$m_0$	$9.108 \cdot 10^{-31} \text{kg}$
$m_l$	$0.9163 m_0$
$m_t$	$0.1905 m_0$
$\alpha$	$0.5 \text{ eV}^{-1}$

Le choix d'une telle expression analytique présente l'avantage de rendre les simulations Monte Carlo bien plus rapides que celles réalisées à l'aide d'un modèle tabulé, plus précis, tout en gardant une bonne description à basse et moyenne énergie. Nous reviendrons sur le bien fondé de ce choix lors de la présentation des résultats de simulation. La description de la structure de bande étant établie, il est possible d'en déduire la formulation de la vitesse des électrons et des probabilités d'interaction avec le réseau cristallin.

A partir de la relation (II-7) et de l'équation (II-32), on peut exprimer les composantes de la vitesse d'un électron de moment  $\hbar \mathbf{k}$  et d'énergie  $\mathcal{E}$  de façon analytique par:

$$v_l = \frac{\hbar k_l}{m_l (1 + 2\alpha \mathcal{E})} \quad (\text{II-33})$$

$$v_t = \frac{\hbar k_t}{m_t (1 + 2\alpha \mathcal{E})} \quad (\text{II-34})$$

Au passage, il est à noter que la composante longitudinale de la vitesse dépend de la composante transversale de  $\mathbf{k}$  par l'intermédiaire de l'énergie et que de même,  $v_t$  dépend de  $k_l$ .

Afin de faciliter les calculs effectués à partir de l'équation des vallées de bande de conduction, il est utile d'introduire la transformation de Herring-Vogt qui associe à un électron d'énergie  $\mathcal{E}$  et de vecteur d'onde  $\mathbf{k}$  de coordonnées  $k_x, k_y, k_z$  un vecteur  $\mathbf{k}^*$  de coordonnées:

$$\mathbf{k}_x^* = \sqrt{\frac{m_0}{m_x}} \mathbf{k}_x ; \mathbf{k}_y^* = \sqrt{\frac{m_0}{m_y}} \mathbf{k}_y ; \mathbf{k}_z^* = \sqrt{\frac{m_0}{m_z}} \mathbf{k}_z \quad (\text{II-35})$$

où  $m_x, m_y, m_z$  sont les masses effectives selon les différents axes ( $m_l$  ou  $m_t$  selon l'orientation de la vallée). Cette transformation permet d'exprimer l'équation (II-32) sous forme

sphérique

$$\mathcal{E}(1 + \alpha\mathcal{E}) = \frac{\hbar^2 \mathbf{k}^{*2}}{2m_0} \quad (\text{II-36})$$

Par cette même transformation, le volume élémentaire  $d^3k$  devient  $d^3k = \left[\frac{m_d}{m_0}\right]^{3/2} d^3k^*$  où  $m_d = \sqrt[3]{m_i^2 m_l}$  est la masse effective de densité d'état.

## II-7.2 Modélisation des interactions subies par les électrons

Les probabilités d'interactions sont traitées dans le formalisme de la mécanique quantique, à l'aide de la "loi d'or de Fermi" [52], encore appelée théorie des perturbations dépendantes du temps [3] ou théorie des transitions quantiques [2]. Pour les dispositifs étudiés, on peut considérer que le semiconducteur n'est pas dégénéré, ce qui permet de simplifier l'expression des termes de collision (II-5), car:  $(1 - f(\mathbf{r}, \mathbf{k}, t)) \approx 1$

$$\left[\frac{\partial f}{\partial t}\right]_{coll} = \int (f(\mathbf{r}, \mathbf{k}', t) S(\mathbf{k}', \mathbf{k}) - f(\mathbf{r}, \mathbf{k}, t) S(\mathbf{k}, \mathbf{k}')) d^3k' \quad (\text{II-37})$$

Les mécanismes de collisions sont de plusieurs types. Pour le silicium, il s'agit principalement de collisions intravallées provoquées par des phonons acoustiques (collisions élastiques), et intervallées de types f (vers une vallée d'axe orthogonal à la vallée de départ) et g (entre vallées d'axes colinéaires). Il faut ajouter à cela, pour le silicium dopé, les interactions avec les impuretés ionisées.

### II-7.2.1 Collisions intervallées

Pour un phonon d'énergie  $k_B T_{Ph}$ , l'énergie après interaction d'un électron d'énergie incidente  $\mathcal{E}$  est  $\mathcal{E}_a = \mathcal{E} + k_B T_{Ph}$  dans le cas d'une absorption de phonon ou  $\mathcal{E}_e = \mathcal{E} - k_B T_{Ph}$  dans le cas d'une émission. La forme de la bande de conduction étant connue, la probabilité pour un électron d'énergie  $\mathcal{E}$  d'interagir avec un phonon dans une collision de type f ou g peut être déduite.

Probabilité d'absorption:

$$\lambda_a = \frac{n_v m_d^{3/2} q^2 \mathcal{D}^2}{\hbar^2 \sqrt{2} \pi \rho k_B T_{Ph}} \left( \frac{1}{\exp(T_{Ph}/T_L) - 1} \right) (\mathcal{E}_a (1 + \alpha\mathcal{E}_a))^{1/2} (1 + 2(\alpha\mathcal{E}_a)) \quad (\text{II-38})$$

où  $n_v$  est le nombre de vallées de destination possibles (4 pour le type f, 1 pour le type g, la vallée de destination étant respectivement orthogonale ou colinéaire à la vallée de départ) et  $\rho$  la densité du semiconducteur.

Dans l'expression (II-38), on a défini:

$$\lambda_a = \frac{1}{\tau_a} \quad (\text{II-39})$$

$\lambda_a$  est l'inverse du temps moyen  $\tau_a$  entre deux interactions de même type.  $\lambda_a$  est donc le nombre moyen d'interactions par seconde.

Probabilité d'émission:

$$\lambda_e = \frac{n_v m_d^{3/2} q^2 \mathcal{D}^2}{\hbar^2 \sqrt{2} \pi \rho k_B T_{Ph}} \left( \frac{1}{\exp(T_{Ph}/T_L) - 1} + 1 \right) (\mathcal{E}_e (1 + \alpha \mathcal{E}_e))^{1/2} (1 + 2(\alpha \mathcal{E}_e)) \quad (\text{II-40})$$

si  $\mathcal{E} > k_B T_{Ph}$  et  $\lambda_e = 0$  dans le cas contraire.

Les probabilités d'absorption et d'émission sont représentées sur la figure II-2 pour les interactions de type f et sur la figure II-3 pour celles de type g.

### II-7.2.2 Collisions intravallées

L'énergie de l'électron est conservée car il s'agit de chocs quasi-élastiques, et la probabilité d'interaction est:

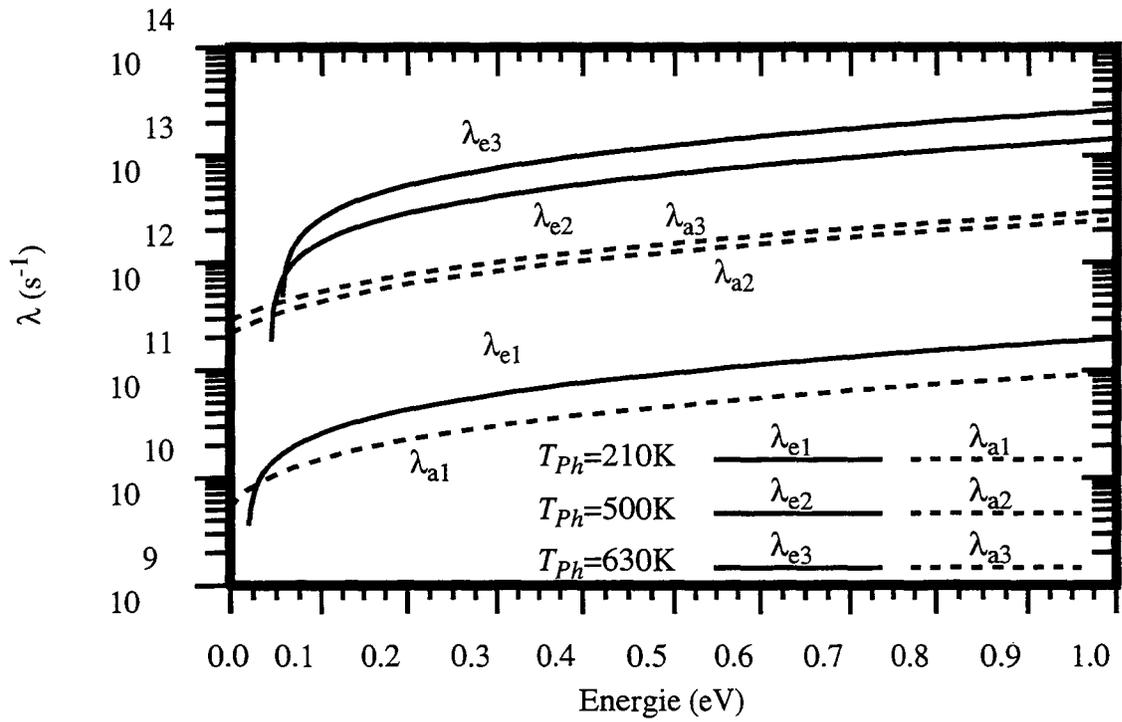
$$\lambda_{ac} = \frac{\sqrt{2} m_d^{3/2} q^2 \mathcal{E}_{ac}^2 k_B T_L}{\hbar^4 \pi \rho v_s^2} (\mathcal{E} (1 + \alpha \mathcal{E}))^{1/2} (1 + 2(\alpha \mathcal{E})) \quad (\text{II-41})$$

$v_s$  étant la vitesse du son dans le cristal. La figure II-4 présente cette probabilité en fonction de l'énergie.

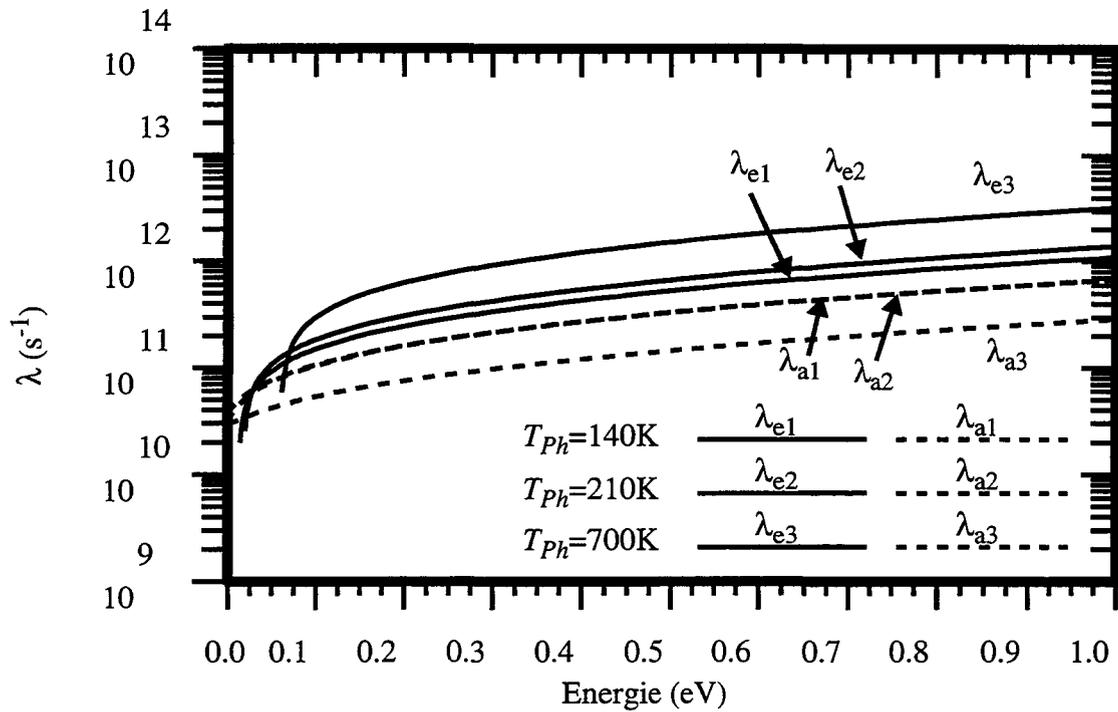
Les modèles de collisions intra et intervallées ont été abondamment étudiés (voir par exemple [53] [55] [56]) et les valeurs classiques données par Canali [54] [43] ont été retenues pour les énergies de phonons et les constantes de couplage associées à la structure de bande d'énergie décrite précédemment. Ces valeurs sont rappelées dans la table 5:

Table 5

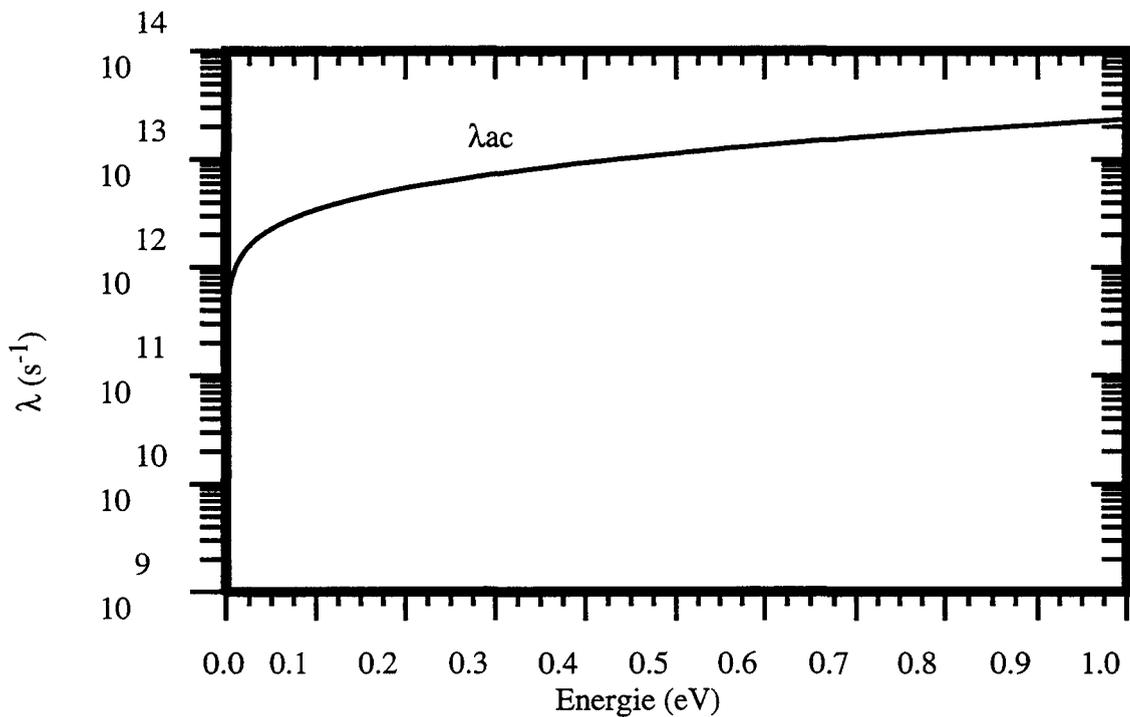
$T_{Ph}$	$\mathcal{D}$
210 K	$1.5 \cdot 10^7 \text{ eV.cm}^{-1}$
500 K	$2.5 \cdot 10^8 \text{ eV.cm}^{-1}$
630 K	$4.0 \cdot 10^8 \text{ eV.cm}^{-1}$
140 K	$5.0 \cdot 10^7 \text{ eV.cm}^{-1}$
210K	$8.0 \cdot 10^7 \text{ eV.cm}^{-1}$
700K	$3.0 \cdot 10^8 \text{ eV.cm}^{-1}$
	$\mathcal{E}_{ac} = 9.0\text{eV}$



**Figure II-2** : Nombres moyens d'interactions de type f par seconde en fonction de l'énergie des électrons: émission et absorption de phonons de différentes énergies.



**Figure II-3 :** Nombres moyens d'interactions de type g par seconde: émission et absorption de phonons de différentes énergies.



**Figure II-4 :** Nombre moyen de collisions intravallées par seconde.

Pour les collisions inter et intra-vallées, on peut considérer que tous les états correspondants à la valeur de l'énergie finale de l'électron sont équiprobables. Cette caractéristique est utilisée lors des simulations Monte Carlo pour déterminer le vecteur d'onde après interaction [6]. Ce point est également utilisé avantageusement pour calculer la mobilité dans le formalisme de dérive-diffusion étendu.

### II-7.2.3 Collisions sur impuretés ionisées

Le modèle de collisions sur impuretés ionisées se révèle être très important, car ce phénomène a une influence non négligeable sur la mobilité des électrons [57] [62], grandeur de premier intérêt pour le couplage entre Monte Carlo et dérive-diffusion. La formulation de la probabilité de ce type de chocs peut-être étudiée à partir des modèles de Conwell-Weisskopf [44], de Brooks-Herring [44], et de la correction proposée par Ridley [60].

Selon l'approche de Brooks-Herring, le potentiel d'interaction correspondant à un ion est de la forme

$$V(\mathbf{r}) = \frac{Zq^2}{\epsilon_{Si}r} \exp(-\beta_s r) \quad (\text{II-42})$$

$\epsilon_{Si}$  est la constante diélectrique du silicium,  $Z$  le nombre de charges de l'ion, et  $\beta_s$  un facteur d'écrantage, qui fait que le potentiel décroît plus vite que l'inverse de la distance  $r$ .

Selon l'approche de Conwell-Weisskopf, un simple potentiel coulombien est utilisé, mais tronqué à une distance  $b$  du centre ionisé. Le paramètre d'impact étant au maximum  $b$ , l'angle minimum de déviation  $\theta_m$  est donné par la formule de Rutherford [52].

$$\tan\left(\frac{\theta_m}{2}\right) = \frac{q^2}{4\pi\epsilon_{Si}2b\mathcal{E}(1 + \alpha\mathcal{E})} \quad (\text{II-43})$$

Il est possible d'unifier les modèles pour mener les calculs en gardant le potentiel écranté et tronqué. On retrouve les modèles de Brooks-Herring et Conwell-Weisskopf correspondants aux cas particuliers  $b \rightarrow \infty$  et  $\beta_s = 0$ , respectivement [44].

Le calcul rigoureux de la probabilité d'interaction associée à un tel potentiel est simplifié par l'utilisation de la transformation de Herring-Vogt, mais c'est seulement au prix d'une

approximation sur la forme des vallées que l'on retrouve une expression similaire à celle fournie par Jacoboni [44] pour le taux d'interaction.

$$\lambda_i = \frac{2^{3/2} \pi Z^2 N_I q^4 \gamma^{1/2}(\mathcal{E}) (1 + 2\alpha\mathcal{E})}{(4\pi\epsilon_{Si})^2 m_d^{1/2}} \times \left( \frac{1}{4(1 - \cos\theta_m) \gamma^2(\mathcal{E}) + 2\mathcal{E}_\beta \gamma(\mathcal{E})} - \frac{1}{8\gamma^2(\mathcal{E}) + 2\mathcal{E}_\beta \gamma(\mathcal{E})} \right) \quad (\text{II-44})$$

où  $\gamma(\mathcal{E}) = \mathcal{E}(1 + \alpha\mathcal{E})$ ,  $\mathcal{E}_\beta = \frac{\hbar^2 \beta_s^2}{2m_d}$  et  $N_I$  est la concentration en impuretés ionisées.

L'expression liée au modèle de Brooks-Herring en est tirée en supprimant l'écrêtage

$$\lambda_{BH} = \frac{2^{3/2} \pi Z^2 N_I q^4 \gamma^{1/2}(\mathcal{E}) (1 + 2\alpha\mathcal{E})}{(4\pi\epsilon_{Si})^2 m_d^{1/2}} \times \left( \frac{1}{\mathcal{E}_\beta^2 + 4\mathcal{E}_\beta \gamma(\mathcal{E})} \right) \quad (\text{II-45})$$

et celle du modèle de Conwell-Weisskopf en annulant l'écrêtage

$$\lambda_{CW} = \frac{2^{1/2} \pi Z^2 N_I b^2 \mathcal{E}^2 (1 + 2\alpha\mathcal{E})}{m_d^{1/2} \gamma^{3/2}(\mathcal{E})} \quad (\text{II-46})$$

Pour une comparaison des modèles de Brooks-Herring et Conwell-Weisskopf, on pourra se référer aux articles de Rode [58] et Chattopadhyay [59]. Outre l'approximation signalée, aboutissant à l'expression (II-44) du taux d'interaction commune aux deux modèles, on peut relever quelques points:

\* Les deux modèles négligent les interactions multiples qui peuvent intervenir à basse température (ce point n'est pas critique pour l'étude de dispositifs à température ambiante) et assimilent les déflexions sur impuretés ionisées à des chocs (durée d'interaction faible devant le temps moyen entre collisions).

A propos du modèle de Conwell-Weisskopf:

\* L'angle de déviation minimal est exprimé à l'aide de la formule de Rutherford, relation de mécanique classique. Cette relation est mise en défaut par l'utilisation de vallées non sphériques.

\* La détermination de la distance d'écrêtage  $b$  donne lieu à plusieurs estimations,

$b = \left(\frac{3}{4\pi N_I}\right)^{1/3}$  [6], ou  $b = \left(\frac{1}{2N_I}\right)^{1/3}$  [2] [43] selon la façon de calculer la distance moyenne entre impuretés.

\* Ce modèle néglige l'écrantage des centres ionisés par les autres électrons libres [58]

\* Si l'écrantage et l'écrêtage sont considérés simultanément, l'écrantage est surestimé[58].

En ce qui concerne le modèle de Brooks-Herring:

\* Il est particulièrement adapté à l'étude des problèmes de transport dans les semiconducteurs non dégénérés, où de plus la concentration en électrons est assez faible pour ne pas mettre en défaut l'approximation de Born. Cela se traduit par le fait que la longueur d'onde de de Broglie est beaucoup plus petite que la longueur d'écrantage ( $k^2 \gg \beta_s^2$ ), ce qui est notamment le cas pour des impuretés diluées.

\* Le modèle complet de Brooks-Herring considère l'écrantage des impuretés par d'autres, ce qui demande de connaître en détail toutes les concentrations. En fait il est préférable d'utiliser l'expression de Dingle [58] [59] pour le facteur d'écrantage, valable quand la concentration en donneurs neutres est négligeable devant la concentration en électrons  $n$ :

$$\beta_s^2 = \frac{q^2 n}{\epsilon_{Si} k_B T} \quad (\text{II-47})$$

\* Le modèle de Brooks-Herring pourrait être mis en défaut dans les zones de forte concentration (supérieure à  $10^{18} \text{cm}^{-3}$ ) où il surestime la mobilité.

\* Il faut remarquer également que dans l'expression du facteur d'écrantage (II-47), la température est la température du gaz d'électron. Le taux de collisions calculé en prenant la température du réseau cristallin est sous-estimée dans le cas d'un transport hors d'équilibre.

\* Il est possible de simplifier encore l'expression du taux d'interaction en utilisant l'approximation de Born:

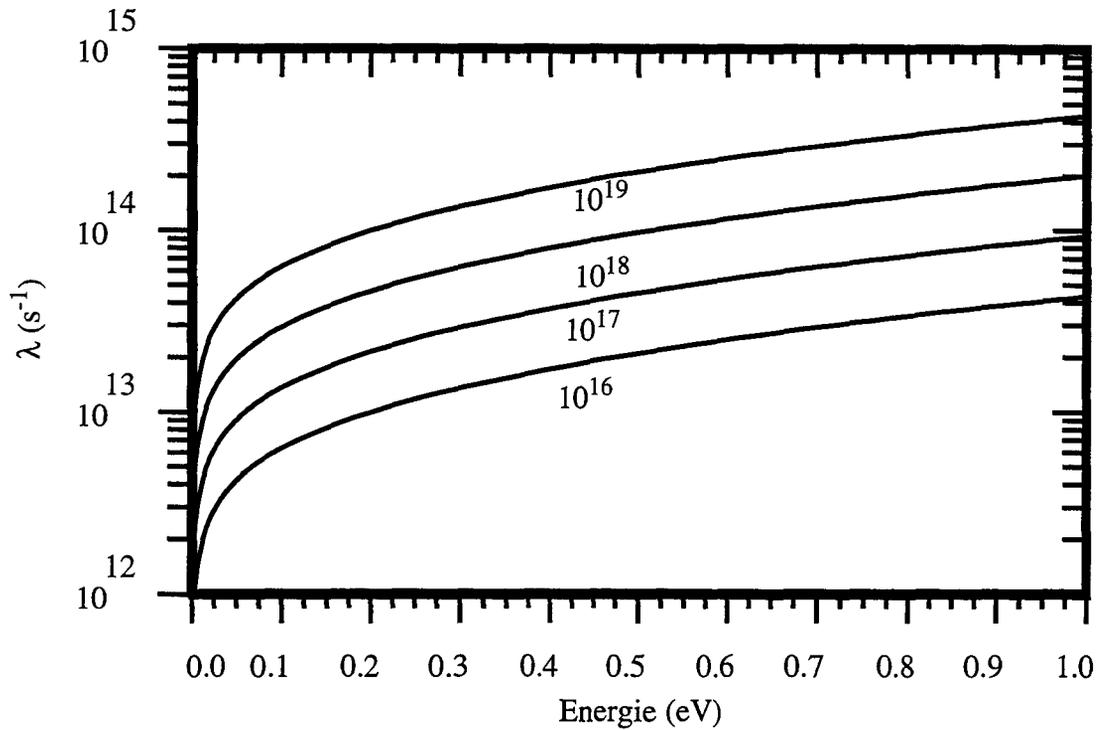
$$\lambda_{BH} = \frac{Z^2 N_I q^2 k_B T m_d^{1/2} (1 + 2\alpha \mathcal{E})}{2^{1/2} \pi \epsilon_{Si} \hbar^2 \gamma^{1/2} (\mathcal{E}) n} \quad (\text{II-48})$$

voire en considérant  $N_I = n$ , rendant ainsi la probabilité indépendante des concentrations [43].

$$\lambda_{BH} = \frac{Z^2 q^2 k_B T m_d^{1/2} (1 + 2\alpha \mathcal{E})}{2^{1/2} \pi \epsilon_{Si} \hbar^2 \gamma^{1/2} (\mathcal{E})} \quad (\text{II-49})$$

On risque alors de sous-estimer la mobilité quand la concentration en impuretés est inférieure à la concentration en électrons.

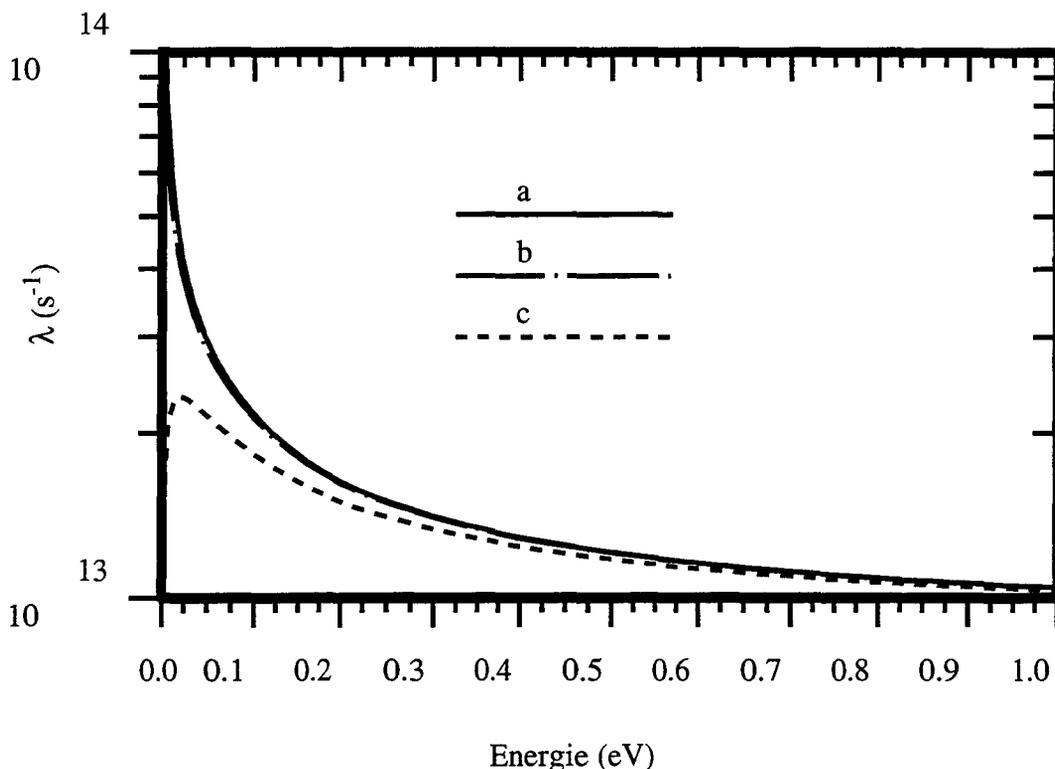
Les figures suivantes présentent les probabilités de collision sur les impuretés ionisées calculées selon les expressions de Conwell-Weisskopf et de Brooks-Herring.



**Figure II-5 :** Taux d'interaction selon l'expression de Conwell-Weisskopf (II-46) pour différentes valeurs de dopage depuis  $N_I = 10^{16} \text{ cm}^{-3}$  jusque  $N_I = 10^{19} \text{ cm}^{-3}$ .

Un inconvénient lié à l'utilisation de ce modèle lors des simulations Monte Carlo réside dans le fait que la probabilité calculée dans les zones fortement dopées est élevée. Ceci amène à consacrer un temps de calcul relativement important à ce type d'interaction. Par contre, le modèle de Brooks-Herring fournit des valeurs de probabilité élevées uniquement à très faible énergie, comme l'indique la figure II-6. Malgré de telles différences, les deux modèles sont pratiquement équivalents pour décrire la mobilité des électrons. En effet, les déflexions subies par les électrons sont moins importantes dans le cas du modèle de Conwell-Weisskopf que dans celui de Brooks-Herring.

La figure suivante permet de comparer les différentes expressions trouvées pour le modèle de Brooks-Herring, selon les approximations réalisées.



**Figure II-6** : La courbe a donne le taux d'interaction selon l'expression de Brooks-Herring (II-48) avec  $N_I = n$ . Les courbes b et c utilisent l'expression complète (II-45) avec  $N_I = n = 10^{18} \text{ cm}^{-3}$  et  $N_I = n = 10^{19} \text{ cm}^{-3}$  respectivement.

L'expression (II-48) ne conduit donc pas à des erreurs importantes, excepté à de très fortes concentrations, pour les énergies faibles. Cette différence n'a cependant pas beaucoup d'influence sur la mobilité des électrons en équilibre thermique, du fait de la redistribution quasi isotrope des moments due aux collisions. C'est donc cette expression qui est la mieux adaptée à la simulation de dispositifs.

La difficulté de prise en compte d'interactions multiples pose toujours problème, notamment à forte concentration. Une forme de réponse a été proposée par Ridley [60] [61] où un terme correctif permet de ne tenir compte que du centre ionisé le plus proche de l'électron. Il est alors possible de réconcilier les modèles de Brooks-Herring et Conwell-Weisskopf. On doit remarquer que cette approche est traitée dans un cadre classi-

que, utilisant la notion de section de capture, où la difficulté évoquée à propos de la distance moyenne entre impuretés subsiste. L'extension à une structure de bande non sphérique ne peut pas se faire de façon rigoureuse. De plus l'évaluation de la mobilité utilise une intégrale non calculable analytiquement [60], ce qui alourdit les simulations Monte Carlo. Finalement, la formulation de Brooks-Herring a été retenue pour cette étude, malgré ses limites à forte concentration, les autres modèles étant cependant disponibles pour comparaison.

#### II-7.2.4 Interactions entre porteurs

Un autre type d'interaction a donné lieu à des études: l'interaction électron-électron [50] [59]. Il faut noter en premier lieu que si ce type d'interaction est important dans une situation donnée, on atteint alors la limite de validité de l'approximation qui consiste à assimiler les électrons à un gaz dilué, ce qui remet en cause toute la modélisation adoptée ici. De plus, la prise en compte de ce type d'interaction n'a pas beaucoup d'influence sur les moyennes d'énergie et de vitesse des porteurs, les chocs étant élastiques [2], alors que sa mise en oeuvre lors de simulations est assez lourde et augmente notablement les temps de calcul [63] [64]. C'est pour ces raisons que l'interaction électron-électron n'a pas été retenue lors des simulations Monte Carlo liées à cette étude.



## REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE II

- [1] : L. Landau, E. Lifchitz, *Statistical Physics*, part 1, Ed. Pergamon Press, 1980
- [2] : P. Kireev, *La physique des semi-conducteurs*, Ed. Mir, 1975.
- [3] : E. M. Conwell, *High-field transport in semiconductors*, in *Solid State Physics*, supplément 9, Academic Press, 1967
- [4] : G. Baccarani, *Physics of submicron devices*, Proc. VLSI Process and Device Modeling, K.U.L. Summer Course, p1-23, 1983
- [5] : D. A. Antoniadis, J. E. Chung, *Physics and technology of ultra short channel MOS-FETs devices*, IEDM technical digest, p21, 1991
- [6] : C. Jacoboni, P. Lugli, *The Monte Carlo method for semiconductor device simulation*, Ed. S. Selberherr, Springer-Verlag
- [7] : E. M. Conwell, O. Vassel, *High-field transport in n-type GaAs*, Phys. Rev., vol.166, No.3, p797, 15 février 1968
- [8] : H. Lin, N. Goldsman, I. D. Mayergoyz, *An efficient deterministic solution of the space-dependent Boltzmann transport equation for silicon*, Solid-State Elec., vol.35, No.1, p33, 1992
- [9] : H. Lin, N. Goldsman, I. D. Mayergoyz, *Device modeling by deterministic self-consistent solution of Poisson and Boltzmann transport equations*, Solid-State Elec., vol.35, No.6, p769, 1992
- [10] : H. Lin, N. Goldsman, I. D. Mayergoyz, *A direct solution to the space-dependent Boltzmann transport equation in silicon*, in *Simulation of semiconductor devices and processes (SISDEP)*, vol.4, p195, Ed. by W. Fichtner, D. Aemmer, Hartung-Gorre, 1991
- [11] : K. A. Hennacy, N. Goldsman, *A generalized Legendre polynomial / sparse matrix approach for determining the distribution function in non-polar semiconductors*, Solid-State Elec., vol.36, No.6, p869, 1993

- [12] : A. Gnudi, D. Ventura, G. Bacarani, F. Odeh, *Two-dimensional MOSFET simulation by means of a multidimensional spherical harmonics expansion of the Boltzmann transport equation*, p917, in *Microelectronic Engineering 19*, Ed. Elsevier, 1992
- [13] : K. Bløtekjaer, *High-frequency conductivity, carrier waves and acoustic amplification in drifted semiconductor plasmas*, Ericsson Technics, vol. 2, p 125, octobre 1966
- [14] : G. Bacarani, M. R. Wordeman, *An investigation of steady-state velocity overshoot in silicon*, Solid-State Elec. vol.28, No.4, p407, 1985
- [15] : R. Thoma, A. Emunds, B. Meinerzhagen, H.-J. Peifer, W. L. Engl, *Hydrodynamic equations for semiconductors with nonparabolic band structure*, IEEE Trans. Elec. Dev., vol.38, No.6, p1343, juin 1991
- [16] : W. Quade, M. Rudan, E. Schöll, *Hydrodynamic simulation of Impact-ionization effects in p-n junctions*, IEEE Trans. on Computer-aided Design, vol.10, No.10, p1287, octobre 1991
- [17] : D. Chen, E. C. Kan, U. Ravaioli, C.-W. Shu, R. W. Dutton, *An improved energy transport model including nonparabolicity and non-maxwellian distribution effects*, IEEE Elec. Dev. Let., vol.13, No.1, janvier 1992
- [18] : E. Dubois, Thèse de doctorat en sciences des matériaux, Université des Sciences et Techniques de Lille Flandres-Artois, 1990
- [19] : Y.-K. Feng, A. Hintz, *Simulation of submicrometer GaAs Mesfet's using a full dynamic transport model*, IEEE Trans. Elec. Dev., vol.35, No.9, p1419, septembre 1988
- [20] : L. R. Logan, H. H. K. Tang, G. R. Srinivisan, *Analytic solutions to the Boltzmann equation for electron transport in silicon*, Phys. Rev. B, vol.43, No.8, p6581, mars 1991.
- [21] : J. M. Highman, K. Hess, C. G. Hwang, R. W. Dutton, *Coupled Monte Carlo-Drift Diffusion analysis of hot electron effects in MOSFET's*, IEEE Trans. Elec. Dev., vol.36, No.5, p930, mai 1989

- [22] : R. Thoma, A. Emunds, B. Meinerzhagen, H.-J. Peifer, W. L. Engl, *A generalized hydrodynamic model capable of incorporating Monte Carlo results*, IEDM 89, p139, 1989
- [23] : S.-C. Lee, T.-W. Tang, *Transport coefficients for a silicon hydrodynamic model extracted from inhomogeneous Monte-Carlo calculations*, Solid-State Elec. vol.35, No.4, p 561, 1992
- [24] : M. A. Stettler, M. A. Alam, M. S. Lundstrom, *A critical examination of the assumptions underlying macroscopic transport equations for silicon devices*, IEEE Trans. Elec. Dev., vol.40, No.4, p733, avril 1993
- [25] : S. Selberherr, *Analysis and simulation of semiconductor devices*, Ed. Springer Verlag, 1984.
- [26] : H. C. de Graaff, F. M. Klaassen, *Compact transistor modelling for circuit design*, Ed. by S. Selberherr, Spinger Verlag, 1990
- [27] : P. T. Landsberg, *D gradv or grad(Dv)*, J. Appl. Phys., vol.56, No.4, p1119, août 1984
- [28] : S. Bandyopadhyay, M. E. Klausmeier-Brown, C. M. Maziar, S. Datta, M. S. Lundstrom, *A rigorous technique to couple Monte Carlo and drift-diffusion models for computationnally efficient device simulation*, IEEE Trans. Elec. Dev., vol.34, No.2, p392, février 1987
- [29] : H. Kosina, S. Selberherr, *Efficient coupling of Monte Carlo and drift-diffusion method with applications to MOSFETs*, Extended Abstracts of the 22nd Conference on Solid State Devices and Materials, Sendai, 1990, p139
- [30] : H. Kosina, S. Selberherr, *Analysis of filter techniques for Monte Carlo device simulation*, in *Simulation of semiconductor devices and processes*, vol.4, p195, Ed. by W. Fichtner, D. Aemmer, Hartung-Gorre, 1991
- [31] : E. Dubois, P.-H. Bricout, R. Fauquembergue, D. Collard, *Monte Carlo/drift-diffusion coupling: non-equilibrium transport in deep submicron MOSFETs*, IEDM 91 technical digest, p519, 1991

- [32] : P.-H. Bricout, E. Dubois, R. Fauquembergue, *Transport simulation in submicron MOSFETs by coupling Monte Carlo and drift-diffusion*, NASECODE IX conf., p24, 1993
- [33] : P.-H. Bricout, E. Dubois, R. Fauquembergue, *Simulation of non-equilibrium transport in deep submicron MOSFETs*, ESSDERC'93, Editions Frontières, p609, 1993
- [34] : E. Constant, *Non steady state carrier transport in semiconductor in perspective with submicron devices*. Topics in Applied Physics : Hot electron transport in semiconductor, vol.58, Ed. by L. Reggiani, Springer-Verlag.
- [35] : W. Shockley, W. T. Read, *Statistics of the recombination of holes and electrons*, Phys. Rev. vol.87, No.5, p835-842, 1952
- [36] : R. N. Hall, *Electron-hole recombination in germanium*, Phys. Rev. vol.87, p387, 1952
- [37] : J. G. Fossum, D. S. Lee, *A physical model for the dependance of carrier lifetime on doping density in nondegenerate silicon*, Solid State Elec. vol.25, No.8, p741-747, 1982
- [38] : J. Dziewior, W. Schmid, *Auger coefficients for highly doped and highly excited silicon*, Appl. Phys. Lett. vol.31, p346-348, 1977
- [39] : E. Dubois, D. Collard, B. Baccus, D. Morel, *IMPACT 3.3 user's guide*, ISEN, janvier 1991
- [40] : K. Yamaguchi, *Field-dependent mobility model for two-dimensional numerical analysis of MOSFETs*, IEEE Trans. Elec. Dev., vol.26, p1068, 1979
- [41] : K. Yamaguchi, *A mobility model for carriers in the MOS inversion layer*, IEEE Trans. Elec. Dev., Vol.30, No.6, p658, juin 1983
- [42] : S. M. Sze, *Physics of semiconductor devices*, John Wiley & Sons, 1981
- [43] : Zimmermann, Thèse de doctorat d'état, Université des Sciences et Techniques de Lille Flandres-Artois, 1980
- [44] : C. Jacoboni, L. Reggiani, *The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials*, Rev. of Modern Phys., vol.55, No.3, p645, juillet 1983

- [45] : F. Venturi, E. Sangiorgi, R. Brunetti, C. Jacoboni, B. Riccò, *Monte Carlo simulation of electron heating in scaled deep submicron MOSFETs*, IEDM technical digest, p485, 1989
- [46] : F. Venturi, E. Sangiorgi, R. Brunetti, C. Jacoboni, B. Riccò, *Monte Carlo simulation of electron heating in scaled deep submicron MOSFETs*, IEEE CAD 10, No.10, p1276, octobre 1991
- [47] : E. Sangiorgi, M. R. Pinto, F. Venturi, W. Fichtner, *A hot-carrier analysis of submicrometer MOSFET's*, IEEE Elec. Dev. Let., Vol.9, No.1, p13, janvier 1988
- [48] : R. Brunetti, C. Jacoboni, F. Venturi, E. Sangiorgi, B. Riccò, *A many-band silicon model for hot-electron transport at high energies*, Solid-State Elec. vol.32, No.12, p1663, 1989
- [49] : F. Venturi, E. Sangiorgi, R. Brunetti, W. Quade, C. Jacoboni, B. Riccò, *Monte Carlo simulations of high energy electrons and holes in SI-n-MOSFETs*, Proc. NUPAD, 1990
- [50] : M. V. Fischetti, S. E. Laux, *Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects*, Phys. Rev. B, vol.38, No.14, p9721, novembre 1988
- [51] : M. V. Fischetti, S. E. Laux, D. J. Frank, *Monte Carlo analysis of semiconductor devices: the Damocles program*, IBM J. Res. Develop. vol.34, No.4, p466, juillet 1990
- [52] : K. Seeger, *Semiconductor Physics*, Solid State Sciences No. 40, Springer Verlag, 1982
- [53] : R. Brunetti, C. Jacoboni, F. Nava, L. Reggiani, *Diffusion coefficient of electrons in silicon*, J. Appl. Phys. vol.52, No.11, p 6713, novembre 1981
- [54] : C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, A. Alberigi-Quaranta, Phys. Rev. B, vol.12, p2265, 1975
- [55] : J. Bude, *Scattering mechanisms for semiconductor transport calculation*, in *Monte Carlo device simulation: full band and beyond*, Ed. by K. Hess Kluwer Academic Publishers, 1991

- [56] : R. Y. Chen, D.-S. Pan, *A Monte Carlo study of many-valley effects on the temporal electron transient transport in silicon*, J. Appl. Phys., vol.70, p4938, novembre 1991
- [57] : D. K. Ferry, *Extension of the Monte Carlo simulation in semiconductors to fast processes*, in *Monte Carlo device simulation: full band and beyond*, Ed. by K. Hess Kluwer Academic Publishers, 1991
- [58] : D. L. Rode, *Low-field electron transport*, in *Semiconductors and semimetals*, vol.10, p1, Academic Press, 1975
- [59] : D. Chattopadhyay, H. J. Queisser, *Electron scattering by ionized impurities in semiconductors*, Rev. of Modern Phys. vol.53, No.4, part I, p745, octobre 1981
- [60] : B. K. Ridley, *Reconciliation of the Conwell-Weisskopf and Brooks-Herring formulae for charged-impurity scattering in semiconductors: third body interference*, J. Phys. C, vol.10, p1589, 1977
- [61] : T. G. Van de Roer, F. P. Widdershoven, *Ionized impurity scattering in Monte Carlo calculations*, J. Appl. Phys., vol.59, p813, février 1986
- [62] : L. E. Kay, T.-W. Tang, *An improved ionized-impurity scattering model for Monte Carlo simulations*, J. Appl. Phys., vol.70, p1475, août 1991
- [63] : C. Moglestue, *A self-consistent Monte Carlo particle model to analyse semiconductor microcomponents of any geometry*, IEEE Tran. Computer-aided Design, vol.5, No.2, p 326, avril 1986
- [64] : A. Hasegawa, K. Miyatsuji, K. Taniguchi, C. Hamaguchi, *A new Monte Carlo simulation of hot electron transport with electron-electron scattering*, Solid State Elec., vol. 31, No.3/4, p547, 1988

## CHAPITRE III

### MISE EN OEUVRE DES METHODES DE SIMULATION

#### III-1 CHAINE DE SIMULATION

Un des principaux objectifs de la simulation des dispositifs microélectroniques est l'optimisation de la fabrication des composants, qui passe par une compréhension détaillée des phénomènes liés à leur fonctionnement. L'évaluation précise du comportement d'un transistor demande une connaissance toute aussi précise de sa géométrie et des profils de dopage. La simulation des étapes de fabrication d'un composant est appelée ici simulation de procédé, alors que le terme simulation de dispositif est réservé au calcul du comportement électrique. Pour réaliser une simulation du comportement électrique d'un dispositif, il est possible de le décrire de façon purement géométrique, en définissant les profils de dopage analytiquement. Cette solution a l'avantage d'être simple, mais ne permet pas de prendre en compte des variations de paramètres technologiques, ni des topologies complexes. Une autre solution est d'utiliser un simulateur de procédé en amont de la simulation de dispositif. Dans cette étude, les dispositifs sont ainsi simulés à partir de données technologiques. Les différents programmes utilisés font donc partie d'une chaîne de simulation présentée ci-dessous.

Les deux premiers simulateurs sont les programmes IMPACT ([1] [2] [3]). Dans le chapitre précédent ont été présentés les modèles utilisés d'une part pour résoudre les équations de dérive-diffusion étendues, et d'autre part pour réaliser des simulations Monte Carlo. Ce chapitre est consacré aux techniques de numérisation et de résolution correspondant à ces deux types de simulation, ainsi qu'à la méthode de couplage.

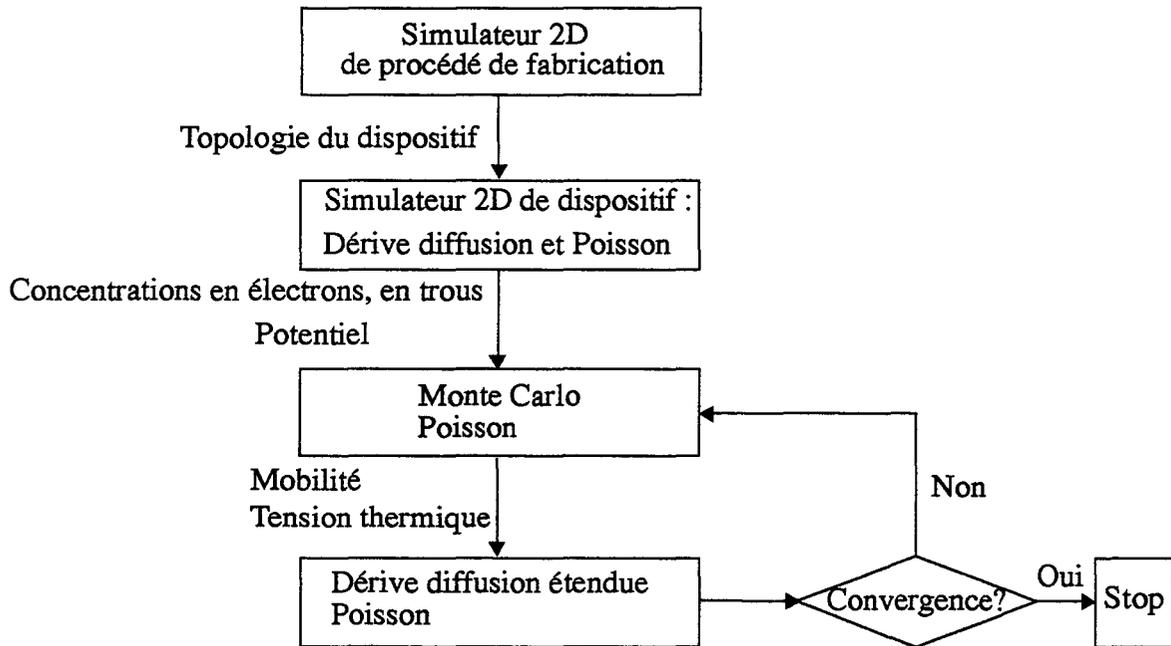


Figure III-1 : Programmes de simulation.

### III-2 DERIVE-DIFFUSION

La simulation bidimensionnelle d'un dispositif microélectronique à l'aide du formalisme de dérive-diffusion est maintenant établie de façon classique ([3] [4]) et se fait typiquement par une méthode d'éléments ou de différences finies. Ces méthodes de résolution nécessitent la discrétisation des équations de Poisson et de courant des porteurs:

\* équation de Poisson

$$\nabla_{\mathbf{r}} \left( \varepsilon \nabla_{\mathbf{r}} \Psi \right) = -q (p - n + N_D - N_A) \quad (\text{III-1})$$

\* Continuité du courant des électrons

$$\nabla_{\mathbf{r}} \mathbf{J}_n = -q (G - R) \quad (\text{III-2})$$

\* Continuité du courant des trous

$$\nabla_{\mathbf{r}} \mathbf{J}_p = q (G - R) \quad (\text{III-3})$$

Le choix effectué lors de l'élaboration du programme IMPACT qui est utilisé ici est de résoudre les équations régissant le comportement électrique des dispositifs par la méthode des différences finies. Il était donc tout à fait logique de s'appuyer sur ce travail pour y intégrer la résolution des équations du formalisme de dérive-diffusion étendu, pour lequel l'expression de la densité de courant des électrons est rappelée ici:

$$\mathbf{J}_n = -qn\mathbf{V} = nq\tilde{\mu}\mathbf{E} + nq\tilde{\mu}\nabla_{\mathbf{r}}\tilde{U} + q\tilde{\mu}\tilde{U}\nabla_{\mathbf{r}}n \quad (\text{III-4})$$

Dans la mesure où cette étude est restreinte aux transistors de type nMOS, les courants de trous sont traités de façon classique. Seule l'équation de courant des électrons utilise le formalisme étendu. Afin d'établir cette dernière équation, des hypothèses simplificatrices sont formulées sur les tenseurs  $\tilde{\mu}$  et  $\tilde{U}$ . En l'occurrence, ces tenseurs sont considérés comme diagonaux [5] [6]. Le bien-fondé de cette approximation sera discuté par la suite, et sa validité sera vérifiée par des simulations Monte Carlo pour le tenseur de tension thermique  $\tilde{U}$ . La densité de courant des électrons peut alors être exprimée sous la forme:

$$J_x = -qnV_x = -nq\mu_{xx}\frac{\partial\Psi}{\partial x} + nq\mu_{xx}\frac{\partial U_{xx}}{\partial x} + q\mu_{xx}U_{xx}\frac{\partial n}{\partial x} \quad (\text{III-5})$$

où  $x$  est la direction de l'axe considéré, horizontal ou vertical. Cette équation peut être discrétisée et résolue de façon classique, en utilisant un changement de variable:

$$\Psi_x^* = \Psi - U_{xx} \quad (\text{III-6})$$

$$J_x = -qnV_x = -nq\mu_{xx}\frac{\partial\Psi_x^*}{\partial x} + q\mu_{xx}U_{xx}\frac{\partial n}{\partial x} \quad (\text{III-7})$$

Cette dernière équation est formellement semblable à l'équation de dérive-diffusion usuelle. Le simulateur de dispositif utilisant la méthode des différences finies pour résoudre l'équation de Poisson et les équations de continuité de courant peut donc résoudre également les équations du formalisme étendu, dans la mesure où les termes de mobilité  $\mu_{xx}$  et de tension thermique  $U_{xx}$  sont connus. Ces termes peuvent être évalués par la méthode de Monte Carlo.

A partir de ces équations développées pour le courant d'électrons, il est aisé de retrouver les

équations décrivant les courants de trous, si l'on souhaite utiliser la même démarche pour simuler des phénomènes où le transport hors d'équilibre des trous est important (transistors PMOS).

### III-3 SIMULATION MONTE CARLO

Comme cela a déjà été mentionné, la méthode de Monte Carlo permet de simuler l'historique d'une particule ou d'un ensemble de particules. Utiliser un ensemble de particules revient à considérer un échantillon, le plus représentatif possible, de la population d'électrons étudiée. Cela se traduit sur le plan informatique par le stockage des informations concernant les porteurs de charge dans des tableaux assez volumineux. Pour étudier des phénomènes stationnaires, il est également possible de ne simuler qu'une seule particule [8] [9] [7], en utilisant la propriété d'ergodicité. L'objectif final étant ici de simuler un dispositif complet, c'est naturellement la simulation d'un ensemble de particules qui a été retenue. Ce choix permet de plus de simuler plus facilement l'évolution temporelle de la population d'électrons. Un deuxième choix concerne la détermination de la durée des vols libres. Une première technique consiste à tirer au sort cette durée en utilisant les probabilités d'interaction. Ce procédé s'accorde assez mal à la simulation d'un ensemble de particules, et on lui préfère généralement une discrétisation régulière du temps qui augmente quelque peu la quantité de calculs nécessaire, mais simplifie la gestion de l'ensemble des particules. A chaque pas de temps  $\Delta t$ , l'action du champ électrique sur les particules est calculée par le biais de trajectoires balistiques, puis les probabilités des différentes interactions sont évaluées, et un tirage au sort permet de déterminer si la particule subit une interaction ou continue sa trajectoire sans modification [8] [10].

L'organigramme des simulations Monte Carlo est présenté de façon simplifiée par la figure III-2. Le potentiel est régulièrement mis à jour au cours de la simulation par la résolution de l'équation de Poisson, de façon à prendre en compte les modifications du champ électrique introduites par le déplacement des électrons. Cette résolution demande l'évaluation des concentrations en électrons, compte tenu des nouvelles positions des particules. A la fin de la simulation Monte Carlo, certaines grandeurs peuvent être calculées en fonction de la position dans le dispositif, telles que la vitesse et l'énergie moyennes des électrons, la mobilité

et la tension thermique, ainsi que le courant le long du canal.

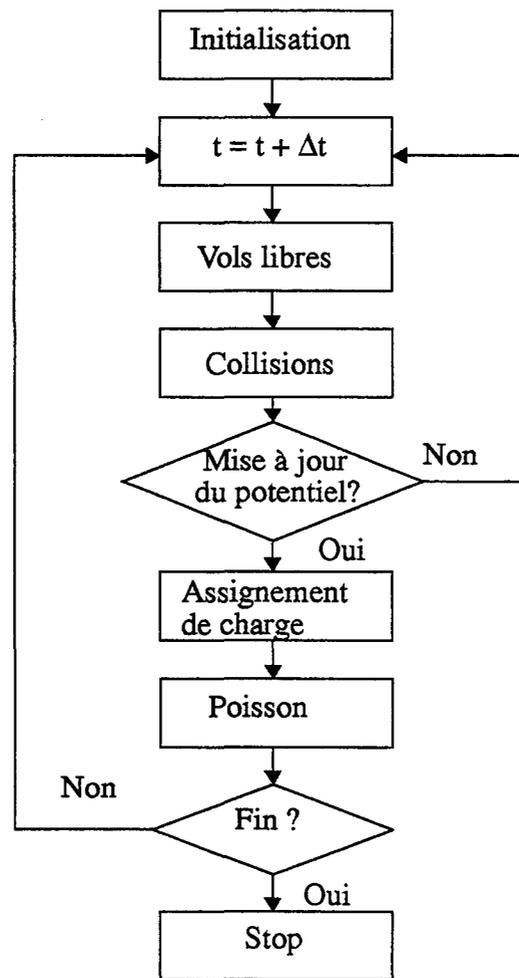


Figure III-2 : Organigramme simplifié du programme de simulation Monte Carlo

### III-3.1 Vol libre

A partir de l'état initial d'un électron, vecteur d'onde  $\mathbf{k}_i$  et position  $\mathbf{r}_i$ , il s'agit de trouver la position  $\mathbf{r}_f$  et le vecteur d'onde  $\mathbf{k}_f$  en fin de vol libre. La relation entre la vitesse et le vecteur d'onde peut être exprimée de la façon suivante, où  $x$  est une direction principale des vallées d'énergie:

$$v_x = \frac{\hbar k_x}{m_x (1 + 2\alpha \mathcal{E})} \quad (\text{III-8})$$

L'influence des forces sur le moment est décrite par la relation:

$$\frac{d\hbar\mathbf{k}}{dt} = \mathbf{F}_E = -q\mathbf{E} \quad (\text{III-9})$$

En considérant le champ électrique constant pendant l'intervalle de temps  $\Delta t$ , le vecteur d'onde à la fin du vol libre est donné par l'équation:

$$\mathbf{k}_f = \mathbf{k}_i - \frac{q\mathbf{E}}{\hbar}\Delta t \quad (\text{III-10})$$

La relation (III-8) permet d'évaluer les vitesses initiale  $\mathbf{v}_i$  et finale  $\mathbf{v}_f$ . La trajectoire parabolique de la particule lors du vol libre est alors approximée comme suit:

$$\mathbf{r}(t) = \mathbf{r}_i + \mathbf{v}_i \cdot t + \frac{1}{2} \frac{(\mathbf{v}_f - \mathbf{v}_i)}{\Delta t} t^2 \quad (\text{III-11})$$

et la position finale est donc:

$$\mathbf{r}_f = \mathbf{r}_i + \frac{1}{2} (\mathbf{v}_i + \mathbf{v}_f) \Delta t \quad (\text{III-12})$$

### III-3.2 Calcul des probabilités

Pour chaque particule, il est nécessaire de déterminer les probabilités des différents types d'interactions. Ces probabilités sont considérées comme indépendantes du champ électrique, comme cela a été remarqué au chapitre précédent. De plus, si l'on excepte les collisions sur impuretés ionisées, elles sont également indépendantes des dopages et peuvent donc être calculées une fois pour toutes, en fonction uniquement de l'énergie des particules (II-38) (II-40) (II-41). La probabilité d'interaction avec les impuretés ionisées, si elle ne peut pas être calculée complètement, peut être préparée: il est possible de calculer une fois pour toutes le terme qui dépend de l'énergie (II-49). Pour économiser le temps de calcul, un système d'index et d'interpolation a été utilisé, permettant de tabuler les probabilités, tout en limitant la taille mémoire nécessaire au stockage de ces informations.

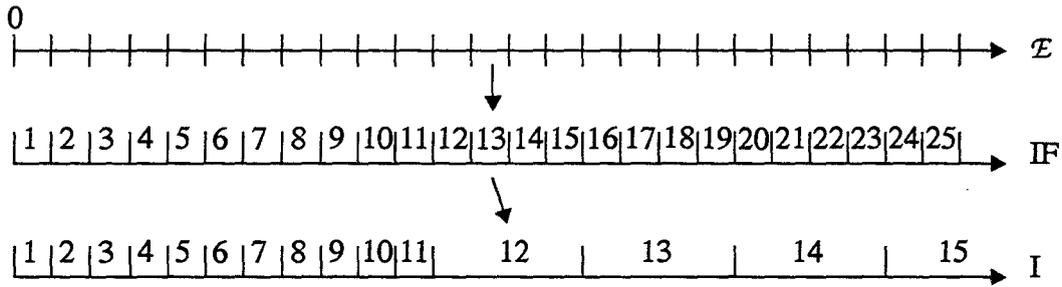


Figure III-3 : Présentation schématique du système d'index utilisé.

Comme la figure III-3 l'indique, la plage d'énergie utilisée est découpée en petits intervalles égaux, repérés par un index IF. La valeur de l'énergie  $\mathcal{E}$  permet de trouver directement cet index IF auquel est associé un second index I. Les taux de collision  $\lambda_i$ , associés à l'énergie  $\mathcal{E}$  sont alors interpolés à partir des valeurs tabulées d'index I et I+1. Les valeurs de l'index I ne sont pas réparties uniformément dans l'intervalle d'énergie utilisé, ce qui autorise une tabulation plus fine pour les valeurs d'énergie faibles, où les taux d'interaction varient plus fortement.

A partir du temps moyen entre deux interactions de type i,

$$\tau_i = \frac{1}{\lambda_i} \quad (\text{III-13})$$

la probabilité qu'au moins une interaction de ce type ait lieu pendant une durée  $\Delta t$  est évaluée à l'aide de la relation suivante:

$$P_i(\Delta t) = 1 - \exp\left(-\frac{\Delta t}{\tau_i}\right) = 1 - \exp(-\lambda_i \Delta t) \quad (\text{III-14})$$

En considérant un intervalle de temps  $\Delta t$  petit devant  $\tau_i$ , il est possible de donner une valeur approchée de cette probabilité:

$$P_i(\Delta t) \approx \frac{\Delta t}{\tau_i} = \lambda_i \Delta t \quad (\text{III-15})$$

Une fois déterminées les différentes probabilités d'interaction  $\lambda_i \Delta t$ , le tirage au sort d'un nombre pseudo-aléatoire  $r$  compris entre 0 et 1 permet de choisir l'événement subi par la particule:

collision de type 1 si  $r$  est compris entre 0 et  $\lambda_1 \Delta t$ ,

collision de type  $n$  si  $r$  est compris entre  $\sum_1^{n-1} \lambda_i \Delta t$  et  $\sum_1^n \lambda_i \Delta t$ ,

pas de collision si  $r$  est supérieur à la somme cumulée des  $\lambda_i \Delta t$ .

Pour limiter le temps passé à trouver l'événement à traiter, il est préférable de tester en premier les cas les plus probables: "pas d'interaction", collision sur impureté ionisée...

Le nombre  $r$  tiré au sort est en fait pseudo-aléatoire: la fonction qui le génère utilise la fonction modulo du calculateur. La suite de nombres fournis est donc parfaitement déterministe, ce qui facilite grandement la programmation et la mise au point du programme. Les nombres obtenus sont uniformément répartis entre 0 et 1, comme le seraient les valeurs prises par une variable aléatoire de densité constante.

### III-3.3 Traitement des interactions: choix du vecteur d'onde après collision

La grandeur effectivement tirée au sort est le vecteur  $\mathbf{k}^*$  obtenu à partir de  $\mathbf{k}$  par la transformation de Herring-Vogt. Si toutes les valeurs de  $\mathbf{k}$  correspondant à l'énergie de la particule  $\mathcal{E}$  après l'interaction sont équiprobables, il est en effet plus facile de choisir  $\mathbf{k}^*$  sur une sphère que  $\mathbf{k}$  sur une ellipse. En travaillant en coordonnées sphériques, cela revient à choisir  $\cos \theta$  entre -1 et 1 et  $\varphi$  entre 0 et  $2\pi$ . C'est ce qui est réalisé pour les interactions mettant en jeu les phonons acoustiques ainsi que les interactions de types f et g.

En ce qui concerne les collisions sur impuretés ionisées, toutes les valeurs de  $\mathbf{k}^*$  ne sont pas équiprobables après l'interaction. En travaillant en coordonnées sphériques avec comme direction de référence celle de  $\mathbf{k}^*$  avant le choc, on peut déterminer la fonction de répartition de  $\cos \theta$  de façon approchée. En tirant au sort un nombre  $r$  entre 0 et 1, on lui fait correspondre  $\cos \theta$ :

pour le modèle de Brooks-Herring;

$$\cos \theta = 1 - \frac{2r}{1 + \frac{4\gamma(\mathcal{E})}{\mathcal{E}_\beta} (1-r)} \quad (\text{III-16})$$

pour le modèle de Conwell-Weisskopf, où l'angle minimum de déviation  $\theta_m$  est donné par la formule de Rutherford (II-43);

$$\cos \theta = \frac{2 \cos \theta_m - (1 + \cos \theta_m) r}{2 - (1 + \cos \theta_m) r} \quad (\text{III-17})$$

et pour le modèle de Ridley:

$$\cos \theta = 1 + \frac{\mathcal{E}_\beta}{2\gamma(\mathcal{E})} + \frac{A \frac{\mathcal{E}_\beta}{2\gamma(\mathcal{E})}}{-A + \gamma(\mathcal{E}) \mathcal{E}_\beta \ln \left( r + (1-r) \exp \left( \frac{4A}{\mathcal{E}_\beta (\mathcal{E}_\beta + 4\gamma(\mathcal{E}))} \right) \right)} \quad (\text{III-18})$$

avec

$$A = \frac{Z^2 N_I q^4 b}{8\pi \epsilon_{Si}^2} \quad (\text{III-19})$$

expression qui se simplifie en se réduisant à l'expression de Brooks-Herring quand le terme de l'exponentielle est petit devant 1.

Le vecteur d'onde après la collision est trouvé en procédant à la transformation de Herring-Vogt inverse, après tirage d'un angle  $\varphi$  entre 0 et  $2\pi$ .

### III-3.4 Choix de la valeur de $\Delta t$

Le choix du pas de temps de la simulation Monte Carlo doit obéir à plusieurs impératifs. Le pas de temps doit d'abord permettre d'évaluer de façon suffisamment précise les trajectoires des particules: le champ électrique doit effectivement pouvoir être considéré comme constant pendant l'intervalle de temps  $\Delta t$ , la distance parcourue pendant un vol libre doit être faible. Le pas de temps doit de plus être petit devant les temps moyens entre interactions  $\tau_i$  pour que l'équation (III-15) soit valide. Enfin  $\Delta t$  doit être suffisamment grand pour que les interactions puissent être considérées comme instantanées et assimilées à des collisions. La valeur retenue est ici  $\Delta t = 2 \cdot 10^{-15}$  s. Cette valeur est au moins 10 à 100 fois plus petite que les temps entre interactions. Pour une particule de vitesse voisine de  $10^7$  cm/s la longueur d'un vol libre est de 0.2 nm, ce qui est satisfaisant pour l'approximation réalisée sur le champ électrique. Par contre, la valeur de  $\Delta t$  n'est pas grande devant les durées d'interaction. On peut cependant obtenir une bonne description de l'évolution de l'ensemble des particules, mais sans décrire finement les trajectoires de chacune d'elles. Le choix de cet intervalle de temps permet de simuler des phénomènes transitoires se déroulant typiquement sur une échelle de temps de l'ordre de la picoseconde en un nombre raisonnable d'étapes: 500 à 1000 pas de temps.

### III-3.5 Phénomène de survitesse

Les éléments précédents sont suffisants pour simuler le comportement électrique du silicium uniformément dopé. Ce sont de telles simulations qui ont permis de valider les paramètres des modèles utilisés pour décrire les interactions subies par les électrons à partir de données expérimentales [11] [7]. Ces simulations peuvent mettre en évidence le phénomène de survitesse: les électrons soumis à un fort champ électrique peuvent dépasser la vitesse de saturation avant de revenir en équilibre avec le réseau cristallin. C'est ce qui peut se produire dans le canal d'un transistor nMOS suffisamment court, quand la polarisation crée de forts champs électriques. La figure III-4 présente l'évolution de la vitesse moyenne des électrons soumis à l'instant 0 à un champ électrique constant.

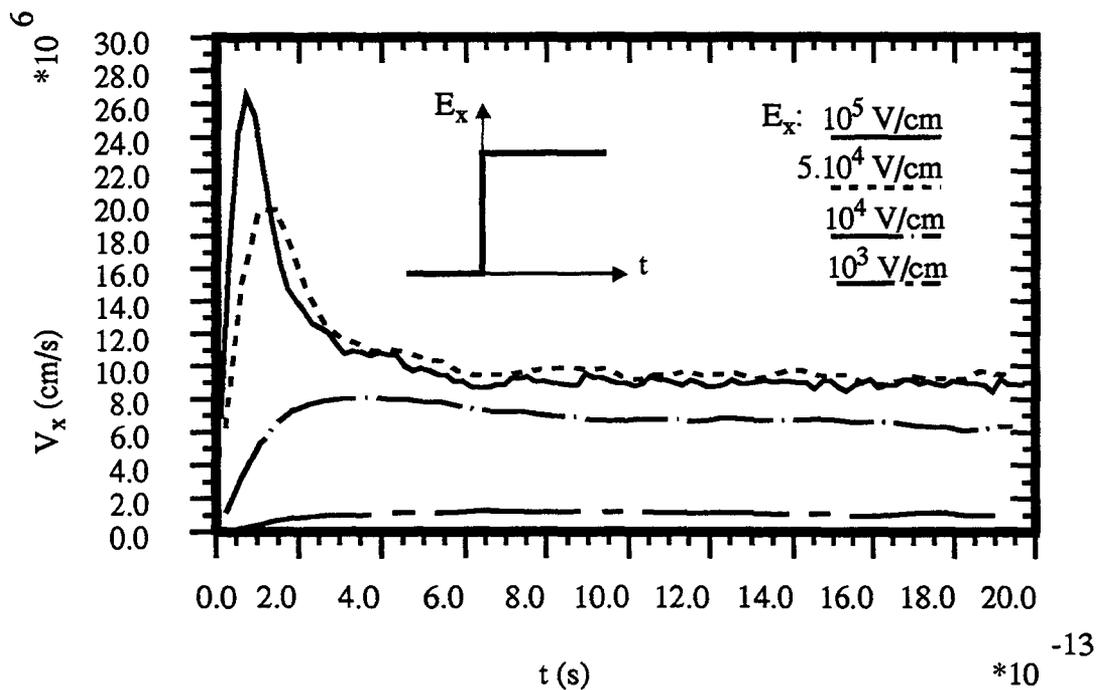
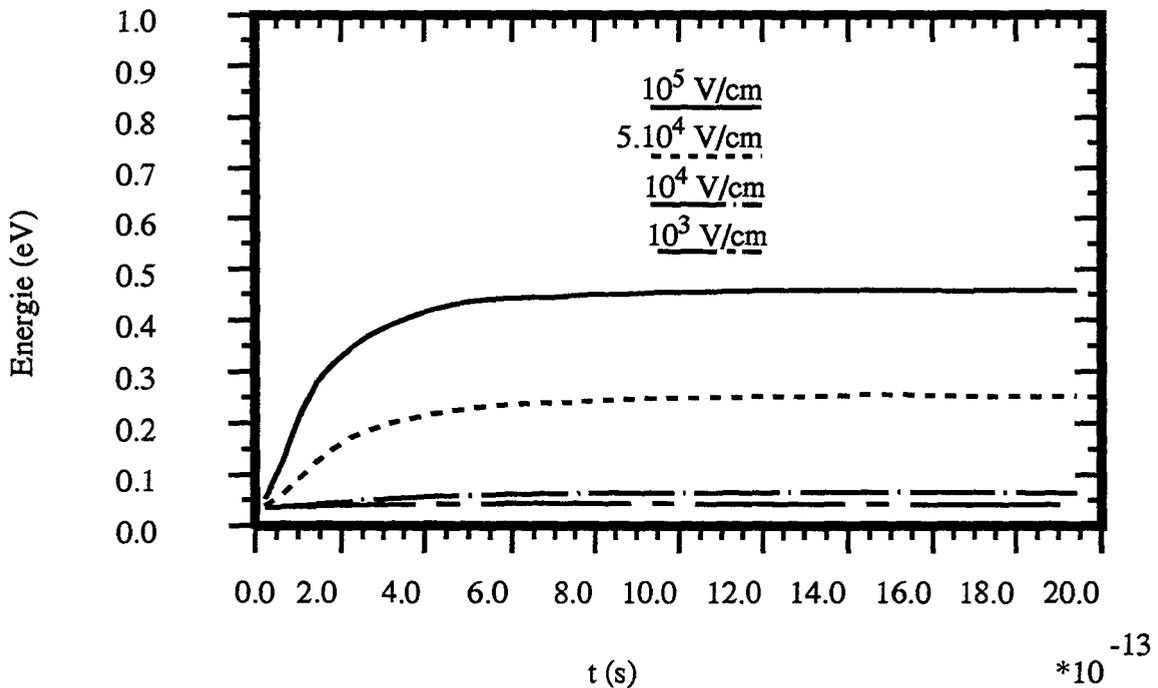


Figure III-4 : Evolution de la vitesse moyenne des électrons soumis à un champ électrique dans la direction  $x <100>$ . Les électrons sont initialement en équilibre avec le réseau. Le dopage est de  $10^{16}$  at/cm<sup>3</sup>.

Pour un champ de  $10^4$  V/cm, on observe déjà une augmentation de la vitesse au-delà de la vitesse moyenne des électrons en équilibre avec le réseau. Pour des valeurs de champ supé-

rieures, il y a clairement un pic de survitesse puis un retour à la vitesse de saturation, voisine de  $10^7$  cm/s, ce qui concorde bien avec les résultats présentés par Yoshii [36] et Bacarani [12]. Cet effet est principalement dû à la réorientation des mouvements des électrons dans la direction du champ électrique. L'énergie cinétique, quant à elle, croît lentement vers sa valeur finale, ainsi que l'indique la figure III-5: la relaxation en énergie est plus lente que la relaxation en vitesse.



**Figure III-5 :** Evolution de l'énergie moyenne des électrons soumis à un champ électrique dans la direction  $x \langle 100 \rangle$ . Les électrons sont initialement en équilibre avec le réseau. Le dopage est de  $10^{16}$  at/cm<sup>3</sup>.

Les valeurs finales de vitesse de dérive sont en excellent accord avec celles déduites des courbes présentées par Canali [11] et Fischetti [38], reportées dans le tableau ci-dessous.

**Table 1:**

$E_x$	$V_x$ [11] [38]
$10^3$ V/cm	$10^6$ cm/s
$10^4$ V/cm	$6 \cdot 10^6$ cm/s
$5 \cdot 10^4$ V/cm	$9 \cdot 10^6$ cm/s
$10^5$ V/cm	$9 \cdot 10^6$ cm/s

Les valeurs finales d'énergie sont par contre un peu inférieures, comme l'indique le tableau suivant.

**Table 2:**

$E_x$	Energie [11] [38]	Energie (figure III-5)
$10^3$ V/cm	$\sim 0.04$ eV	0.04 eV
$10^4$ V/cm	$\sim 0.06$ eV	0.06 eV
$5 \cdot 10^4$ V/cm	$\sim 0.3$ eV	0.25 eV
$10^5$ V/cm	$\sim 0.6$ eV	0.46 eV

### III-4 SIMULATION MONTE CARLO DE DISPOSITIF

La simulation d'un dispositif microélectronique par la méthode de Monte Carlo demande des éléments supplémentaires par rapport à la simulation du matériau semiconducteur pour décrire le dispositif, à savoir sa géométrie, ses frontières, les profils de dopage et des informations sur les différents matériaux utilisés. Le dispositif est décrit à l'aide d'un maillage rectangulaire non uniforme, utilisé pour résoudre les équations de dérive-diffusion et de Poisson. Ce maillage et les informations qui lui sont associées permettent de localiser les particules, de déterminer les forces et les interactions qu'elles subissent: champ électrique, collisions sur impuretés ionisées dont la probabilité dépend des dopages. D'autre part, la simulation Monte Carlo doit prendre en compte les conditions aux frontières: électrodes,

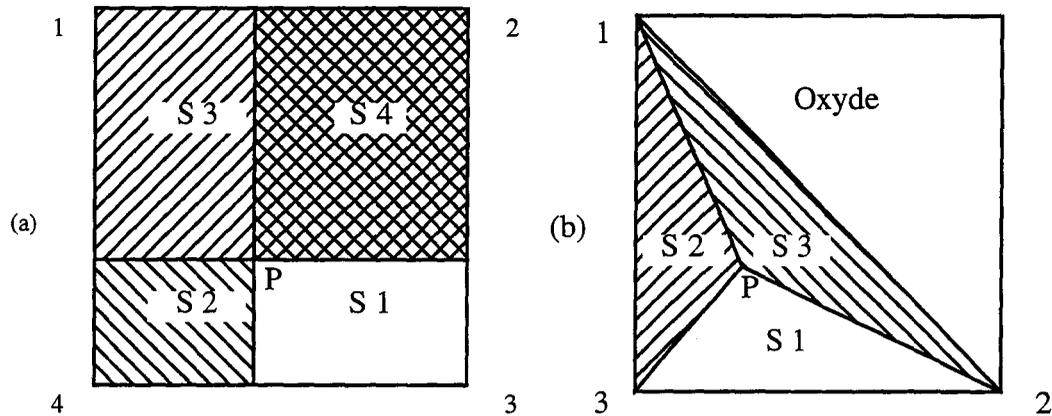
interface Si/SiO<sub>2</sub>, limites du domaine de simulation. Enfin, les grandeurs associées aux particules sont utilisées pour évaluer certaines grandeurs macroscopiques sur les noeuds du maillage, notamment les concentrations en électrons, les termes de mobilité, de tension thermique. Ces différents points font l'objet des paragraphes suivants. De plus, une stratégie dite de "multiplication de particules" sera discutée. Cette technique est destinée à économiser le temps de simulation en limitant le nombre de particules dans les zones fortement dopées.

### III-4.1 Localisation des particules

Après chaque vol libre, il est nécessaire de déterminer dans quelle maille chaque particule termine sa trajectoire. La solution retenue pour traiter cette question est l'utilisation d'un système d'index horizontaux et verticaux. Le maillage principal, non uniforme, est superposé à un maillage auxiliaire, uniforme et très fin, dont le pas typique est de quelques angströms. La position permet de trouver directement la maille fine dans laquelle se trouve la particule, et les index en x et y donnent les indices de la maille principale recherchée [13]. Le principe est identique à celui de la tabulation des probabilités.

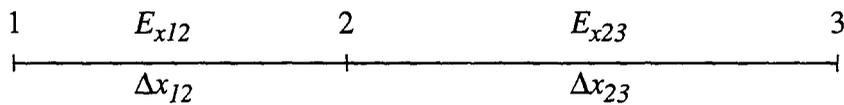
### III-4.2 Interpolation des grandeurs macroscopiques

La détermination de la probabilité d'interaction avec des impuretés ionisées rend nécessaire la connaissance des concentrations en dopants et en électrons à l'emplacement de chaque particule. Ces grandeurs sont estimées par interpolation à partir des concentrations aux noeuds du maillage. Cette interpolation est faite à partir des valeurs aux quatre noeuds des mailles rectangulaires, sauf pour les mailles comportant un segment oblique d'interface Si/SiO<sub>2</sub>, où seulement trois noeuds sont utilisés. Le poids associé au noeud n, qui est la fraction S<sub>n</sub> de la surface totale de la maille comme indiqué par la figure III-6, est calculé à partir de la position de la particule et des coordonnées des noeuds.



**Figure III-6** : Poids utilisés pour l'interpolation à l'emplacement de la particule P, dans les mailles rectangulaires (a), et triangulaires (b)

La résolution de l'équation de Poisson permet de connaître le potentiel aux noeuds du maillage, et donc le champ électrique entre les noeuds.



**Figure III-7** : Interpolation du champ électrique selon la direction x.

Les valeurs du champ électrique entre 1 et 2 d'une part, 2 et 3 d'autre part, calculés aux milieux des segments joignant les noeuds sont:

$$E_{x12} = -\frac{\Psi_2 - \Psi_1}{\Delta x_{12}} ; E_{x23} = -\frac{\Psi_3 - \Psi_2}{\Delta x_{23}} \quad (\text{III-20})$$

L'estimation du champ électrique à l'intérieur de la maille est rendue délicate par le phénomène de "force auto-induite" [17], artéfact dû à l'interpolation, notamment dans le cas d'un maillage non uniforme. Plusieurs méthodes ont été proposées [17], et l'on a retenu ici la relation suivante pour estimer la valeur du champ au noeud 2, qui permet de prendre en compte la différence de taille de mailles voisines:

$$E_2 = \frac{E_{x12}\Delta x_{23} + E_{x23}\Delta x_{12}}{\Delta x_{12} + \Delta x_{23}} \quad (\text{III-21})$$

En procédant de la même façon, il est possible de calculer à chaque noeud les composantes horizontales et verticales du champ électrique. Ceci permet ensuite d'évaluer les composantes du champ à l'emplacement d'une particule P avec le schéma d'interpolation rectangulaire de la figure III-6. Pour les mailles voisines de l'interface Si/SiO<sub>2</sub>, il est nécessaire de tenir compte des permittivités différentes du silicium et de l'oxyde lors de la résolution de l'équation de Poisson [3] et de l'interpolation, ce qui est fait en utilisant le déplacement électrique  $\mathbf{D} = \epsilon\mathbf{E}$ , où la permittivité  $\epsilon$  vaut  $\epsilon_{Si} = 11,7\epsilon_0$  dans le silicium et  $\epsilon_{Ox} = 3,9\epsilon_0$  dans l'oxyde.

### III-4.3 Evaluation de grandeurs macroscopiques

Les grandeurs associées aux particules, position, vecteur d'onde, vitesse et énergie sont utilisées pour calculer des valeurs moyennes associées aux noeuds du maillage telles que la concentration en électrons, la vitesse moyenne, les termes de tension thermique et de mobilité. Plusieurs méthodes ont été développées [8]: attribution des valeurs associées à une particule au noeud le plus proche, répartition entre les noeuds les plus proches, avec différentes pondérations possibles [8] [36] [37]. La solution retenue consiste en une répartition entre les noeuds voisins, utilisant les mêmes poids que pour l'interpolation des concentrations, comme indiqué à la figure III-6. Par exemple, pour le calcul de la concentration en électrons, la fraction  $S_n$  de la charge de la particule est affectée au noeud numéro  $n$ . Cette répartition se fait sur trois ou quatre noeuds selon la géométrie de la maille où se trouve la particule. On obtient ainsi à chaque noeud des grandeurs moyennes reflétant le comportement des particules.

Pour obtenir des résultats précis, il est possible de simuler de nombreuses particules, l'écart-type des valeurs calculées étant inversement proportionnel à la racine carrée du nombre de particules prises en compte. Une autre solution, lors de la simulation d'un régime stationnaire, est d'utiliser la propriété d'ergodicité et de réaliser des moyennes temporelles, ce qui limite les ressources informatiques à mettre en oeuvre, en évitant de gérer un nombre trop important de particules.

### III-4.4 Conditions aux frontières

Les frontières du domaines où évoluent les électrons se répartissent en plusieurs catégories: limites du domaine de simulation, interface Si/SiO<sub>2</sub>, et électrodes. Dans la mesure où un vol libre peut mener une particule hors du domaine de simulation, il est utile d'ajouter une "marge" autour du dispositif simulé. On utilise une distance  $d_{max}$  qu'aucune particule ne peut parcourir pendant le pas de temps  $\Delta t$  retenu. Cette distance est la largeur de la marge entourant le dispositif. Afin de traiter les conditions aux frontières avec efficacité, on associe à chaque maille du dispositif les frontières à prendre en compte. Ainsi, toutes les mailles situées à une distance inférieure à  $d_{max}$  d'une frontière sont repérées par un numéro qui indique les tests et le traitement qui doivent être effectués. Les particules se situant dans ces mailles sont testées, les autres non.

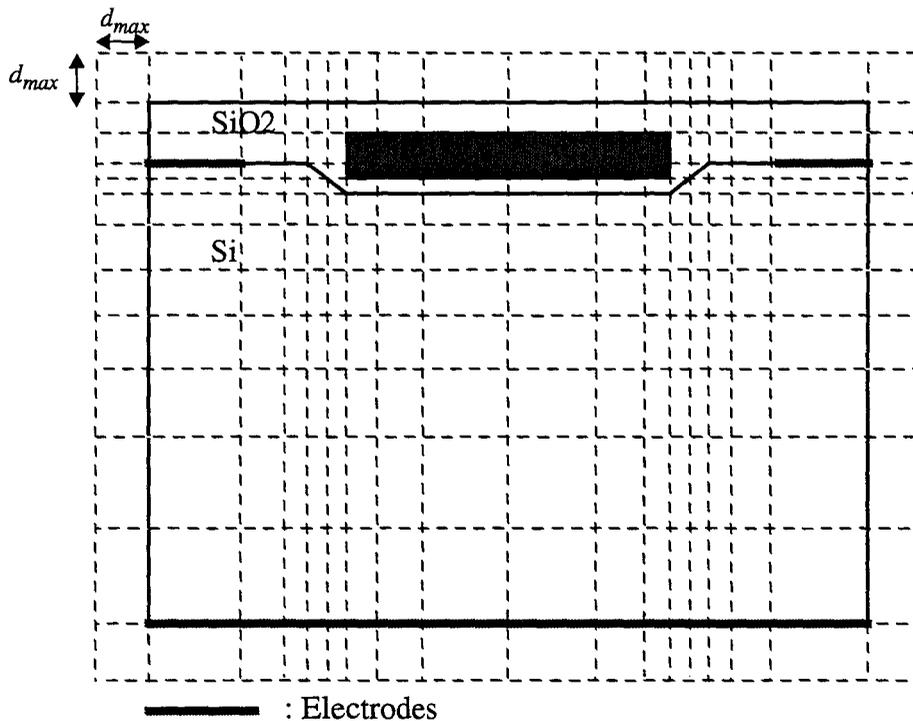


Figure III-8 : Maillage schématique d'un transistor MOS.

### Limites du domaine de simulation

Pour ces limites (frontières gauche et droite du dispositif), les simulations classiques de dérive-diffusion imposent une condition de miroir: pas de courant à travers la frontière, champ électrique normal à la frontière nul. Ces mêmes conditions se traduisent pour les simulations Monte Carlo par une simple réflexion des particules vers l'intérieur du dispositif.

### Interface Si/SiO<sub>2</sub>

Afin de rendre possible la simulation de transistors non plans, l'interface entre silicium et oxyde est modélisée par une succession de segments éventuellement diagonaux joignant deux noeuds consécutifs du maillage. Tester la collision d'une particule avec l'interface revient à rechercher une éventuelle intersection de sa trajectoire avec les segments proches. Afin de limiter le temps passé à traiter les particules proches de l'interface, qui sont nombreuses, les numéros des segments à tester sont associés aux mailles proches de l'interface (distance inférieure à  $d_{max}$ ). Une fois déterminé un point d'intersection, la particule subit une réflexion vers le silicium, ou une diffusion [14]. Dans ce dernier cas, la direction de la vitesse après collision contre l'interface est tirée au sort, et toutes les directions possibles dirigées vers le silicium sont équiprobables. Pour des segments d'interface obliques, qui ne sont pas selon une orientation  $\langle 100 \rangle$  du cristal, il est difficile de définir une réflexion, car les vecteurs  $v$  et  $k$  ne sont pas modifiés de la même façon. Une solution possible est de toujours faire subir aux particules une diffusion. Un autre procédé permet de conserver en partie le caractère spéculaire des collisions sur les segments obliques. Il consiste en une réflexion effectuée sur la direction du vecteur vitesse et en une correction de la norme du vecteur d'onde et de la vitesse, de façon à garantir la conservation de l'énergie. Cette façon de traiter les collisions à l'interface est en fait un moyen pratique de modéliser l'influence de la surface entre silicium et oxyde sur la mobilité des particules. Des collisions purement spéculaires n'ont aucune influence sur la mobilité, car seule la composante de la vitesse orthogonale à l'interface est modifiée lors de la réflexion: elle change simplement de signe. Les diffusions de particules à l'interface réduisent par contre la mobilité: la vitesse

des particules est réorientée de façon aléatoire, et la vitesse moyenne diminue. La répartition entre collisions spéculaires et diffusions est choisie de façon à refléter les variations de mobilité dans les dispositifs, les phénomènes physiques en cause étant bien plus complexes [14] [15] [16] [17] [18].

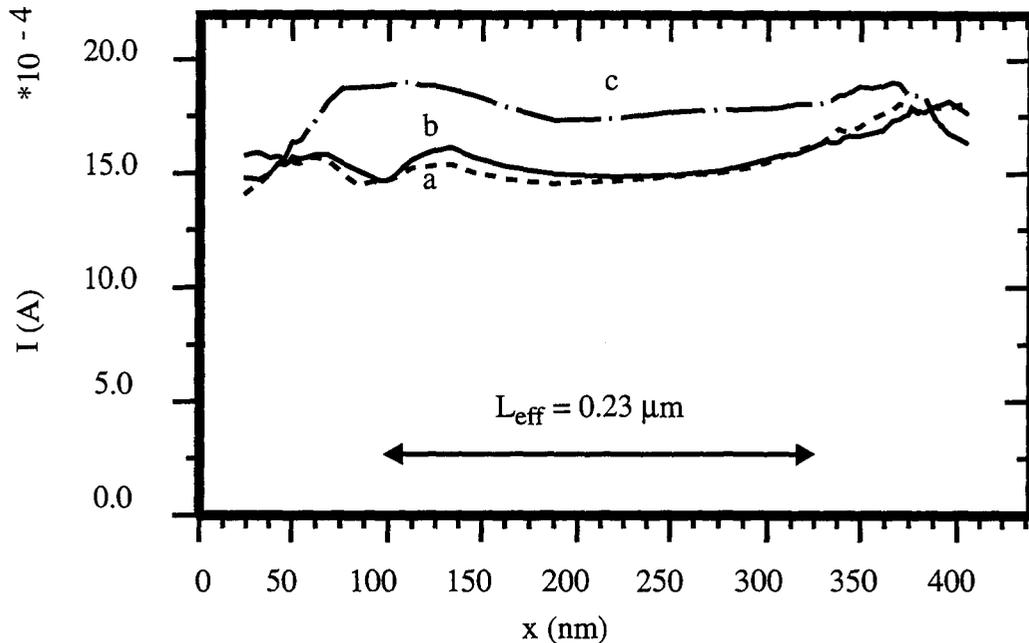
### Electrodes

Les électrodes sont considérées comme des contacts ohmiques dont il est possible de simuler le comportement par l'utilisation des conditions de Dirichlet [4] [19] [20]. L'obtention d'un contact ohmique se fait classiquement par l'utilisation d'un fort dopage à l'interface avec le contact métallique. La chute de potentiel à l'électrode est faible devant les différences de potentiel dans d'autres parties du dispositif, la région proche du contact est électriquement neutre et les porteurs y sont à l'équilibre thermique. Pour simuler des dispositifs en régime statique, le potentiel  $\Psi$  est considéré comme constant le long de l'électrode, et la concentration en électrons  $n$  est corrigée de façon à maintenir l'électroneutralité près du contact. Une façon de réaliser cette dernière condition est d'injecter des particules le long de l'électrode pour rétablir une charge totale nulle. L'énergie des particules injectées est prise à l'équilibre thermique,  $\mathcal{E} = \frac{3}{2}k_B T$  et le vecteur d'onde des particules injectées est tiré au sort, orienté vers l'intérieur du domaine silicium du dispositif [8]. Cela met en défaut certaines approches utilisant des fenêtres où le domaine de simulation est limité à la zone hors d'équilibre [23]. En effet, la prise en compte des conditions aux frontières impose de simuler au moins une partie des zones de source et drain à fort dopage, où l'électroneutralité est vérifiée [9] [21] [22] [24].

Lors de la simulation d'un dispositif MOS, un certain nombre de particules représentant des électrons quittent le domaine à travers l'électrode de drain. Si des particules sont injectées pour imposer l'électroneutralité sur toutes les électrodes, le nombre de particules à l'intérieur du dispositif peut évoluer au cours de la simulation, et il n'y a plus conservation du nombre de charges total [25] [26] [27] [28]. Trois stratégies de traitement des contacts ohmiques ont été testées.

- a) La première consiste à imposer l'électroneutralité dans les mailles d'électrodes de source et drain. En cas d'excédent de particules, certaines sont retirées des mailles, alors qu'en cas de déficit, de nouvelles particules sont injectées le long des électrodes.
- b) La deuxième stratégie garantit la conservation de la charge totale: l'électroneutralité est maintenue à la source comme dans le cas précédent, et le nombre de particules est corrigé côté drain de façon à garder le même nombre total de particules lors de la simulation.
- c) Le troisième traitement diffère du second uniquement par le fait que les particules sont injectées dans l'ensemble du volume des mailles de l'électrode de source, et pas seulement en surface, le long du contact.

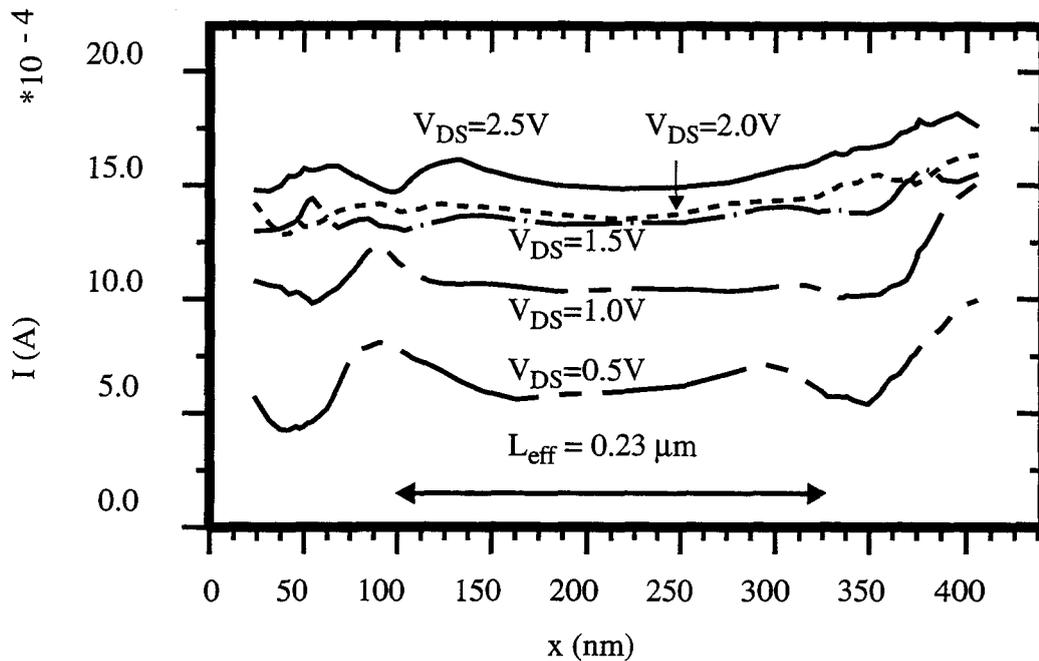
La structure utilisée pour le test des conditions imposées aux électrodes est un transistor nMOS, de longueur effective de canal  $L_{\text{eff}} = 0.23\mu\text{m}$ , d'épaisseur d'oxyde  $t_{\text{ox}} = 45\text{\AA}$ . Le dopage maximal des zones de source et drain est de  $10^{19}\text{cm}^{-3}$ . Les données technologiques liées à la fabrication de ce transistor seront reprises au chapitre suivant. Elles sont tirées de l'article de Sai-Halasz et al. [29], portant sur des dispositifs fabriqués par IBM. Les contacts de source et drain ont été pris verticaux pour les simulations, afin de pouvoir comparer les résultats avec ceux de Fischetti, Laux et Frank [30]. La largeur du transistor est  $W = 1\mu\text{m}$ , et les tensions de polarisation sont  $V_{\text{GS}} = 2.5\text{V}$  et  $V_{\text{DS}} = 2.5\text{V}$ . Les simulations Monte Carlo portent sur environ 24000 particules. Après une picoseconde de simulation à partir de conditions initiales d'énergie et de vitesse arbitraires, les résultats sont évalués sous forme de moyennes calculées sur un deuxième intervalle de temps de une picoseconde.



**Figure III-9** : Courant calculé le long du transistor pour les trois méthodes de traitement des électrodes: a- électroneutralité des mailles d'électrodes; charges injectées le long de l'électrode. b- conservation du nombre de particules; charges injectées le long de l'électrode. c- conservation du nombre de particules; charges injectées dans les mailles d'électrode.

La façon de simuler l'injection des électrons à partir de la source peut influencer de façon notable le courant dans le dispositif. L'injection de particules le long de l'électrode permet d'obtenir une meilleure conservation du courant le long du dispositif comme le montre la comparaison des courbes a et b d'une part, et c d'autre part. La variation du courant dans la région du drain peut être attribuée au fait que l'échelle de temps de la méthode de Monte Carlo ne permet pas une actualisation rapide des concentrations en porteurs, ce qui conduit à des phénomènes transitoires pour le champ électrique et le courant lors des simulations de dispositifs. Le fait d'imposer l'électroneutralité de la région de drain ou la conservation du nombre total de particules n'a par contre que peu d'influence sur le courant global. Le nombre de particules reste stable pour la simulation a: 24000 au départ, entre 23500 et 24000 au cours de la simulation.

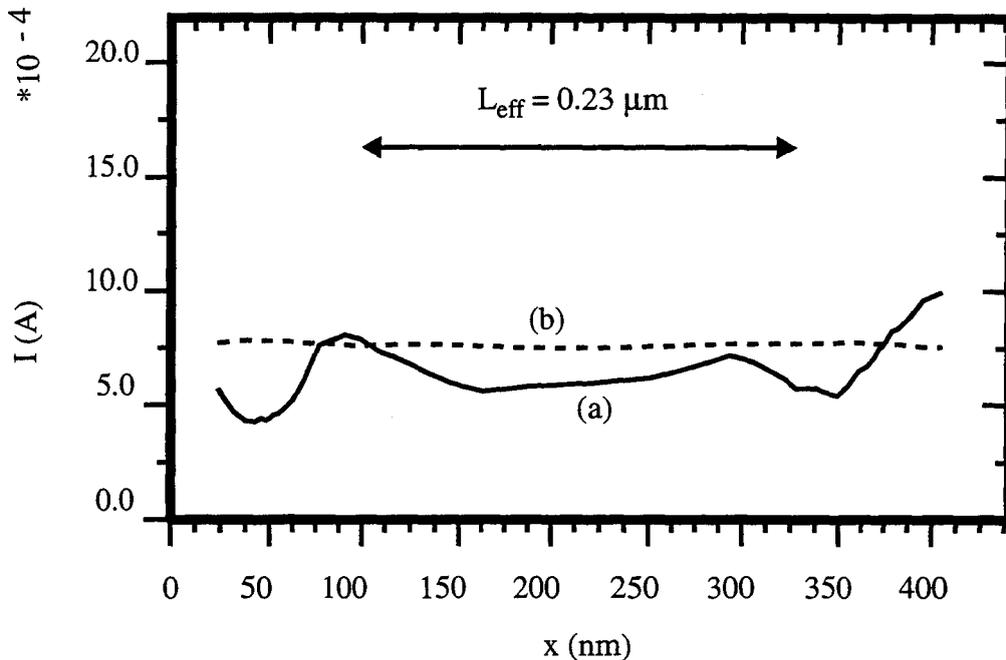
La figure III-10 présente le courant estimé le long du canal pour le même transistor, à différentes tensions de drain. Les fluctuations de concentrations en porteurs de charge dans les zones fortement dopées de source et de drain entraînent des variations de courant le long du dispositif, d'autant plus importantes que la polarisation est faible.



**Figure III-10** : Courant calculé le long du transistor pour différentes tensions de drain.

$V_{GS} = 2.5V$ . Pour les électrodes: conservation du nombre de particules; charges injectées le long de la source.

Les fluctuations observées le long du canal sont dues au fait que le système simulé ne parvient pas à une situation vraiment stable pendant la durée de la simulation, ici 2ps. Pour une simulation plus longue, par exemple 20 ps, où les moyennes sont calculées pendant les dix dernières picosecondes, elles disparaissent pratiquement, comme l'indique la figure III-11. L'ensemble des particules simulées a alors atteint un régime stationnaire à partir des conditions initiales arbitraires.



**Figure III-11** : Courant calculé le long du transistor pour  $V_{GS} = 2.5V$ ,  $V_{DS} = 0.5V$ .

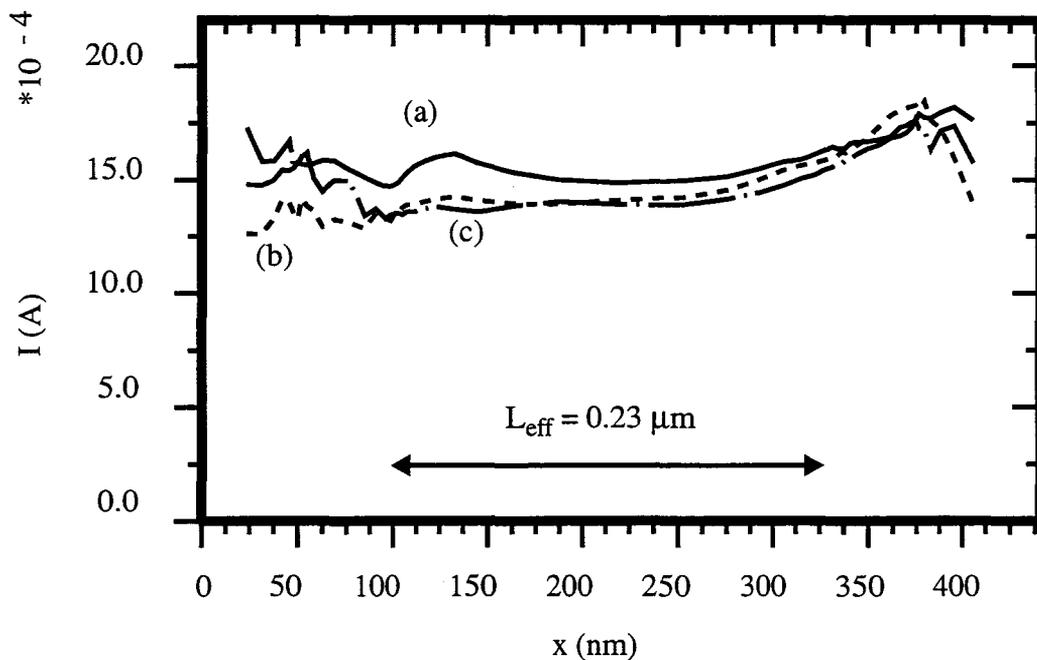
a: simulation de 2 ps, moyenne pendant 1ps; b: simulation de 20 ps, moyenne pendant 10 ps.

La méthode retenue pour traiter les électrodes consiste à garantir l'électroneutralité du contact de source en injectant des particules le long de l'électrode, d'énergie  $\mathcal{E} = \frac{3}{2}k_B T$ , et de corriger les mailles du drain de façon à maintenir constant le nombre total de particules dans le dispositif, ce qui se traduit par la conservation de la charge totale. Dans la mesure où les phénomènes de génération par impact ne sont pas pris en compte lors des simulations Monte Carlo, il n'y a pas de courant au contact du substrat. Cette électrode n'est prise en considération que lors de la résolution de l'équation de Poisson.

#### III-4.5 Gestion des particules: technique de multiplication

La simulation d'un dispositif par la méthode de Monte Carlo nécessite la prise en compte d'au moins une partie des zones fortement dopées, près des contacts ohmiques. Une part importante du temps de simulation est donc consacrée à traiter de nombreuses par-

ticules près des électrodes, où les électrons sont normalement en équilibre avec le réseau cristallin, le champ électrique étant faible. Il est possible de limiter l'étendue des zones ohmiques dans le domaine de simulation, ce qui est une première façon de réduire les temps de calcul. Une deuxième technique consiste à utiliser des particules de différents poids [9] [10] [33]. Une particule représente alors plus d'électrons dans les zones fortement dopées que dans les régions où le transport hors d'équilibre intervient. De façon plus générale, on privilégie certaines occurrences [31] et on multiplie le nombre de particules pour étudier les cas rares [9] [30] [32]: survitesse, haute énergie. Pour la simulation d'un dispositif MOS, cela revient par exemple à diviser les particules qui passent de la source au canal en plusieurs particules de poids inférieur, et à regrouper les particules qui entrent dans la zone de drain. Le canal, où peut dominer le transport hors d'équilibre est simulé par des particules, plus nombreuses, représentant chacune moins d'électrons.

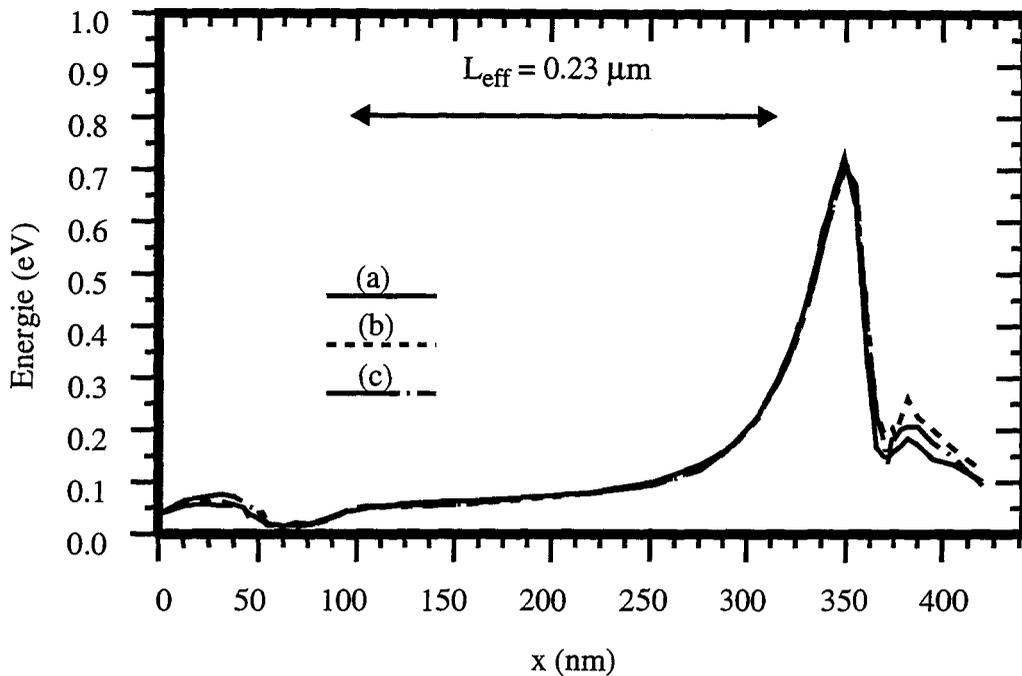


**Figure III-12** : Courant calculé le long du transistor pour  $V_{GS} = 2.5\text{V}$ ,  $V_{DS} = 2.5\text{V}$ .

- a: sans multiplication avec 24000 particules;
- b: avec multiplication et environ 24000 particules; minimum 20 par maille
- c: avec multiplication et environ 12000 particules; minimum 20 par maille.

Une condition nécessaire au bon déroulement des simulations est de garantir un nombre suffisant de particules dans les zones fortement dopées, afin de limiter les fluctuations dues au mouvement de particules de poids important. La solution retenue consiste à imposer dans les mailles des zones de source et drain, un nombre minimal de particules de même poids, qui conditionne leur éclatement ou leur regroupement. Pour les simulations dont les résultats sont présentés ici, les particules peuvent avoir un poids de 1, comme celles qui sont dans le canal, ou des poids 2, 4 ou 8 fois plus importants dans les zones de source et drain. Le nombre des particules de poids le plus élevé dans une maille doit être supérieur au nombre minimal, seuil qui a été fixé lors des simulations à 20. Si ce n'est pas le cas, elles sont séparées en particules de poids inférieur. Inversement, si elles sont assez nombreuses, elles sont regroupées en particules de poids plus important.

Même si le courant estimé le long du canal est influencé par l'utilisation de la technique de multiplication des particules, comme l'indique la figure III-12, l'énergie des particules le long de l'interface n'est que faiblement altérée, comme le montre la figure III-13. Il est normal que l'algorithme de multiplication ait peu d'influence sous l'interface, car les mailles étant plus fines que dans le reste du dispositif, les particules n'y sont pratiquement pas modifiées.



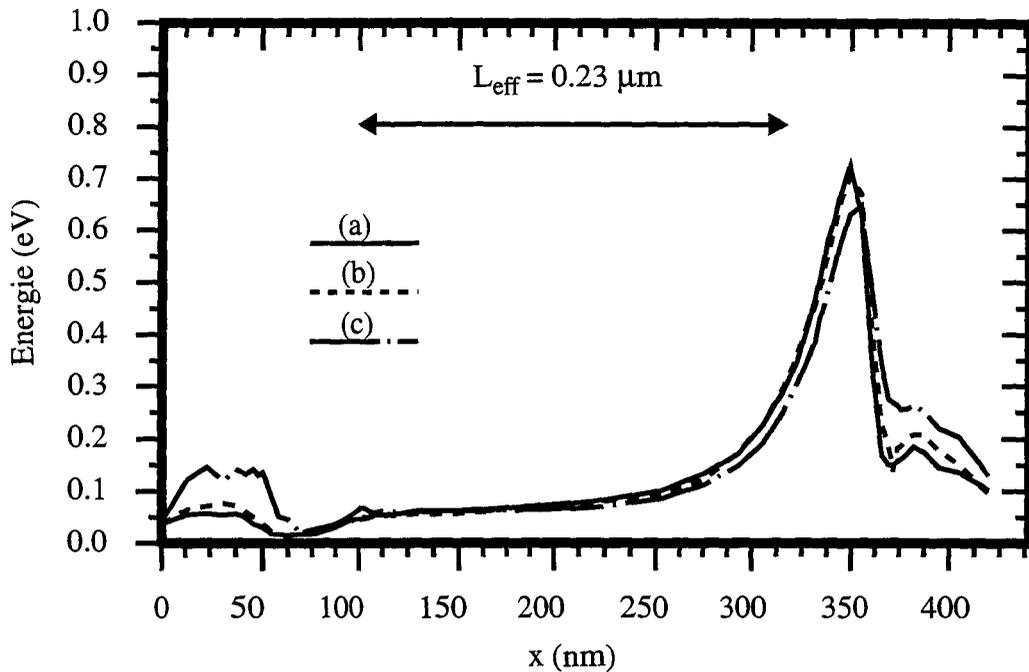
**Figure III-13** : Energie moyenne des électrons calculée sous l'interface, le long du transistor pour  $V_{\text{GS}} = 2.5\text{V}$ ,  $V_{\text{DS}} = 2.5\text{V}$ .

a: sans multiplication avec 24000 particules;

b: avec multiplication et environ 24000 particules; minimum 20 par maille

c: avec multiplication et environ 12000 particules; minimum 20 par maille.

Cette technique de multiplication doit être utilisée avec précaution. Si par exemple le nombre de particules par maille peut devenir petit, il peut en résulter des phénomènes artificiels, correspondant d'une part aux fluctuations introduites par le trop petit nombre de particules simulées, qui ne constituent plus un échantillon représentatif des électrons, et d'autre part aux erreurs introduites en supprimant des particules lors des regroupements. C'est ce qui est observé sur la figure III-14: l'énergie est surestimée dans les zones de source et drain, où l'algorithme de multiplication est utilisé, et légèrement sous-estimé dans le canal.



**Figure III-14** : Energie moyenne des électrons calculée sous l'interface, le long du transistor pour  $V_{GS} = 2.5V$ ,  $V_{DS} = 2.5V$ .

a: sans multiplication avec 24000 particules;

b: avec multiplication et environ 12000 particules; minimum 20 par maille.

c: avec multiplication et environ 12000 particules; minimum 5 par maille.

Si la technique de multiplication de particules est intéressante pour réduire les temps de simulation en consacrant moins de temps aux zones fortement dopées où les électrons sont en équilibre thermique avec le réseau, elle demande des précautions garantissant un échantillonnage suffisant pour décrire le système physique, sans introduire d'artéfacts liés à l'algorithme.

### III-5 COUPLAGE

La démarche de couplage utilise la complémentarité de la technique de Monte Carlo et de la méthode de dérive-diffusion, en combinant les avantages des deux approches. Les simulations Monte Carlo permettent d'évaluer dans les zones où le transport électroni-

que est hors d'équilibre, les paramètres nécessaires à la résolution des équations du formalisme de dérive diffusion étendu. Les valeurs à calculer sont les termes de tension thermique et de mobilité, comme indiqué au chapitre précédent.

### III-5.1 Evaluation des termes du tenseur thermique

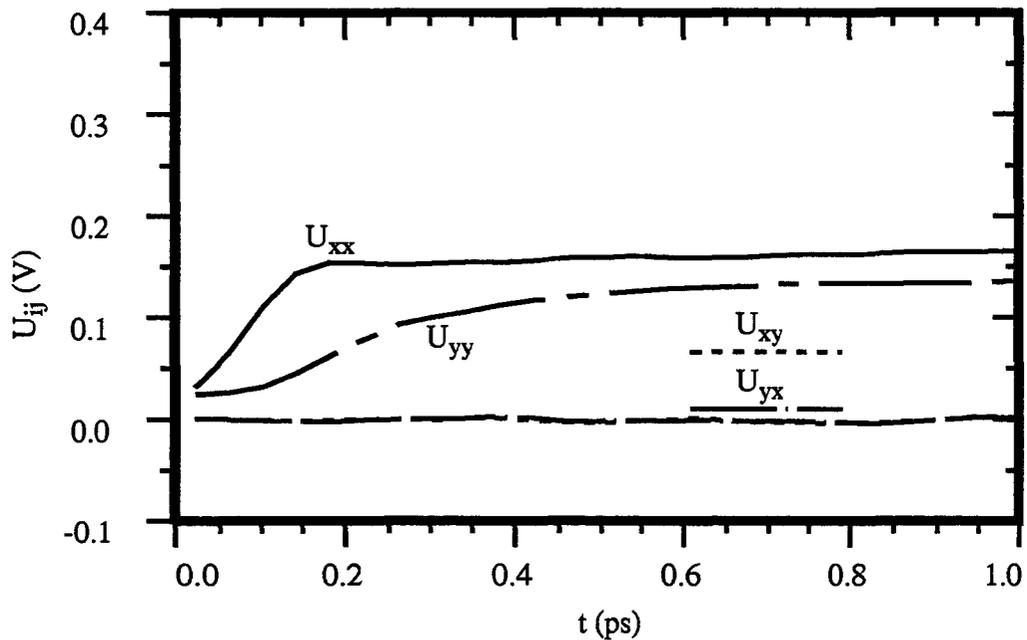
Le calcul des moments successifs de l'équation de transport de Boltzmann fait intervenir le tenseur thermique  $\tilde{U}$  dont les termes sont définis par l'équation suivante [34] [35]:

$$U_{ij} = \frac{1}{qn} \int f(\mathbf{r}, \mathbf{k}, t) \hbar k_i \cdot v_j d^3 k \quad (\text{III-22})$$

Le terme  $qU_{ij}$  est la valeur moyenne du produit de la composante  $i$  du moment par la composante  $j$  de la vitesse des électrons. Dans le cas d'électrons en équilibre avec le réseau cristallin, on retrouve des termes non diagonaux nuls, et pour les termes diagonaux, la relation d'Einstein est vérifiée:

$$U_{ii} = V_T = \frac{k_B T}{q} \quad (\text{III-23})$$

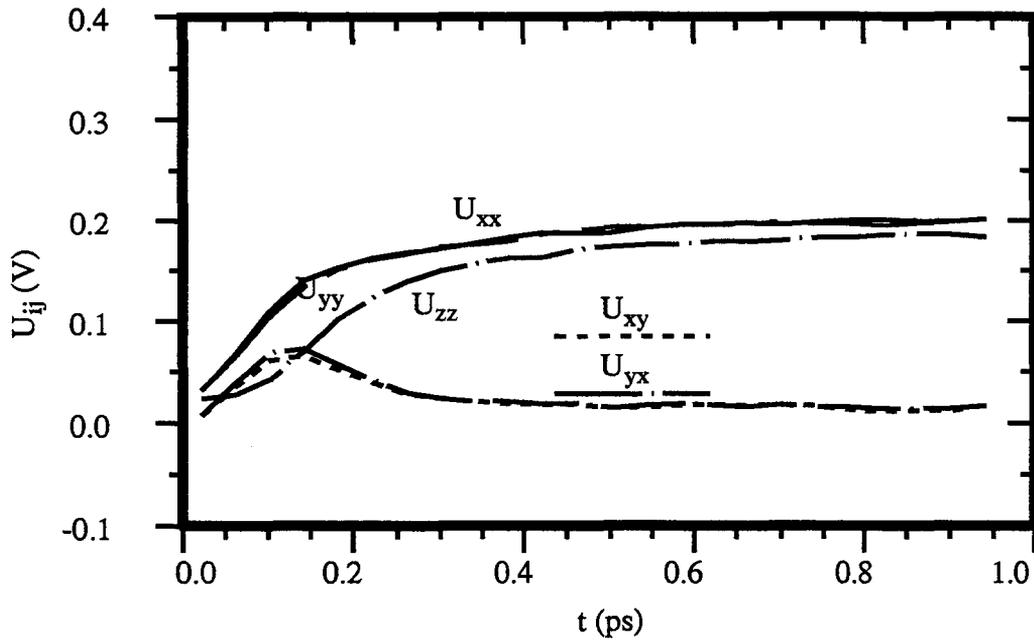
Les différents termes du tenseur  $\tilde{U}$  peuvent être évalués par des moyennes réalisées lors de simulations Monte Carlo. La simulation d'un barreau de silicium uniformément dopé, où les électrons sont soumis à un champ électrique constant, permet de vérifier si les termes non diagonaux sont négligeables. Les résultats d'une simulation de ce type sont présentés sur la figure III-15.



**Figure III-15** : Composantes du tenseur de tension thermique. Dopage:  $10^{16} \text{ cm}^{-3}$ , champ électrique:  $E_x = 5 \cdot 10^4 \text{ Vcm}^{-1}$ ,  $E_y = 0 \text{ Vcm}^{-1}$ .

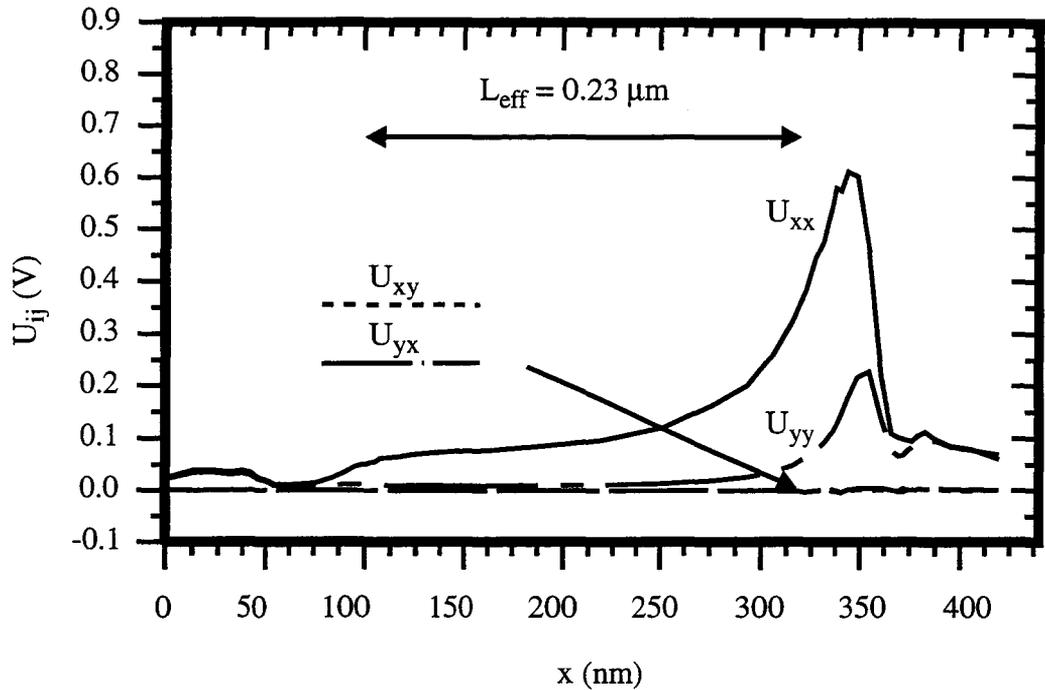
Le terme de tension thermique  $U_{xx}$  qui vaut  $\frac{k_B T}{q} = 25 \text{ mV}$  au début de la simulation augmente à mesure que les électrons acquièrent de l'énergie cinétique. Il en est de même pour le terme  $U_{yy}$ : l'accélération selon l'axe x augmente l'énergie des électrons, ce qui entraîne une plus grande dispersion des vitesses selon l'axe y. Les composantes non diagonales, par contre, restent quasiment nulles.

Une simulation avec des composantes de champ électrique égales selon x et y donne des résultats sensiblement différents (figure III-16), notamment pour les termes non diagonaux du tenseur qui ne sont plus aussi faibles. En régime stationnaire, ils restent cependant négligeables devant les termes diagonaux.



**Figure III-16** : Composantes du tenseur de tension thermique. Dopage:  $10^{16} \text{ cm}^{-3}$ , champ électrique:  $E_x = 5.10^4 \text{ Vcm}^{-1}$ ,  $E_y = 5.10^4 \text{ Vcm}^{-1}$ .

Ces résultats ne sont pas directement transposables à la simulation d'un transistor. Pour un dispositif orienté  $\langle 100 \rangle$ , même si le champ électrique selon la direction  $y$  peut être important, la vitesse des électrons selon l'axe  $y$  sera limitée par la présence de l'interface. La situation est alors différente du barreau de silicium simulé ici, et l'on pourra négliger l'influence des termes non diagonaux du tenseur de tension thermique. La figure III-17 présente les différents termes du tenseur, évalués sous l'interface, pour le transistor présenté précédemment. Les valeurs sont calculées aux noeuds du maillage, à partir des vitesses et des moments des particules, selon le schéma d'affectation présenté au paragraphe III -4.3. D'une part, comme cela apparaît déjà sur les figures précédentes, les termes diagonaux sont différents, ce qui met en défaut l'approximation faite habituellement à propos de la température du gaz électronique, considérée comme scalaire [35]. D'autre part, même pour de forts champs électriques selon  $x$  et  $y$ , les composantes non diagonales du tenseur restent négligeables, ce qui permet d'utiliser les équations du formalisme de dérive-diffusion étendu ne faisant intervenir que les termes diagonaux.



**Figure III-17** : Termes de tension thermique calculés sous l'interface, le long du transistor pour  $V_{GS} = 2.5V$ ,  $V_{DS} = 2.5V$ .

### III-5.2 Evaluation de la mobilité

En reprenant les équations du formalisme étendu de dérive-diffusion développées au paragraphe II -4.3 dans le cas du transport quasi-statique, la densité de courant est exprimée par la relation suivante:

$$\mathbf{J}_n = -qn\mathbf{V} = n\tilde{\mu}\mathbf{C}_P = nq\tilde{\mu}\mathbf{E} + nq\tilde{\mu}\nabla_{\mathbf{r}}\tilde{U} + q\tilde{\mu}\tilde{U}\nabla_{\mathbf{r}}n \quad (\text{III-24})$$

où

$$\mathbf{C}_P = \frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} \hbar \mathbf{k} d^3k \quad (\text{III-25})$$

est le taux de changement du moment.

Le tenseur de mobilité relie le courant, et donc aussi la vitesse moyenne des électrons, à  $\mathbf{C}_P$ . Ce vecteur peut être évalué lors de simulations Monte Carlo comme la moyenne de la

perte de moment des particules, due aux collisions. Il s'agit alors d'évaluer l'intégrale:

$$\frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} \hbar \mathbf{k} d^3 k = \frac{1}{n} \iint (f(\mathbf{r}, \mathbf{k}', t) S(\mathbf{k}', \mathbf{k}) - f(\mathbf{r}, \mathbf{k}, t) S(\mathbf{k}, \mathbf{k}')) \hbar \mathbf{k} d^3 k' d^3 k \quad (\text{III-26})$$

où  $S(\mathbf{k}, \mathbf{k}') d^3 k'$  représente le probabilité pour un électron de passer de l'état  $\mathbf{k}$  à un état contenu dans  $d^3 k'$  autour de  $\mathbf{k}'$ .

Une façon de calculer  $C_P$  est de réaliser la moyenne sur tous les états  $\mathbf{k}'$  possibles après un choc, connaissant les mécanismes d'interaction des électrons avec le réseau. Si l'on ne considère que des phénomènes de diffusion isotropes (interactions de type f et g, ou avec phonon acoustique), on obtient l'expression:

$$\frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} \hbar \mathbf{k} d^3 k = \frac{1}{n} \int -\lambda(\mathbf{k}) (f(\mathbf{r}, \mathbf{k}, t)) \hbar \mathbf{k} d^3 k \quad (\text{III-27})$$

où  $\lambda(\mathbf{k})$  est le taux d'interaction cumulé d'une particule de moment  $\hbar \mathbf{k}$ , pour l'ensemble des mécanismes considérés. En effet,  $\lambda(\mathbf{k}) = \int S(\mathbf{k}, \mathbf{k}') d^3 k'$ , et l'intégrale du premier terme de l'équation (III-26) est nul, pour les collisions à caractère isotrope [6].

Les collisions sur impuretés ionisées, n'ayant pas ce caractère isotrope, demandent un traitement particulier. Ce type d'interaction est élastique, donc la norme du moment se conserve, du moins dans le cas d'une structure de bande d'énergie à symétrie sphérique. L'expression donnée par l'équation (III-26) peut alors s'écrire sous la forme [40] [41]:

$$\frac{1}{n} \int \left[ \frac{\partial f}{\partial t} \right]_{coll} \hbar \mathbf{k} d^3 k = \frac{1}{n} \int -\lambda_m(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, t) \hbar \mathbf{k} d^3 k \quad (\text{III-28})$$

où  $\lambda_m(\mathbf{k})$  est le taux de perte du moment, que la connaissance du mécanisme de collision, et donc de  $S(\mathbf{k}, \mathbf{k}')$ , permet d'évaluer selon la relation suivante [42]:

$$\lambda_m(\mathbf{k}) = \int S(\mathbf{k}, \mathbf{k}') (1 - \cos(\mathbf{k}', \mathbf{k})) d^3 k' \quad (\text{III-29})$$

Pour la structure de bande d'énergie retenue dans ce travail, comportant des vallées ellipsoïdales, le calcul est effectué à l'aide de la transformation de Herring-Vogt, qui permet de revenir à un problème de symétrie sphérique, comme pour la détermination du taux d'interaction, au paragraphe II -7.2.

Le taux de perte du moment dû aux interactions sur impuretés ionisées est donné par les

relations suivantes, correspondant aux différents modèles.

Pour le modèle de Brooks-Herring [6]:

$$\lambda_{m(BH)} = \lambda_{BH} \times \frac{N_I}{n} \left( \frac{-2\mathcal{E}_\beta}{\mathcal{E}_\beta + 4\gamma(\mathcal{E})} + \frac{\mathcal{E}_\beta}{2\gamma(\mathcal{E})} \ln \left( 1 + \frac{4\gamma(\mathcal{E})}{\mathcal{E}_\beta} \right) \right) \quad (\text{III-30})$$

$$\text{avec } \lambda_{BH} = \frac{Z^2 q^2 k_B T m_d^{1/2} (1 + 2\alpha\mathcal{E})}{2^{1/2} \pi \varepsilon_{Si} \hbar^2 \gamma^{1/2}(\mathcal{E})} \quad (\text{III-31})$$

Pour le modèle de Conwell-Weisskopf:

$$\lambda_{m(CW)} = \lambda_{BH} \times \left( \frac{N_I \hbar^2 q^2}{4m_d \varepsilon_{Si} k_B T \gamma(\mathcal{E})} \ln \left( \frac{1 - \cos(\theta_m)}{2} \right) \right) \quad (\text{III-32})$$

où l'angle minimum de déviation  $\theta_m$  est donné par la formule de Rutherford (II-43), et  $\lambda_{BH}$  a la même valeur que précédemment.

Pour le modèle de Ridley, le calcul du taux de perte du moment [43] fait intervenir la fonction  $E_1(z)$  définie par:

$$E_1(z) = \int_z^\infty \frac{\exp(-t)}{t} dt \quad (\text{III-33})$$

Il est donc impossible de calculer analytiquement cette expression, et son utilisation intensive lors de simulations Monte Carlo est difficile, à moins d'avoir recours à des formules approchées, ou à des valeurs tabulées.

Finalement, le taux de changement du moment  $C_P$  est donné par l'expression suivante:

$$C_P = -\frac{1}{n} \int (\lambda(\mathbf{k}) + \lambda_m(\mathbf{k})) f(\mathbf{r}, \mathbf{k}, t) \hbar \mathbf{k} d^3 k \quad (\text{III-34})$$

où  $\lambda(\mathbf{k})$  est le taux d'interaction cumulé d'une particule de moment  $\hbar \mathbf{k}$  et  $\lambda_m(\mathbf{k})$  son taux de perte du moment dû aux collisions sur impuretés ionisées.

Lors de simulations Monte Carlo, il est donc possible d'évaluer la vitesse moyenne des électrons  $V$ , et le taux de changement du moment  $C_P$ , qui est la moyenne de  $-(\lambda(\mathbf{k}) + \lambda_m(\mathbf{k})) \hbar \mathbf{k}$ . Il s'agit ensuite, à partir de ces valeurs, d'estimer les termes de mobilité, qui pourront être utilisés pour résoudre les équations de dérive-diffusion étendues.

En reprenant l'équation (III-24), la relation entre  $V$ ,  $C_P$  et la mobilité est donnée par:

$$V = -\frac{1}{q}\tilde{\mu}C_P \quad (\text{III-35})$$

ce qui, pour un problème bidimensionnel amène le système de deux équations suivant:

$$\begin{aligned} V_x &= -\frac{1}{q}\mu_{xx}C_{P_x} - \frac{1}{q}\mu_{xy}C_{P_y} \\ V_y &= -\frac{1}{q}\mu_{yx}C_{P_x} - \frac{1}{q}\mu_{yy}C_{P_y} \end{aligned} \quad (\text{III-36})$$

Ce système fait intervenir quatre inconnues, qui sont les termes du tenseur de mobilité, et ne peut donc pas être résolu sans imposer de contrainte supplémentaire. Le moyen le plus simple de décrire, pour une position donnée, la relation entre la vitesse moyenne des électrons et le taux de changement du moment est de considérer que les termes non diagonaux sont nuls; on obtient alors:

$$\mu_{xx} = -\frac{qV_x}{C_{P_x}} ; \mu_{yy} = -\frac{qV_y}{C_{P_y}} \quad (\text{III-37})$$

Lors de simulations utilisant le formalisme classique de dérive-diffusion, la mobilité est considérée comme une fonction locale du champ électrique et des concentrations en dopants, comme cela a été présenté au chapitre précédent. Lorsqu'il s'agit de prendre en compte des phénomènes de transport non stationnaires, il n'est plus possible de considérer la mobilité en un point du dispositif comme une fonction de variables locales. En effet les termes de mobilité dépendent de la distribution des électrons à la position où ils sont évalués, cette distribution étant elle-même reliée à l'ensemble du transport électronique. Ainsi, les composantes de vitesse et de taux de changement dans les égalités (III-37) sont liées au comportement de l'ensemble du dispositif. Ces égalités permettent, lors de simulations utilisant le formalisme étendu de dérive-diffusion, de retrouver les mêmes valeurs de densité de courant que par la méthode de Monte Carlo. L'utilisation des termes de mobilité ainsi évalués dans les équations de dérive-diffusion étendues permet d'y introduire les phénomènes de transport non stationnaires, de façon cohérente avec le formalisme utilisé.

A titre d'illustration, la valeur de mobilité évaluée lors d'une simulation Monte Carlo est comparée à celle extraite d'une courbe représentant la vitesse de dérive des électrons en fonction d'un champ électrique uniforme. La valeur calculée à l'aide de l'équation (III-37) pour un dopage uniforme  $N_I = 10^{16} \text{ cm}^{-3}$  et un champ uniforme  $E_x = 10^3 \text{ Vcm}^{-1}$  est  $\mu_{xx} = 1,13 \cdot 10^3 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  ce qui concorde tout à fait avec celle donnée par la pente de la courbe représentée sur la figure III-18 à champ faible qui donne  $\mu_{xx} = 1,1 \cdot 10^3 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . Ces deux estimations de la mobilité à champ faible sont elles-mêmes cohérentes avec les modèles de mobilité couramment utilisés [45] [4] [3].

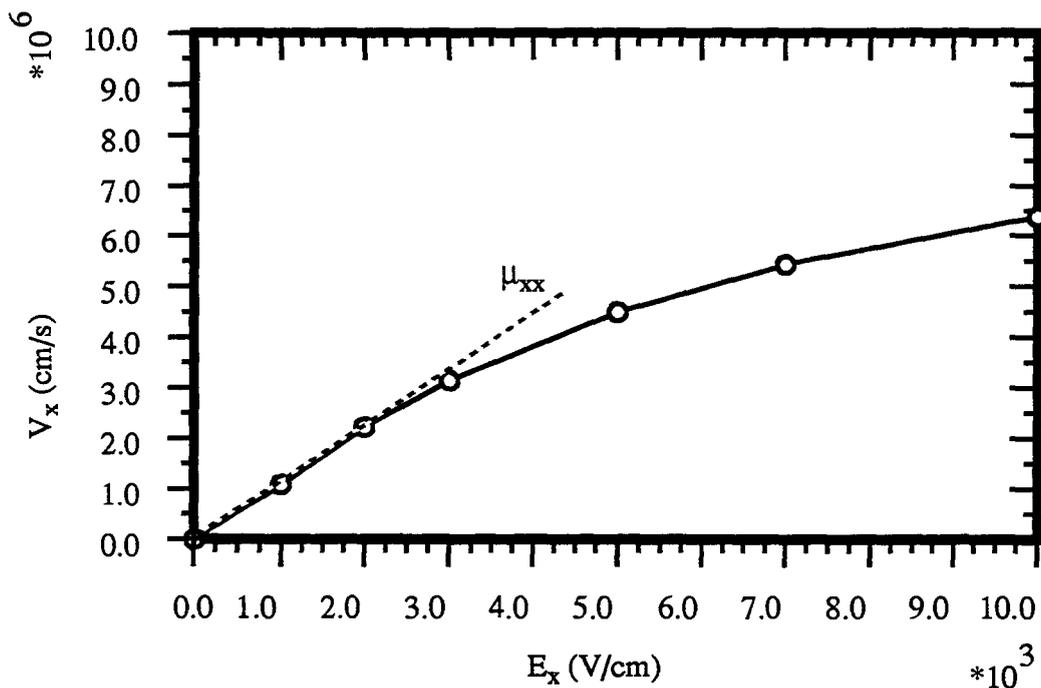
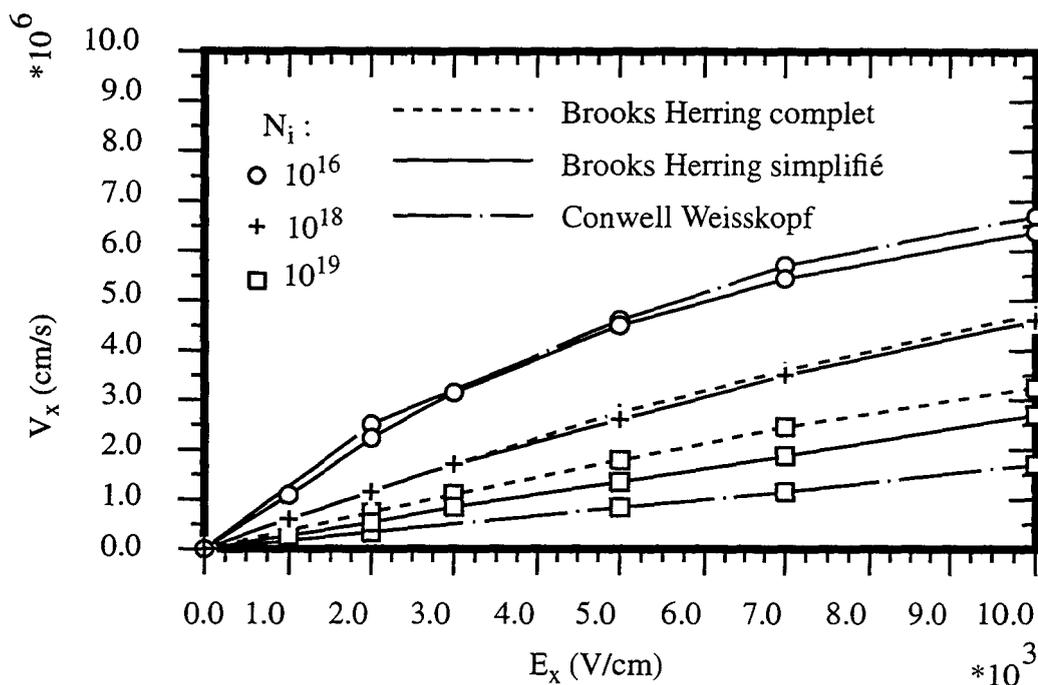


Figure III-18 : Vitesse de dérive des électrons en fonction du champ électrique.

$$N_I = 10^{16} \text{ cm}^{-3}.$$

La figure suivante présente les résultats obtenus à l'aide de trois modèles de collisions sur impuretés ionisées: celui de Conwell Weisskopf, le modèle de Brooks-Herring complet décrit au chapitre précédent (II-45) et sa version simplifiée pour laquelle la longueur d'onde de de Broglie de l'électron est négligée devant la longueur d'écrantage (II-48)



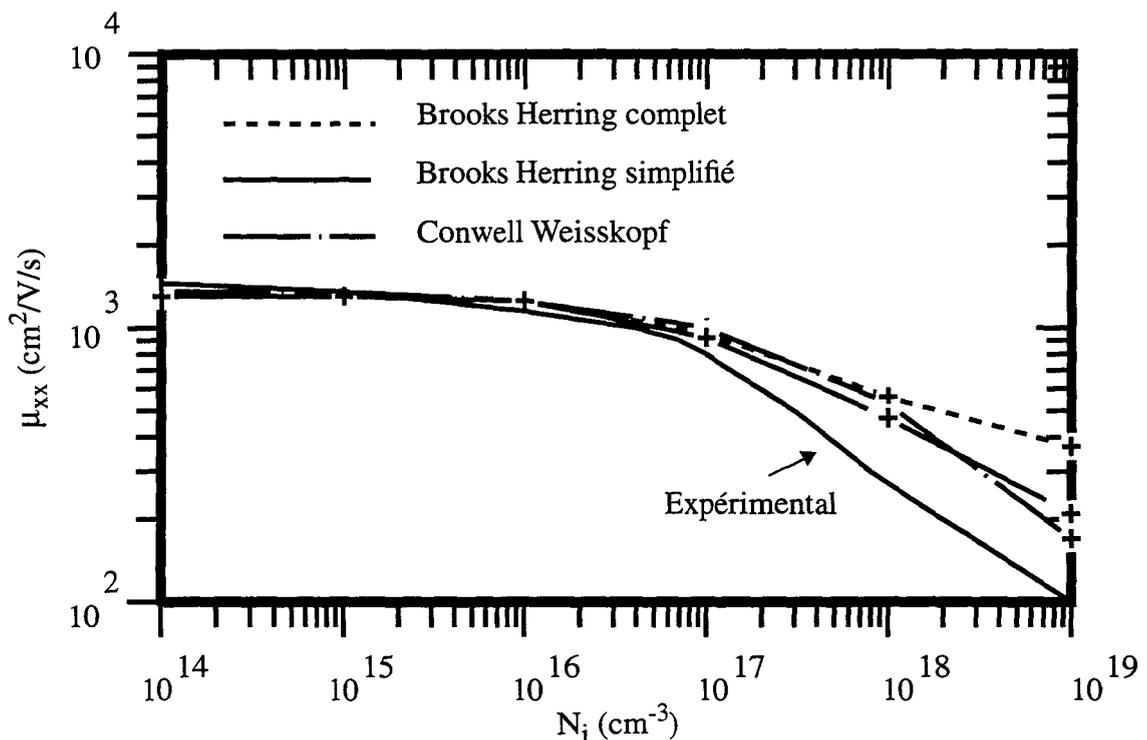
**Figure III-19** : Vitesse de dérive des électrons en fonction du champ électrique pour différents dopages, selon le modèle d'interaction sur impureté ionisée utilisé.

Comme cela a déjà été remarqué, les modèles de Brooks Herring complet et simplifié sont équivalents pour les concentrations en impuretés inférieures à  $10^{18} \text{ cm}^{-3}$ . Par contre, pour  $N_I = 10^{19} \text{ cm}^{-3}$ , le modèle simplifié sous-estime la vitesse moyenne par rapport au modèle complet. Une autre difficulté liée à l'utilisation du modèle simplifié à très forte concentration est que la mobilité évaluée à l'aide du taux de perte du moment ne correspond pas à la pente de la courbe représentant la vitesse en fonction du champ. Cette expression simplifiée, si elle permet de diminuer la durée des simulations Monte Carlo, peut poser problème dans le cadre du couplage Monte Carlo-dérive diffusion étendu, les valeurs de mobilité à forte concentration n'étant pas cohérentes lors des deux types de simulation.

Si pour les faibles concentrations, les modèles de Brooks-Herring et de Conwell-Weisskopf sont équivalents, comme l'indiquent les courbes correspondant à  $N_I = 10^{16} \text{ cm}^{-3}$ , ils donnent des résultats différents à forte concentration. La mobilité évaluée à l'aide du modèle de Conwell-Weisskopf est nettement inférieure pour  $N_I = 10^{19} \text{ cm}^{-3}$ .

La figure III-20 présente la mobilité évaluée lors de simulations Monte Carlo à partir du

taux de perte du moment et de la vitesse moyenne, ainsi que des valeurs expérimentales. Les simulations sont réalisées en considérant un dopage uniforme et un faible champ électrique. La première courbe est conforme à la mobilité calculée directement à partir du modèle de Brooks-Herring [44]. La mobilité décroît à forte concentration, mais moins que ce qui est indiqué par la courbe expérimentale. Pour  $N_I = 10^{19} \text{ cm}^{-3}$ , la mobilité est environ trois fois plus élevée que la valeur expérimentale [45], ce qui est lié à la difficulté de modélisation des collisions sur impuretés ionisées à très forte concentration, dans le cas d'un semiconducteur dégénéré. Les valeurs fournies par les simulations Monte Carlo sont par contre correctes depuis les faibles concentrations jusqu'aux valeurs légèrement supérieures à  $N_I = 10^{17} \text{ cm}^{-3}$ . Etant le plus proche des valeurs expérimentales, c'est le modèle de Conwell-Weisskopf qui semble le mieux adapté, malgré un écart de presque un facteur deux à forte concentration.



**Figure III-20** : Mobilité en fonction de la concentration en dopants ( $E_x = 2 \cdot 10^3 \text{ V cm}^{-1}$ ).

La courbe expérimentale provient de *Physics of semiconductor devices*, Sze [45].

Les expressions de  $\mu_{xx}$  et  $\mu_{yy}$  données en (III-37) introduisent une difficulté quant à leur évaluation lors de simulations Monte Carlo. En effet, si la densité de courant  $J$  est orientée dans l'une des directions  $x$  ou  $y$ , l'un des termes  $C_{Py}$  ou  $C_{Px}$  présents au dénominateur est nul. Si le courant est orienté selon l'axe  $x$ , la mobilité dans la direction  $y$  n'est pas définie, le mouvement moyen selon l'axe  $y$  étant nul. Lorsque ces valeurs de mobilité sont utilisées pour résoudre les équations de dérive-diffusion étendues, par exemple dans le cas d'un courant selon la direction  $x$ , la connaissance de la mobilité  $\mu_{xx}$  est importante pour déterminer la densité de courant selon  $x$ . Par contre, la densité de courant  $J_y$  évaluée à l'aide de la relation (III-24) est nulle quelle que soit la valeur affectée au terme  $\mu_{yy}$ , puisque les contributions du champ électrique et des gradients de tension thermique et de concentration selon  $y$  se compensent. Lorsque cette situation se présente lors de simulations Monte Carlo de dispositif, la valeur attribuée à la mobilité selon  $y$  est celle qui est évaluée selon  $x$ :  $\mu_{yy} = \mu_{xx}$ . Dans l'hypothèse inverse d'un courant uniquement selon  $y$ , la mobilité selon  $x$  serait prise égale à la mobilité selon  $y$ . Pour ces cas particuliers d'orientation du courant quasiment selon un axe  $\langle 100 \rangle$ , cela revient à considérer la mobilité comme un scalaire [46] [47], ce qui est le plus souvent une bonne approximation, dans la mesure où les expressions de  $\mu_{xx}$  et  $\mu_{yy}$  montrent que ceux-ci dépendent plus de l'énergie des électrons, que de la direction de leur mouvement. En effet, pour observer une différence notable entre les deux termes, due à l'anisotropie des vallées d'énergie, il faudrait une importante redistribution des électrons privilégiant certaines vallées, afin que la vitesse moyenne et le moment moyen aient des directions différentes.



### REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE III

- [1] : D. Collard, B. Baccus, E. Dubois, D. Morel, *Impact 1-2-3: an integrated 2D process/device simulator for MOS technology*, in lecture notes of the software forum, NASECODE VI conf., p16-31, 1989
- [2] : B. Baccus, D. Collard, E. Dubois, D. Morel, *IMPACT 4 - a general two-dimensional multilayer process simulator*, proc. of SISDEP III conf., p255-266, Ed. by G. Bacarani and M. Rudan, Bologna, 1988
- [3] : E. Dubois, Thèse de doctorat en sciences des matériaux, Université des Sciences et Techniques de Lille Flandres-Artois, 1990
- [4] : S. Selberherr, *Analysis and simulation of semiconductor devices*, Ed. Spinger Verlag, 1984.
- [5] : S. Bandyopadhyay, M. E. Klausmeier-Brown, C. M. Maziar, S. Datta, M. S. Lundstrom, *A rigorous technique to couple Monte Carlo and drift-diffusion models for computationally efficient device simulation*, IEEE Trans. Elec. Dev., vol.34, No.2, p392, février 1987
- [6] : H. Kosina, S. Selberherr, *Efficient coupling of Monte Carlo and drift-diffusion method with applications to MOSFETs*, Extended Abstracts of the 22nd Conference on Solid State Devices and Materials, Sendai, 1990, p139
- [7] : Zimmermann, Thèse de doctorat d'état, Université des Sciences et Techniques de Lille Flandres-Artois, 1980
- [8] : C. Jacoboni, P. Lugli, *The Monte Carlo method for semiconductor device simulation*, Ed. S. Selberherr, Springer-Verlag
- [9] : E. Sangiorgi, B. Riccò, F. Venturi, *MOS2: An efficient MOnTe Carlo Simulator for MOS devices*, IEEE Trans. on CAD, vol.7, No.2, février 1988
- [10] : F. Venturi, R. K. Smith, E. C. Sangiorgi, M. R. Pinto, B. Riccò, *A general purpose device simulator coupling Poisson and Monte Carlo transport with applications to deep submicron MOSFETs*, IEEE Trans. on CAD, vol.8, No.4, avril 1989

- [11] : C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, A. Alberigi-Quaranta, *Electron drift velocity in silicon*, Phys. Rev. B, vol.12, p2265, 1975
- [12] : G. Baccarani, *Physics of submicron devices*, Proc. VLSI Process and Device Modeling, K.U.L. Summer Course, p1-23, 1983
- [13] : G. Hobler, S. Selberherr, *Monte Carlo simulation of ion implantation into two- and three-dimensional structures*, IEEE Trans. on CAD, vol.8, No.5, mai 1989
- [14] : M. Charef, Thèse de doctorat de troisième cycle, Université des Sciences et Techniques de Lille , 1983
- [15] : Chu-Hao, J. Zimmermann, M. Charef, R. Fauquembergue, E. Constant, *Monte Carlo study of two-dimensional electron gas transport in Si-MOS devices*, Solid State Elec., vol.28, No.8, p733, 1985
- [16] : Y.-J. Park, T.-W. Tang, D. H. Navon, *Monte Carlo surface scattering simulation in MOSFET structures*, IEEE Trans. Elec. Dev., vol.30, No.9, p1110, septembre 1983
- [17] : M. V. Fischetti, S. E. Laux, *Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects*, Phys. Rev. B, vol.38, No.14, p9721, novembre 1988
- [18] : E. Sangiorgi, M. R. Pinto, *A semi-empirical model of surface scattering for Monte Carlo simulation of silicon n-MOSFETs*, IEEE Trans. Elec. Dev., vol.39, No.2, p356, février 1992
- [19] : G. Nanz, *A critical study of boundary conditions in device simulation*, proc. of SIS-DEP IV conf., p321, Ed. by W. Fichtner, D. Aemmer, Zurich, Hartung-Gorre, 1991
- [20] : G. Nanz, P. Dickinger, S. Selberherr, *Calculation of contact currents in device simulation*, IEEE Trans. on CAD, vol.11, No.1, janvier 1992
- [21] : K. Kato, *Hot-carrier simulation for MOSFETs using a high-speed Monte Carlo method*, IEEE Trans. Elec. Dev., vol.35, No.8, p1344, août 1988
- [22] : P. Lugli, C. Jacoboni, *Monte Carlo simulation of semiconductor devices: a critical review*, ESSDERC87, p861, Ed. by G. Soncini and P.U. Calzolari, Elsevier Science Publishers, 1988

- [23] : P. T. Nguyen, D. H. Navon, T.-W. Tang, *Boundary conditions in regional Monte Carlo device analysis*, IEEE Trans. Elec. Dev., vol.32, No.4, p783, avril 1985
- [24] : D. Y. Cheng, C. G. Hwang, R. W. Dutton, *PISCES-MC: a multiwindow, multimethod 2-D device simulator*, IEEE Trans. on CAD, vol.7, No.9, p1017, septembre 1988
- [25] : P. Hesto, J. F. Pone, R. Castagné, *Unidimensional particule model of conduction in n type GaAs at low temperature: contact influence*, Journal de Physique, supp. C7 au No.10, tome 42, octobre 1981
- [26] : L. Rajaonarison, P. Hesto, J. F. Pone, P. Dollfus, *Monte Carlo simulation of submicronic devices and processes*, proc. of SISDEP IV conf., p513, Ed. by W. Fichtner, D. Aemmer, Zurich, Hartung-Gorre, 1991
- [27] : P. Lugli, U. Ravaioli, D. K. Ferry, *Monte Carlo simulation of contacts in submicron devices*, American Institute of Physics conf. proc., No.122, *The physics of VLSI*, p162, Palo Alto 1984
- [28] : U. Ravaioli, P. Lugli, M. A. Osman, D. K. Ferry, *Advantages of collocation methods over finite differences in one-dimensional Monte Carlo simulations of submicron devices*, IEEE Trans. Elec. Dev., vol.32, No.10, p2097, octobre 1985
- [29] : G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, E. Ganin, S. Rishton, D. S. Zicherman, H. Schmid, M. R. Polcari, H. Y. Ng, P. J. Restle, T. H. P. Chang, R. H. Denard, *Design and experimental technology for 0.1- $\mu\text{m}$  gate-length low-temperature operation FET's*, IEEE Elec. Dev. Let., vol.8, No.10, octobre 1987
- [30] : M. V. Fischetti, S. E. Laux, D. J. Frank, *Monte Carlo analysis of semiconductor devices: the Damocles program*, IBM J. Res. Develop. vol.34, No.4, p466, juillet 1990
- [31] : C. Jacoboni, *A new approach to Monte Carlo simulation*, IEDM 89, p469
- [32] : A. Phillips, P. J. Price, *Monte Carlo calculations of hot electron energy tails*, Appl. Phys. Lett. vol.54, p528 1977
- [33] : Y. Yamada, S. Ikeda, N. Shimojoh, *The particle simulation of self-aligned GaAs MESFETs with a submicrometer gate-length*, ESSDERC87, p875, Ed. by G. Soncini and P.U. Calzolari, Elsevier Science Publishers, 1988

- [34] : S.-C. Lee, T.-W. Tang, *Transport coefficients for a silicon hydrodynamic model extracted from inhomogeneous Monte-Carlo calculations*, Solid-State Elec. vol.35, No.4, p 561, 1992
- [35] : M. A. Stettler, M. A. Alam, M. S. Lundstrom, *A critical examination of the assumptions underlying macroscopic transport equations for silicon devices*, IEEE Trans. Elec. Dev., vol.40, No.4, p733, avril 1993
- [36] : A. Yoshii, M. Tomizawa, K. Yokoyama, *Accurate modeling for submicrometer-gate Si and GaAs MESFET's using two-dimensional particle simulation*, IEEE Trans. Elec. Dev., vol.30, No.10, p1376, octobre 1983
- [37] : M. Tomizawa, K. Yokoyama, A. Yoshii, *Nonstationary carrier dynamics in quarter-micron Si MOSFET's*, IEEE CAD, vol.7, No.2, p254, février 1988
- [38] : M. V. Fischetti, *Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures-Part I: homogeneous transport*, IEEE Trans. Elec. Dev., vol.38, No.3, p634, mars 1991
- [39] : M. V. Fischetti, S. E. Laux, *Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures-Part II: sub-micrometer MOSFET's*, IEEE Trans. Elec. Dev., vol.38, No.3, p650, mars 1991
- [40] : P. Kireev, *La physique des semi-conducteurs*, Ed. Mir, 1975.
- [41] : Y.-K. Feng, A. Hintz, *Simulation of submicron GaAs Mesfet's using a full dynamic transport model*, IEEE Trans. Elec. Dev., vol.35, No.9, p1419, septembre 1988
- [42] : E. M. Conwell, *High-field transport in semiconductors*, in *Solid State Physics*, supplément 9, Academic Press, 1967
- [43] : B. K. Ridley, *Reconciliation of the Conwell-Weisskopf and Brooks-Herring formulae for charged-impurity scattering in semiconductors: third body interference*, J. Phys. C, vol.10, p1589, 1977
- [44] : L. E. Kay, T.-W. Tang, *An improved ionized-impurity scattering model for Monte Carlo simulations*, J. Appl. Phys., vol.70, p1475, août 1991
- [45] : S. M. Sze, *Physics of semiconductor devices*, John Wiley & Sons, 1981

- [46] : H. Kosina, P. Lindorfer, S. Selberherr, *Monte Carlo - Poisson coupling using transport coefficients*, ESSDERC91, p53, Elsevier Science Publishers, 1991
- [47] : H. Kosina, S. Selberherr, *A Monte Carlo MOSFET simulator based on a new method for the Poisson-transport iteration*, Proc. NUPAD 92



## CHAPITRE IV

### ETUDE DE TRANSISTORS NMOS FORTEMENT SUBMICRONIQUES

#### IV-1 INTRODUCTION

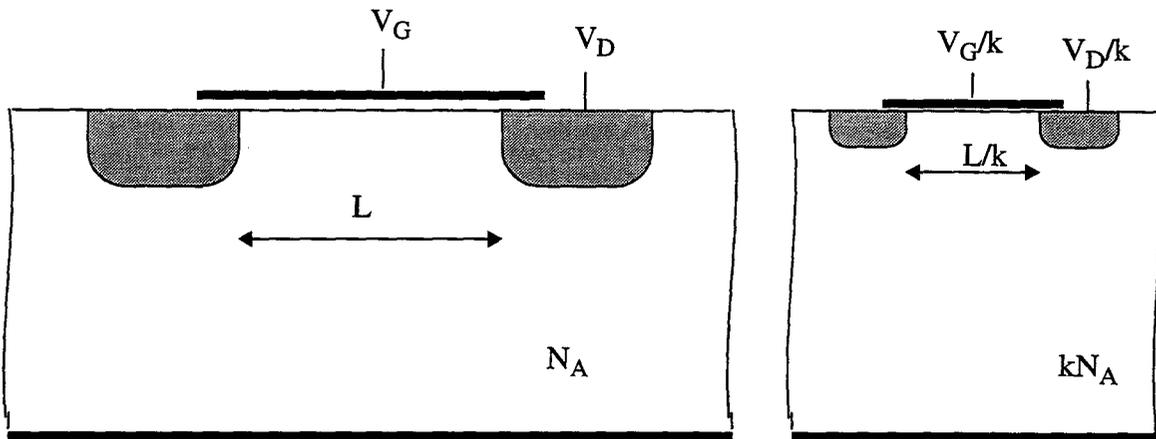
La recherche de plus grandes performances, en termes de rapidité et de compacité, conduit à réduire toujours plus la taille des dispositifs microélectroniques élémentaires. Des dispositifs plus petits et plus proches les uns des autres autorisent en effet des vitesses de fonctionnement plus élevées, liées à l'augmentation de la vitesse intrinsèque des transistors et à la diminution des capacités parasites [1] [2]. D'autre part, l'intégration d'un grand nombre de composants sur un même circuit intégré permet de traiter des tâches plus complexes. Cette recherche de miniaturisation s'accompagne d'un certain nombre de défis technologiques concernant la fabrication des circuits: qualité des matériaux, précision, fiabilité et reproductibilité des étapes technologiques, pour ne citer que ces aspects. Un travail important doit également être réalisé concernant l'interconnection des dispositifs, avec là aussi des contraintes technologiques liées à la précision et à la fiabilité. Les problèmes de consommation et de dissipation de chaleur liés à la densité croissante des circuits intégrés doivent aussi être résolus.

La conception des composants doit être optimisée, de façon à fixer leurs performances dans une plage connue. Ceci demande l'étude et la compréhension des phénomènes intervenant dans le comportement électrique. Cette compréhension et l'optimisation des dispositifs utilisent largement des outils de simulation, concernant à la fois les étapes technologiques de fabrication et le fonctionnement des composants. Les simulateurs sont cependant limités par les modèles sous-jacents et peuvent dans certains cas ne fournir que des informations qualitatives sur le comportement des composants microélectroniques. Les

simulateurs développés dans le cadre de ce travail, par le degré de sophistication des modèles utilisés et par leur relative simplicité d'utilisation, sont des outils efficaces pour étudier les dispositifs dont les longueurs caractéristiques sont inférieures au micron. Ils permettent d'en simuler le fonctionnement électrique et d'étudier en détail les phénomènes de transport qui s'y produisent. Ce chapitre présente les résultats obtenus concernant l'étude de transistors nMOS plans submicroniques.

## IV-2 STRATEGIE DE MINIATURISATION

La miniaturisation des transistors MOS est basée sur une réduction des dimensions caractéristiques d'un facteur  $k$ , qui se traduit par un changement d'échelle. La règle de miniaturisation la plus classique consiste à appliquer le même coefficient de réduction dans toutes les directions. L'épaisseur d'oxyde, la longueur et la largeur du canal, les profondeurs de jonction sont ainsi réduites d'un facteur  $k$ . L'objectif étant dans ce cas de conserver un comportement de type "canal long", le dopage de substrat  $N_A$  est multiplié par  $k$  et les tensions de polarisation diminuées d'un facteur  $k$ , ce qui permet de garder un champ électrique maximal constant, et de réduire la largeur des zones déplétées entourant les jonctions de source et drain du même facteur  $k$ . La figure IV-1 résume ces différentes transformations. Pour obtenir un transistor "à l'échelle", il est aussi nécessaire de réduire la tension de seuil d'un facteur  $k$ , de façon cohérente avec la diminution de la plage de tension de grille utilisée [1] [3]. Cet ajustement de tension de seuil peut être réalisé par une implantation supplémentaire de dopants. Pour cette stratégie de miniaturisation dite "à champ maximal constant", le courant de saturation et les temps de commutation sont divisés par  $k$ . La puissance dissipée par une porte logique est divisée par  $k^2$ , ainsi que la surface qu'elle occupe. La puissance dissipée par unité de surface est ainsi maintenue constante. Si les dimensions des interconnexions sont elles aussi réduites d'un facteur  $k$ , les constantes de temps de ces lignes sont malheureusement inchangées [1] [2], ce qui illustre l'importance des lignes de connection pour les performances d'un circuit.



**Figure IV-1 :** Schéma de réduction à l'échelle  $k$  d'un transistor MOS, d'après Sze [3]

D'autres stratégies de miniaturisation sont possibles. Pour des raisons de compatibilité de circuits logiques, la réduction peut être recherchée à tension constante, ce qui permet de conserver les mêmes tensions d'alimentation et des niveaux logiques identiques. Différents facteurs peuvent être appliqués aux dopages et aux épaisseurs d'oxyde, selon les performances recherchées [4] [5]. Ces approches se traduisent cependant par l'apparition de forts champs électriques qui en limitent l'utilisation.

Les différentes stratégies de miniaturisation se heurtent à des problèmes technologiques tels que la réduction de l'épaisseur d'oxyde de grille. Il est en effet difficile de réaliser des oxydes minces très uniformes. Par ailleurs, les potentiels de contacts ne sont pas toujours modifiables. Un certain nombre d'autres paramètres ne peuvent être réduits à l'échelle, et la miniaturisation introduit alors des effets négatifs.

En première approximation, la pente sous le seuil  $S$  qui traduit le pouvoir de coupure du transistor est inchangée lors d'une réduction à champ maximal constant. En effet,  $S$  peut s'exprimer sous la forme [3]:

$$S \approx \frac{k_B T}{q} \ln(10) \left( 1 + \frac{C_D}{C_{ox}} \right) \quad (\text{IV-1})$$

et les capacités de déplétion  $C_D$  et d'oxyde  $C_{ox}$  sont réduites toutes les deux d'un facteur  $k$ .

$S$  a une limite physique qui est, pour une température de 300K:

$$S_{min} = \frac{k_B T}{q} \ln(10) \approx 60 \text{ mV/dec} \quad (\text{IV-2})$$

Sous la tension de seuil, la tension de grille doit diminuer d'au moins 60 mV pour réduire le courant d'un facteur 10, mais des valeurs de 80 mV à 90 mV sont courantes. Lors de la miniaturisation, la fourchette de tension diminue, et le courant résiduel traversant le transistor bloqué augmente de concert avec la réduction des dimensions [6]. Pour améliorer le pouvoir de coupure du transistor, il faudrait réduire le dopage du substrat afin de diminuer la capacité  $C_D$  [3]. Or les concentrations sont multipliées par le facteur d'échelle  $k$  dans l'optique d'une miniaturisation à champ maximal constant. Il y a donc un compromis à réaliser en ce qui concerne la concentration de substrat.

La valeur de la tension de seuil doit également être d'autant mieux maîtrisée que la plage de tension de grille est réduite, ce qui pose problème dans la mesure où elle dépend de la température, et surtout des fluctuations de dopage [7].

Enfin, l'augmentation du champ électrique vertical due au dopage plus important a des répercussions sur la mobilité dans le canal, qui est réduite, et donc sur la conductivité du canal [6] [10]. Cette augmentation du champ électrique vertical peut de plus introduire des effets quantiques susceptibles de modifier les caractéristiques électriques [8]. Enfin, le champ électrique vertical doit être diminué pour limiter l'injection de charge à travers l'oxyde par effet tunnel [13].

La miniaturisation d'un dispositif MOS se traduit finalement par un compromis entre les différents effets cités. Selon les objectifs recherchés, plusieurs stratégies sont possibles:

- \* réduire la température de fonctionnement, ce qui permet d'obtenir une transconductance plus élevée qu'à température ambiante [9] [13]. Cela permet également de diminuer le courant dans le transistor bloqué en améliorant le pouvoir de coupure du transistor [7] et en augmentant la tension de seuil [3].

- \* mettre à profit le phénomène de survitesses pour compenser la réduction de mobilité due au champ électrique vertical [10] [13] [14] [15]

\* jouer sur la géométrie des zones de source et de drain pour avoir des jonctions peu profondes et limiter les effets canaux courts [6] [16] [23], tout en utilisant des dopages élevés et des siliciures pour diminuer les résistances d'accès [18] [19] [20]

\* modifier la géométrie du canal, dans le profil de dopage en produisant une faible concentration en surface [16] [17] [21] [22], ou dans sa forme en réalisant une grille enterrée [24] [25] [26] [27]

### IV-3 PRINCIPALES ETAPES TECHNOLOGIQUES DE FABRICATION

Les résultats de simulation exposés ci-après correspondent au développement de dispositifs nMOS submicroniques présentés par Sai-Halasz [13] [14] [15]. Cette technologie a été retenue parmi d'autres dans la mesure où les articles cités fournissent des données permettant la simulation des étapes de fabrication [28] et présentent les caractéristiques électriques de transistors réalisés dans les laboratoires d'IBM. Cette même technologie a par ailleurs donné lieu à des simulations utilisant la méthode de Monte Carlo, par Fischetti, Laux et Frank [29], à l'aide du programme Damoclès. Ce simulateur utilise une structure de bande d'énergie tabulée, valable même à haute énergie, et prend en compte les différentes interactions de façon plus précise, ce qui demande d'importants moyens informatiques. Le programme Monte Carlo développé dans le cadre du présent travail, plus simple, fournit cependant des résultats tout à fait comparables.

Les principales données relatives à la fabrication [13] qui définissent la géométrie des dispositifs sont résumées ainsi:

\* substrat de type p, de résistivité  $2\Omega\text{cm}$ , soit une concentration en bore d'environ  $6.5 \cdot 10^{15}\text{cm}^{-3}$

\* implantation de bore pour l'ajustement de la tension de seuil et le contrôle du perçage. Doses de  $2.5$  à  $5.5 \cdot 10^{12}\text{cm}^{-2}$ , énergie de 15 et 30 keV. La simulation utilise une implantation de  $5.5 \cdot 10^{12}\text{cm}^{-2}$ , sous une énergie de 30 keV

\* réalisation de l'oxyde de grille, d'épaisseur comprise entre 33 et 45 Å. L'épaisseur retenue pour la simulation est 45 Å.

\* réalisation des extensions de source et drain par implantation d'antimoine. La dose est de

$1.5 \cdot 10^{14} \text{cm}^{-2}$  et l'énergie de 16 keV pour la simulation, alors que pour les essais réalisés, la dose atteint quelques  $10^{14} \text{cm}^{-2}$  et l'énergie descend à 10 keV

\* fabrication des espaceurs à l'aide d'un dépôt de nitrure de 100nm suivi d'une gravure

\* implantation d'arsenic pour les zones de source et drain, sous une énergie de 20 keV, pour des doses de 2 à  $5 \cdot 10^{15} \text{cm}^{-2}$  (simulation:  $2 \cdot 10^{15} \text{cm}^{-2}$ )

\* budget thermique d'activation des dopants de source et drain compris entre 30mn à  $900^\circ\text{C}$  et 20mn à  $850^\circ\text{C}$

\* la profondeur finale des jonctions de source et de drain est d'environ 100nm, et de 50nm pour les extensions de type LDD<sup>1</sup>

La simulation des étapes de fabrication décrites ci-dessus à l'aide du simulateur de procédé IMPACT2 aboutit au profil de dopage suivant, sur lequel la double diffusion arsenic/antimoine apparaît clairement:

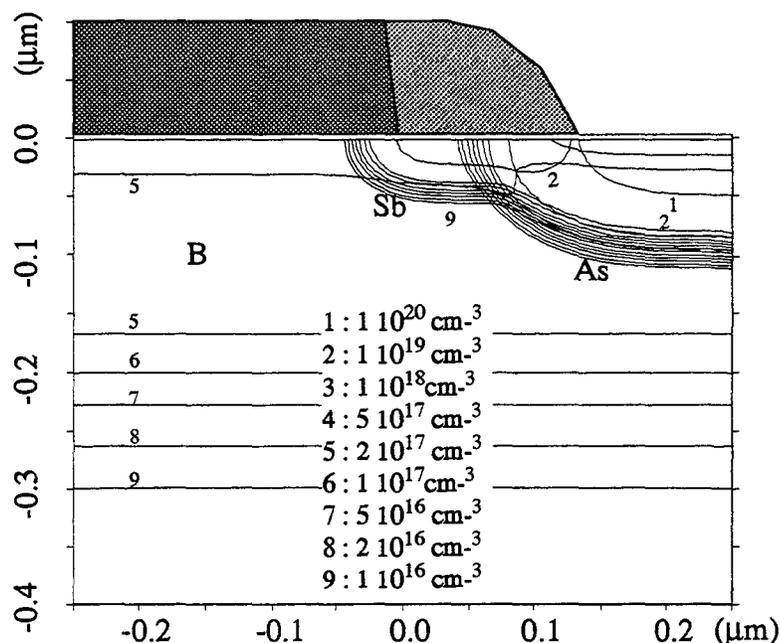


Figure IV-2 : Résultat de simulation bidimensionnelle des étapes de fabrication du transistor NMOS, d'après [13]. (Le dispositif étant symétrique, seul un demi transistor est simulé.)

1. LDD: Lightly Doped Drain. L'utilisation de deux implantations permet de réduire le champ maximal par un passage progressif aux fortes concentrations des régions de source et de drain. De plus, il est possible d'obtenir une jonction fine près du bord de grille tout en limitant les résistances d'accès grâce aux zones plus profondes [6].

## IV-4 PHENOMENE DE SURVITESSE

### IV-4.1 Dépendance par rapport à la longueur

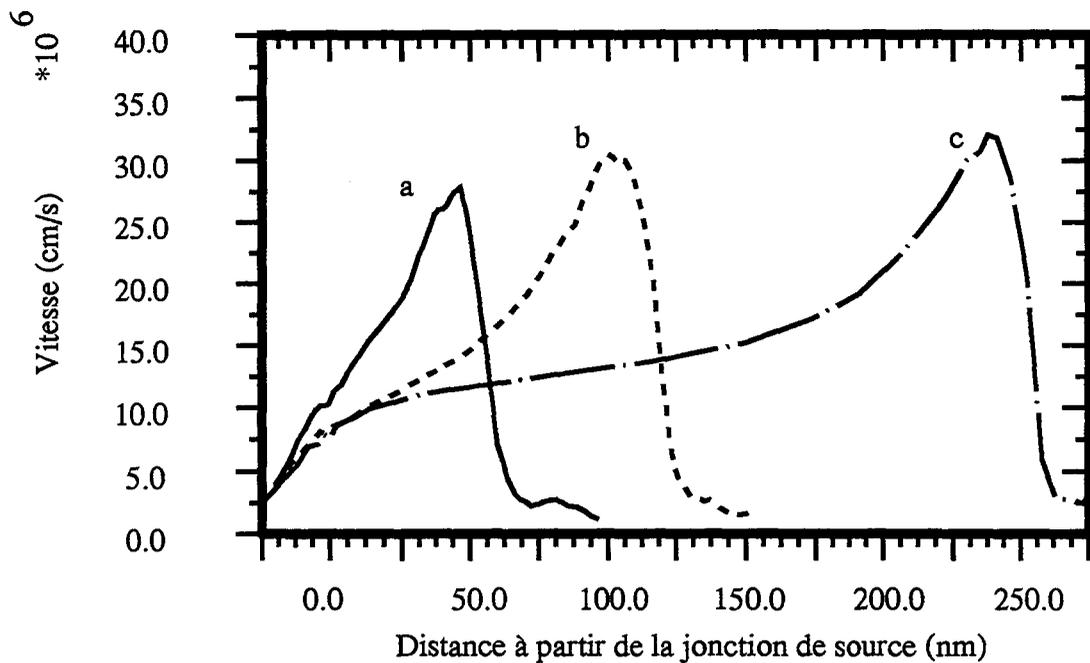
Les transistors présentés par Sai-Halasz ont des transconductances élevées à basse température (77K - azote liquide) et à température ambiante, ce qui s'explique par l'influence du phénomène de survitesse sur le courant source-drain [14] [15]. La simulation de tels transistors par la méthode de Monte Carlo permet de mettre clairement en évidence le fait que les électrons ont dans une partie du canal une vitesse supérieure à la vitesse de saturation [29] [30]. Des transistors de longueurs de grille différentes, dont le profil de dopage est décrit par la figure IV-2 ont ainsi été simulés. Conformément aux règles de miniaturisation, les tensions de polarisation sont réduites à mesure que la longueur effective de canal diminue. Les premières simulations correspondent aux polarisations suivantes:

$$* L_{\text{eff}} = 233\text{nm} \quad V_{\text{GS}} = 2.5\text{V} \quad V_{\text{DS}} = 2.5\text{V}$$

$$* L_{\text{eff}} = 103\text{nm} \quad V_{\text{GS}} = 1.0\text{V} \quad V_{\text{DS}} = 1.0\text{V}$$

$$* L_{\text{eff}} = 43\text{nm} \quad V_{\text{GS}} = 0.7\text{V} \quad V_{\text{DS}} = 0.6\text{V}$$

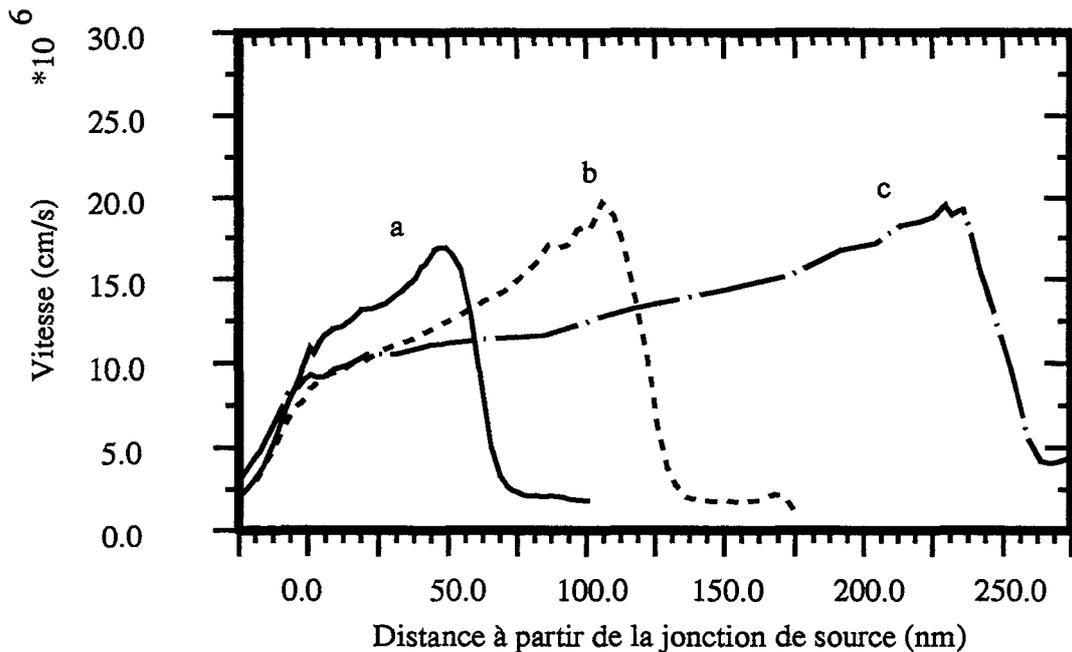
Les courbes de la figure IV-3 représentent la vitesse des électrons juste sous l'interface Si/SiO<sub>2</sub> pour les différentes longueurs de canal, et les suivantes donnent la vitesse moyenne des électrons calculée sur une profondeur de 10nm.



**Figure IV-3 :** Vitesse des électrons sous l'interface Si/SiO<sub>2</sub>. (a:  $L_{\text{eff}} = 43\text{nm}$ , b:  $L_{\text{eff}} = 103\text{nm}$ , c:  $L_{\text{eff}} = 233\text{nm}$ )

Les simulations dont ces résultats sont extraits ont mis en jeu environ 25000 particules pendant deux picosecondes, les moyennes étant effectuées pendant la deuxième picoseconde, afin de s'affranchir des effets transitoires liés au démarrage de la simulation. Il est à noter que lors de ces simulations, les interactions des électrons avec les impuretés ionisées ont été décrites par un modèle de Brooks-Herring simplifié indépendant des concentrations, et que les interactions des électrons avec l'interface ont été traitées comme de simples collisions spéculaires.

Le phénomène de survitesse est présent pour chacune des trois longueurs de canal. Pour un canal très court, les électrons ont une vitesse supérieure à la vitesse de saturation dès la jonction entre source et substrat. La survitesse prend place sur l'ensemble du canal, alors que pour les transistors plus longs, elle n'intervient que près de la zone de drain, sans influence systématique sur le courant de drain. Le second jeu de courbes, représenté en figure IV-4, permet de mieux décrire la vitesse moyenne des électrons dans le canal et l'influence qu'elle peut avoir sur le courant.

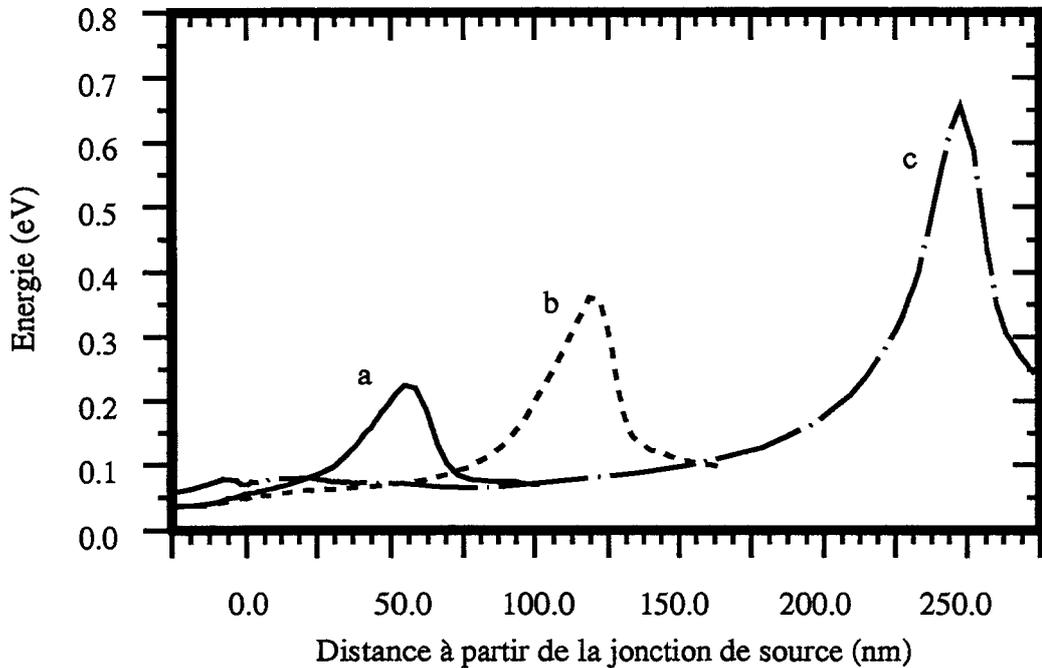


**Figure IV-4 :** Vitesse moyenne des électrons calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub>. (a:  $L_{\text{eff}} = 43\text{nm}$ , b:  $L_{\text{eff}} = 103\text{nm}$ , c:  $L_{\text{eff}} = 233\text{nm}$ )

Là encore, on peut remarquer que pour le transistor le plus court, la survitesse intervient tout le long du canal, ce qui traduit un fonctionnement différent du régime “canal long”. Pour les autres dispositifs, les courbes de vitesse indiquent dans le canal un comportement classique, avec une vitesse avoisinant la valeur de saturation, puis une zone de survitesse.

Ce phénomène de survitesse est interprété à l'aide de la figure IV-5 qui décrit l'évolution de l'énergie moyenne des électrons, également calculée sur une profondeur de 10nm. Les électrons soumis au champ électrique horizontal voient leur vitesse augmenter sous l'action de la force électrique, et surtout, leur trajectoire est réorientée dans la direction du champ. Ces deux effets se conjuguent pour donner une importante valeur à la vitesse moyenne. L'énergie, quant à elle, augmente simplement sous l'influence de la force électrique. On retrouve là l'origine du phénomène de survitesse: la réorientation permet d'obtenir une vitesse moyenne importante, alors que l'énergie reste faible. La durée du transit étant de l'ordre du temps de libre parcours moyen, les électrons peuvent traverser le canal avant de

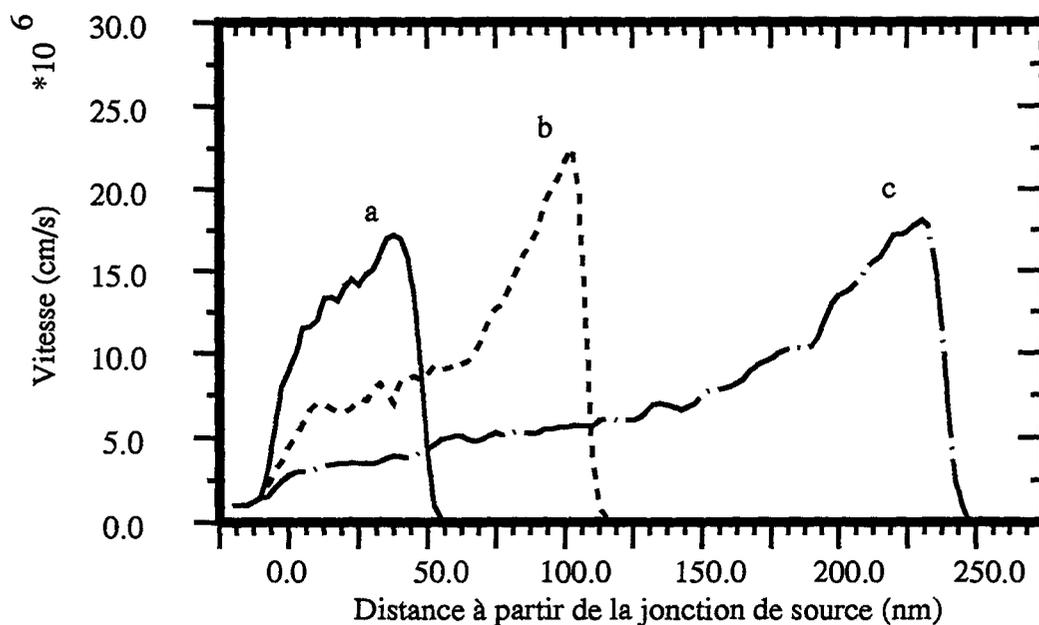
subir des interactions qui les réorientent de façon aléatoire et les ramènent en équilibre thermique<sup>1</sup> avec le réseau cristallin.



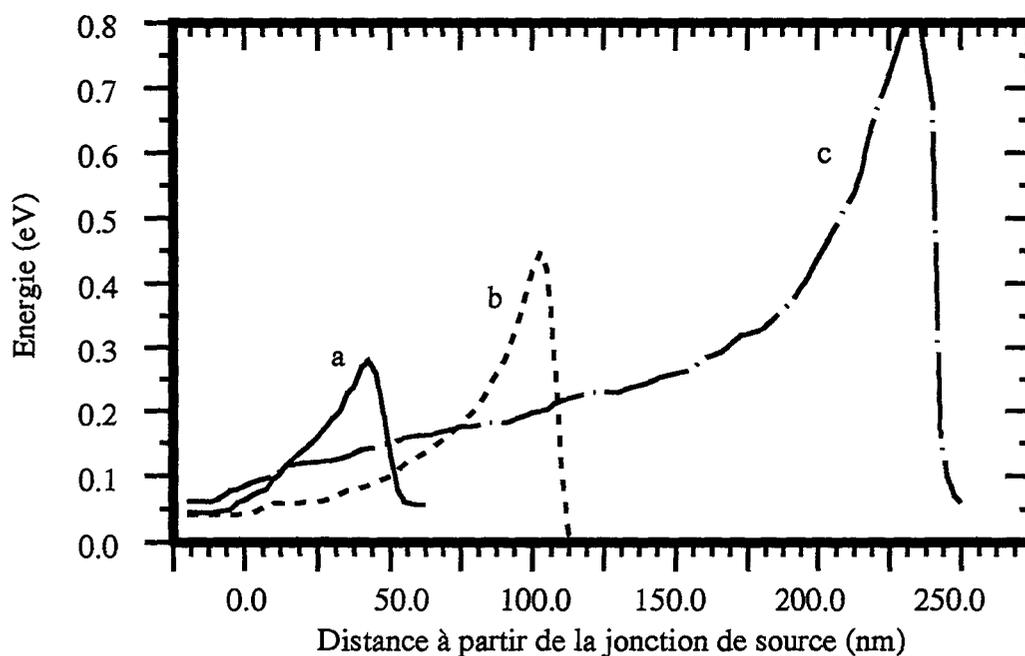
**Figure IV-5 :** Energie moyenne des électrons calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub>. (a:  $L_{\text{eff}} = 43\text{nm}$ , b:  $L_{\text{eff}} = 103\text{nm}$ , c:  $L_{\text{eff}} = 233\text{nm}$ )

La comparaison des courbes présentées précédemment avec celles publiées par Fischetti, Laux et Frank [29] et reportées ci-dessous montre que le simulateur utilisé dans le présent travail donne des résultats assez proches de ceux de Damoclès où la structure de bande est décrite plus finement à hautes énergies, par tabulation. Les différences observées, notamment sur les courbes de vitesse, peuvent provenir de la différence de description des dispositifs autant que des modèles différents inclus dans les simulateurs Monte Carlo.

1. Le terme d'équilibre thermique est utilisé par opposition au transport où le phénomène de survitesse intervient. Le transistor n'est pas en équilibre au sens strict, puisqu'il est soumis à une polarisation non nulle et est traversé par un courant.



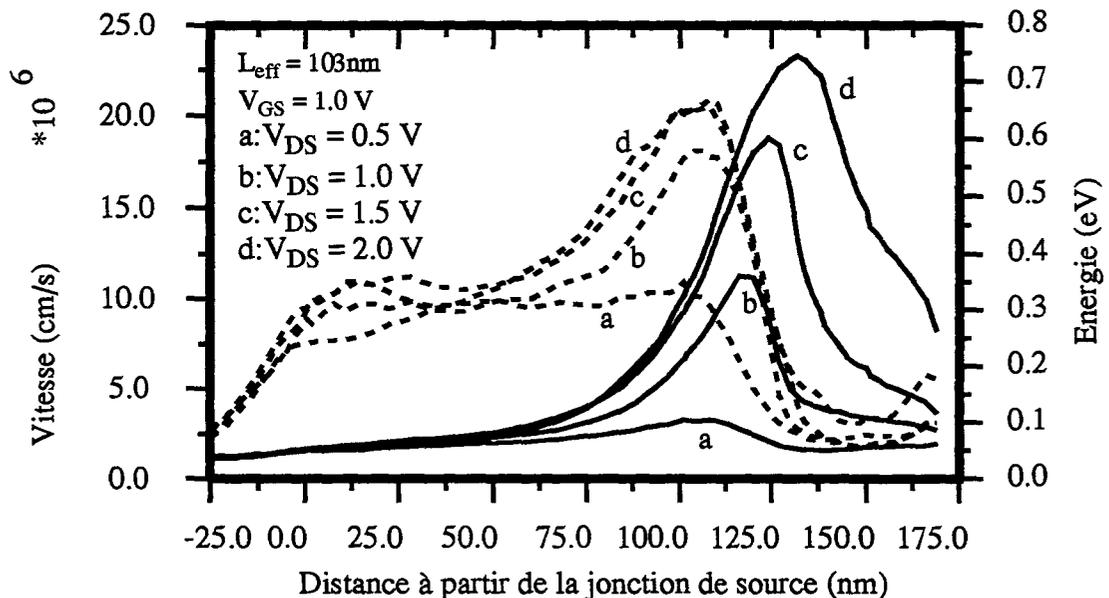
**Figure IV-6 :** Vitesse moyenne des électrons calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub>. (a:  $L_{\text{eff}} = 43\text{nm}$ , b:  $L_{\text{eff}} = 103\text{nm}$ , c:  $L_{\text{eff}} = 233\text{nm}$ ). D'après Fischetti [29].



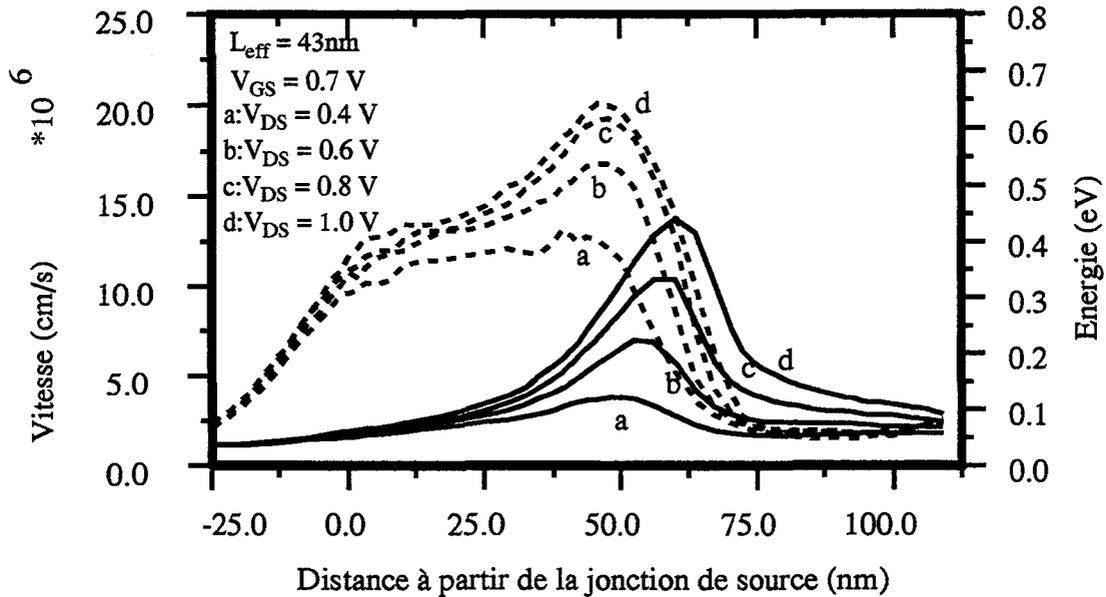
**Figure IV-7 :** Energie moyenne des électrons calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub>. (a:  $L_{\text{eff}} = 43\text{nm}$ , b:  $L_{\text{eff}} = 103\text{nm}$ , c:  $L_{\text{eff}} = 233\text{nm}$ ). D'après Fischetti [29].

#### IV-4.2 Influence de la polarisation

Les transistors de longueurs effectives de canal 43 nm et 103 nm ont été simulés à l'aide de la méthode de Monte Carlo pour différentes tensions de drain, le potentiel de grille étant maintenu constant, supérieur à la tension de seuil. Pour les faibles valeurs de  $V_{DS}$ , les courbes de vitesse et d'énergie indiquent un régime de fonctionnement sans survitesses. La vitesse moyenne est voisine de la vitesse de saturation, et l'énergie reste faible, de l'ordre de 0.1 eV. Lorsque la tension  $V_{DS}$  augmente, le phénomène de survitesses devient de plus en plus important. La vitesse dépasse la vitesse de saturation à l'approche de la jonction de drain, comme le montre la figure IV-8. Ce phénomène s'observe pratiquement dans tout le canal pour le transistor le plus court (figure IV-9).



**Figure IV-8 :** Vitesse et énergie moyennes des électrons (respectivement en pointillés et en traits pleins) calculées sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations.



**Figure IV-9 :** Vitesse et énergie moyennes des électrons (respectivement en pointillés et en traits pleins) calculées sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations.

L'énergie moyenne augmente elle aussi, et le maximum observé dans la zone de drain est de plus en plus éloigné de la jonction à mesure que la tension augmente. La courbe (d) de la figure IV-8 montre une énergie moyenne proche de 0.8 eV pour laquelle des phénomènes dus aux porteurs chauds tels que la génération de paires électrons-trous par impact, et donc l'apparition d'un courant de substrat, devraient être pris en compte. Cependant, la génération par impact et le courant de substrat qui en résulte disparaissent pour les transistors les plus courts. En effet, l'énergie moyenne maximale des électrons diminue à mesure que le transistor est miniaturisé [12], comme l'indiquait déjà la figure IV-5. Ceci est dû à la réduction des tensions de polarisation qui accompagne la diminution de la longueur du canal, dans le cadre de la miniaturisation à champ électrique maximal constant. Ce point est de toute première importance pour l'évaluation de la dégradation du transistor due aux effets de porteurs chauds, même s'il doit être relativisé par la réduction de l'épaisseur d'oxyde, qui par contre favorise l'injection de charges par effet tunnel.

## IV-5 COUPLAGE MONTE CARLO/DERIVE-DIFFUSION: APPLICATION A L'ETUDE DE TRANSISTORS NMOS FORTEMENT SUBMICRONIQUES

### IV-5.1 Extraction des résultats Monte Carlo

Dans le schéma de couplage retenu pour simuler les dispositifs MOS submicroniques, la méthode de Monte Carlo permet d'évaluer les paramètres de tension thermique et de mobilité dans les zones hors d'équilibre thermique. Ces paramètres sont ensuite réutilisés pour résoudre les équations étendues de dérive-diffusion dans l'ensemble du domaine de simulation. Or, la technique de Monte Carlo ne permet de décrire le transport électronique que dans les régions où des particules sont effectivement présentes. Il est donc nécessaire de définir la zone où les paramètres résultant d'une simulation Monte Carlo doivent être pris en compte dans les équations de dérive-diffusion, et la zone où les valeurs analytiques usuelles sont mieux adaptées. Ainsi dans les zones à l'équilibre thermique, les composantes du tenseur de tension thermique sont fixées à  $\frac{k_B T}{q} \approx 26\text{mV}$ , pour une température de 300 K. La mobilité  $y$  est considérée comme scalaire, et calculée à l'aide du modèle de Yamaguchi, tel qu'il est décrit au paragraphe II -6.2.

A partir des résultats Monte Carlo, plusieurs cas sont à considérer. Dans une partie du dispositif simulé, notamment dans la région du substrat, il n'y a aucune particule présente. Il est évident qu'il faut alors prendre la valeur à l'équilibre pour la tension thermique et une formulation analytique de la mobilité.

Dans la partie du dispositif où des particules sont effectivement présentes, les termes de tension thermique et de mobilité calculés aux noeuds du maillage comme des moyennes de grandeurs associés aux particules ne sont à prendre en compte que s'ils sont représentatifs et significatifs. Pour être représentatives, ces moyennes doivent concerner un nombre suffisant de particules. Le critère retenu pour les grandeurs ainsi évaluées en un noeud du maillage est qu'il y ait un nombre minimal de particules associée au noeud par pas de temps. De manière pratique, ce minimum est fixé à une particule par noeud et par pas de temps. Il est toutefois possible de modifier ce critère en fonction du nombre de particules et de la densité du maillage. Pour les noeuds où le nombre de particules est trop faible, les valeurs fournies par les modèles analytiques sont utilisées.

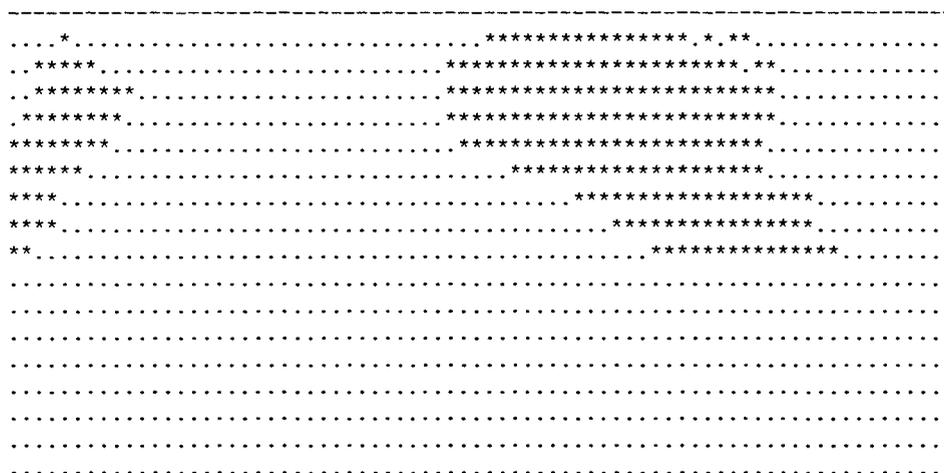
Ce critère de validation est insuffisant pour la mobilité, qui est calculée à partir du quotient de la vitesse moyenne par le taux moyen de perte du moment (paragraphe III -5.2). Or, dans les zones où la vitesse moyenne est faible, il s'agit du quotient de deux nombres petits par rapport aux valeurs usuelles, et de plus "bruités" du fait du caractère aléatoire des simulations Monte Carlo. Le quotient peut alors prendre des valeurs aberrantes, ne représentant en rien le mouvement des particules. Ces valeurs non significatives sont éliminées en fixant une vitesse moyenne minimale. Pour les simulations utilisées dans ce travail, le minimum a été fixé à  $10^6$  cm/s, sauf indication contraire.

Pour les noeuds où la vitesse moyenne dans la direction  $i$  est inférieure au seuil, deux cas se présentent: ou bien la vitesse moyenne dans l'autre direction  $j$  est elle aussi inférieure au seuil, et on considère qu'il y a équilibre thermique, ou bien un courant non négligeable circule selon  $j$ , et la valeur de mobilité dans la direction  $i$  est prise égale à la mobilité selon  $j$  (paragraphe III -5.2).

Les régions où les résultats Monte Carlo sont pris en compte sont résumées schématiquement sur les figures suivantes, dans le cas de la simulation d'un transistor de longueur effective de canal de 43 nm, pour  $V_{DS} = 1V$  et  $V_{GS} = 0.7V$ .

- \* La mobilité selon la direction du canal évaluée lors de la simulation Monte Carlo est généralement validée dans les zones de source et de drain ainsi que dans le canal, comme l'illustre la figure IV-11.
- \* La mobilité selon la direction verticale n'est validée que dans les régions où existe une composante de courant verticale (figure IV-12).
- \* Les composantes de tension thermique  $U_{xx}$  et  $U_{yy}$  sont prises en compte dès que l'un ou l'autre terme de mobilité est validé, c'est-à-dire en général dans la source, le canal et le drain.





**Figure IV-12** : Représentation schématique des zones où la valeur de mobilité  $\mu_{yy}$  prend la valeur calculée lors de la simulation Monte Carlo

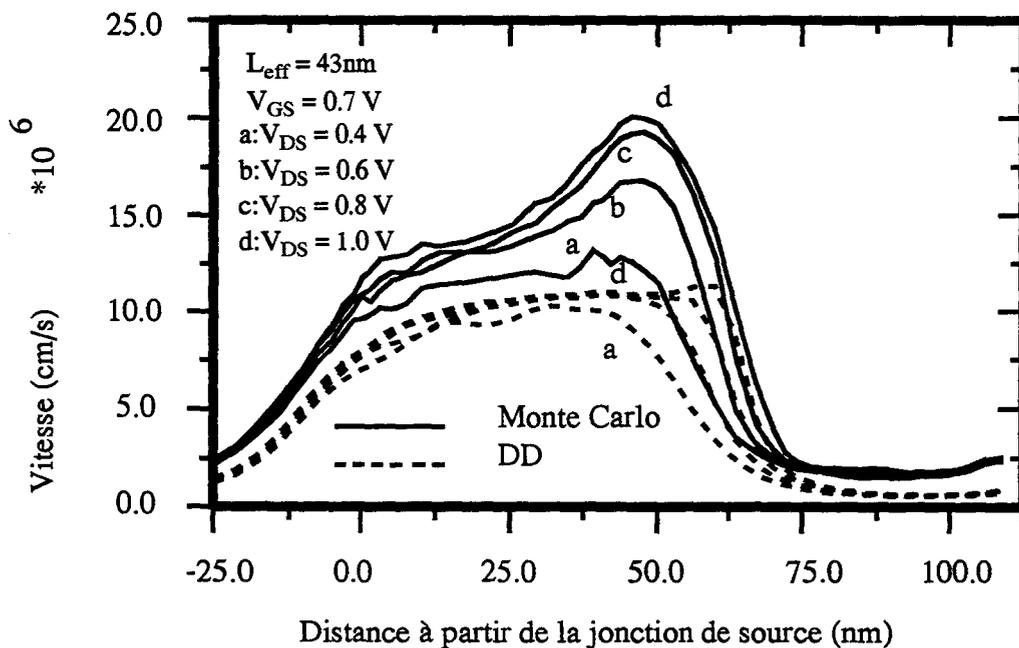
On peut noter sur les deux schémas précédents qu'aux noeuds correspondant aux électrodes ainsi qu'aux noeuds immédiatement voisins, ce sont les valeurs classiques du modèle de dérive-diffusion qui sont utilisées. Les conditions fixées pour la simulation des contacts sont en effet celles de l'équilibre thermique. Les contacts sont constitués ici de trois noeuds de maillage placés verticalement sous l'interface Si/SiO<sub>2</sub>, ce qui correspond à une troncature du domaine de simulation supprimant une partie des zones fortement dopées de source et de drain.

## IV-5.2 Interprétation dans le cadre du formalisme de dérive-diffusion étendu

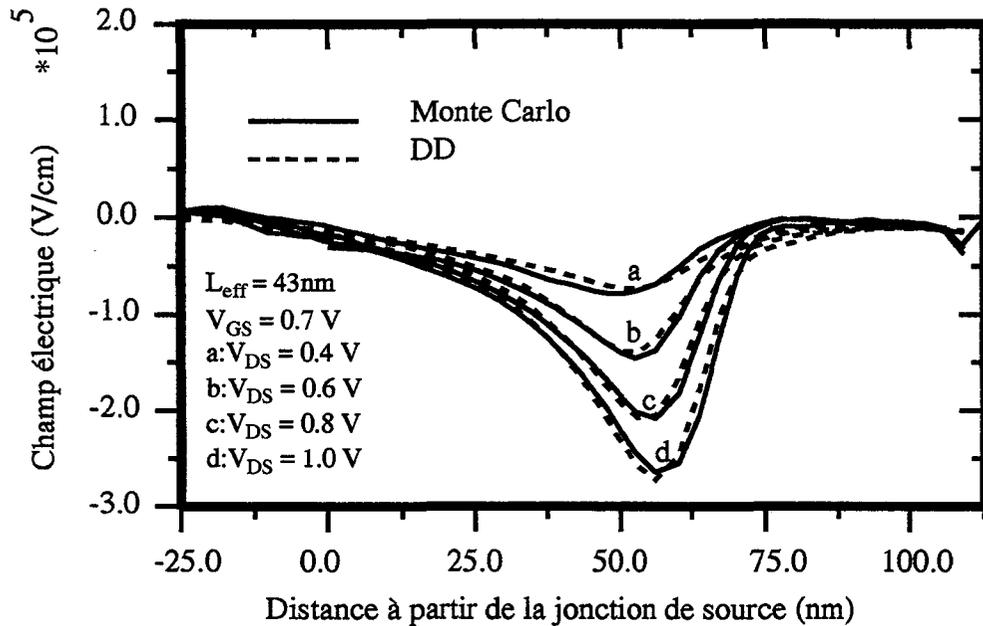
### IV-5.2.1 Vitesse des porteurs

Les phénomènes de transport où intervient la survitesse, tels qu'ils ont été décrits précédemment, ne sont pas pris en compte dans le formalisme classique de dérive-diffusion. Ainsi, l'extraction de la vitesse moyenne des électrons le long du canal à partir des densités de courant et des concentrations, lors de simulations utilisant les équations de dérive-diffusion usuelles, fournit des valeurs de l'ordre de la vitesse de saturation, environ

$10^7$  cm/s, dès que la tension  $V_{DS}$  est assez élevée. La figure IV-13 compare pour différentes polarisations la vitesse moyenne évaluée lors de simulations utilisant les méthodes de Monte Carlo et de dérive-diffusion, pour un transistor de longueur effective de canal de 43 nm. Pour la plus faible valeur de  $V_{DS}$ , 0.4 V, les vitesses moyennes fournies par les deux types de simulation sont presque semblables, ce qui indique un transport proche de l'équilibre thermique. Par contre, les vitesses calculées par les deux techniques diffèrent de façon flagrante pour les fortes valeurs de tension  $V_{DS}$ , alors que le champ électrique auquel sont soumis les électrons est sensiblement identique lors des deux simulations, comme le montre la figure IV-14. Cette carence des modèles classiques peut être corrigée par l'utilisation du formalisme étendu de dérive-diffusion qui permet d'incorporer le phénomène de survitesse, tout en conservant des équations du même type.



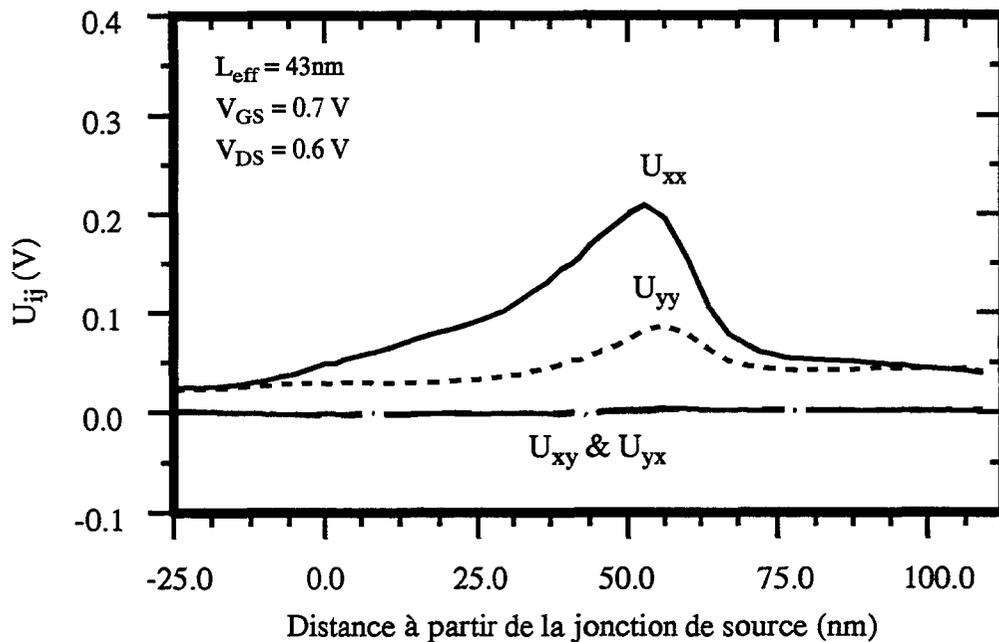
**Figure IV-13 :** Vitesse moyenne des électrons calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations



**Figure IV-14 :** Champ électrique moyen subi par les électrons, calculé sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations

#### IV-5.2.2 Tenseur de tension thermique

La résolution des équations étendues de dérive-diffusion peut se faire si les paramètres de mobilité et de tension thermique sont connus. Ceux-ci sont évalués à l'aide de moyennes calculées lors de simulations Monte Carlo. La figure IV-15 présente les différentes composantes de tension thermique susceptibles d'intervenir dans une simulation bidimensionnelle. Elles sont évaluées le long de l'interface Si/SiO<sub>2</sub>, pour le même transistor de longueur de canal  $L_{\text{eff}} = 43 \text{ nm}$ . Comme cela a été étudié au chapitre précédent, les composantes non diagonales du tenseur sont négligeables.



**Figure IV-15 :** Tenseur de tension thermique moyen, calculé sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub>

Les figures suivantes sont les profils bidimensionnels des composantes  $U_{xx}$  et  $U_{yy}$ , qui permettent de localiser la zone de transport hors d'équilibre. Conformément aux courbes de vitesse et d'énergie, c'est à l'intérieur de la zone de drain, après avoir traversé la jonction, que les électrons atteignent leur énergie maximale. Pour la composante de tension thermique  $U_{xx}$ , où  $x$  est la direction du canal, on retrouve sur la figure IV-16 le maximum situé à l'interface Si/SiO<sub>2</sub>, à l'intérieur de la zone de drain et non à la jonction. Le maximum de la composante  $U_{yy}$ , quant à lui, est éloigné de l'interface, comme l'indique la figure IV-17. Par ailleurs, la valeur des composantes  $U_{xx}$  et  $U_{yy}$  dans la zone de source est d'environ 25 mV, ce qui correspond à une situation de thermalisation. Dans la zone de drain, elle est légèrement supérieure, ce qui traduit un retour progressif à l'équilibre thermique des électrons lors des simulations.

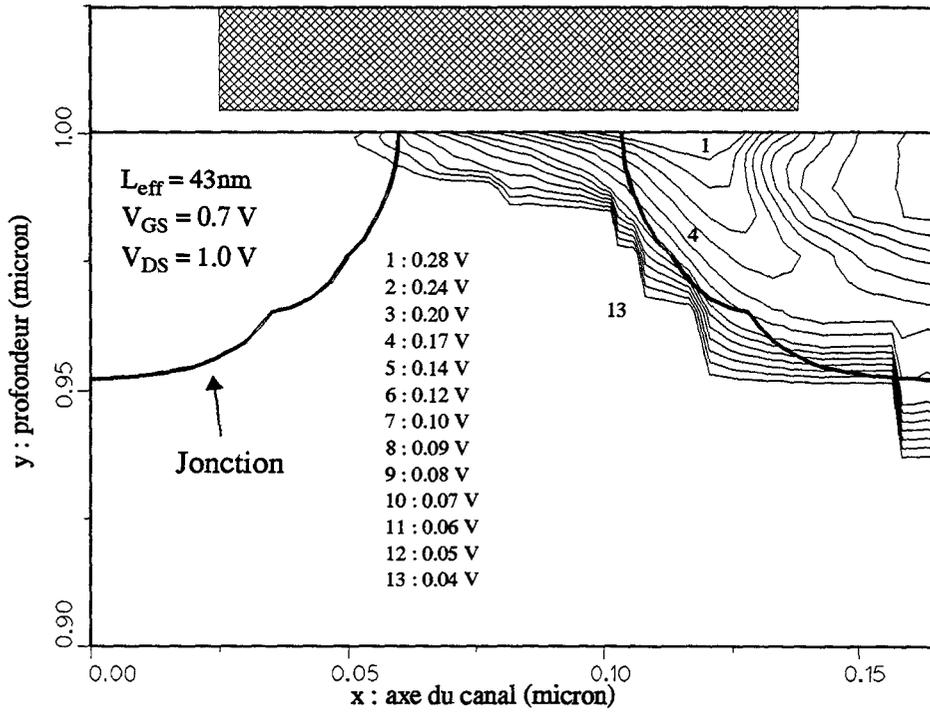


Figure IV-16 : Profil bidimensionnel de la composante  $U_{xx}$  du tenseur de tension thermique  
 (Les points anguleux des isovaleurs sont dus à l'algorithme d'interpolation)

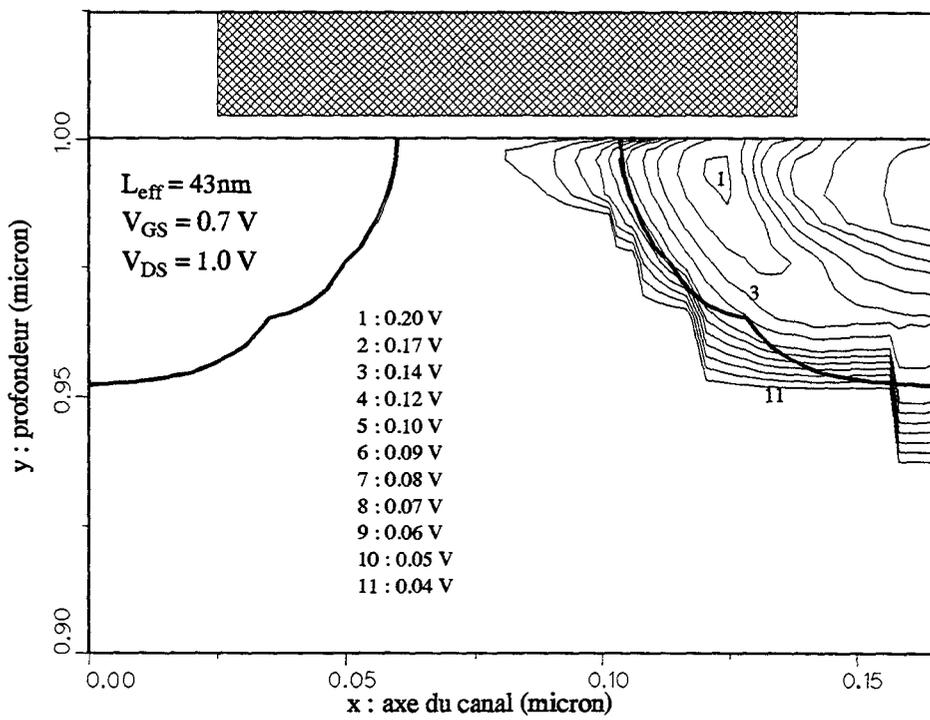
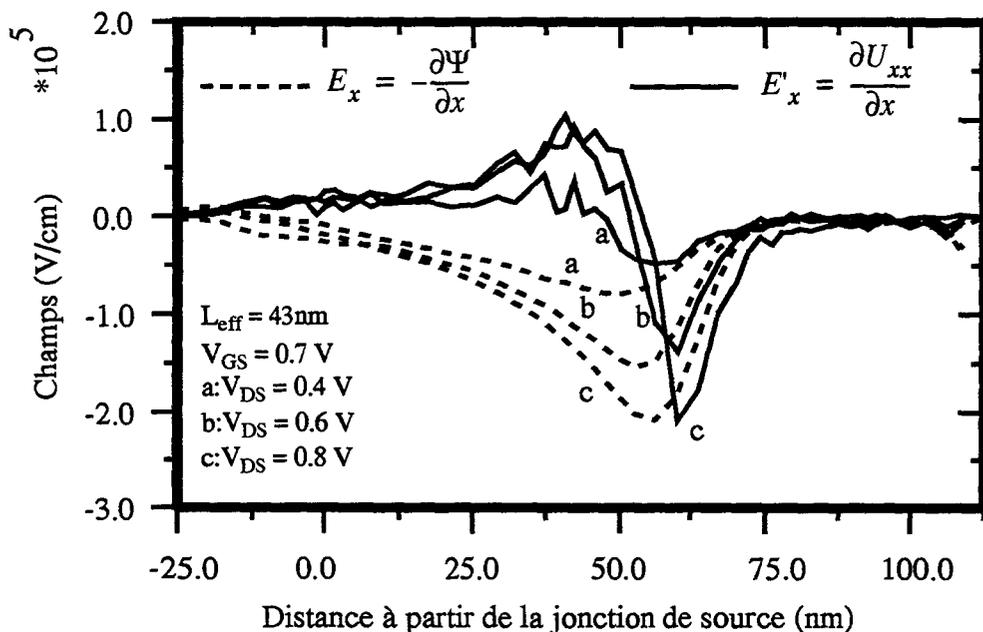


Figure IV-17 : Profil bidimensionnel de la composante  $U_{yy}$  du tenseur de tension thermique

La tension thermique intervient dans les équations étendues de dérive-diffusion de deux manières, ainsi que le rappelle l'expression de la densité de courant d'électrons suivante:

$$J_x = -qnV_x = -nq\mu_{xx}\frac{\partial\Psi}{\partial x} + nq\mu_{xx}\frac{\partial U_{xx}}{\partial x} + q\mu_{xx}U_{xx}\frac{\partial n}{\partial x} \quad (\text{IV-3})$$

Elle est présente d'une part dans le terme de diffusion, donnant le coefficient de diffusivité par la relation d'Einstein généralisée  $D_{xx} = \mu_{xx}U_{xx}$ , et d'autre part sous forme d'un champ supplémentaire  $\frac{\partial U_{xx}}{\partial x}$  auquel sont soumis les électrons. Ce champ est dénommé gradient de tension thermique. Il est aussi appelé champ de gradient d'énergie [31], mais cette appellation n'est pas rigoureuse, l'énergie et la tension thermique étant des grandeurs différentes. La figure IV-18 permet de comparer les composantes selon l'axe x du champ électrique et du champ gradient de tension thermique pour différentes polarisations.

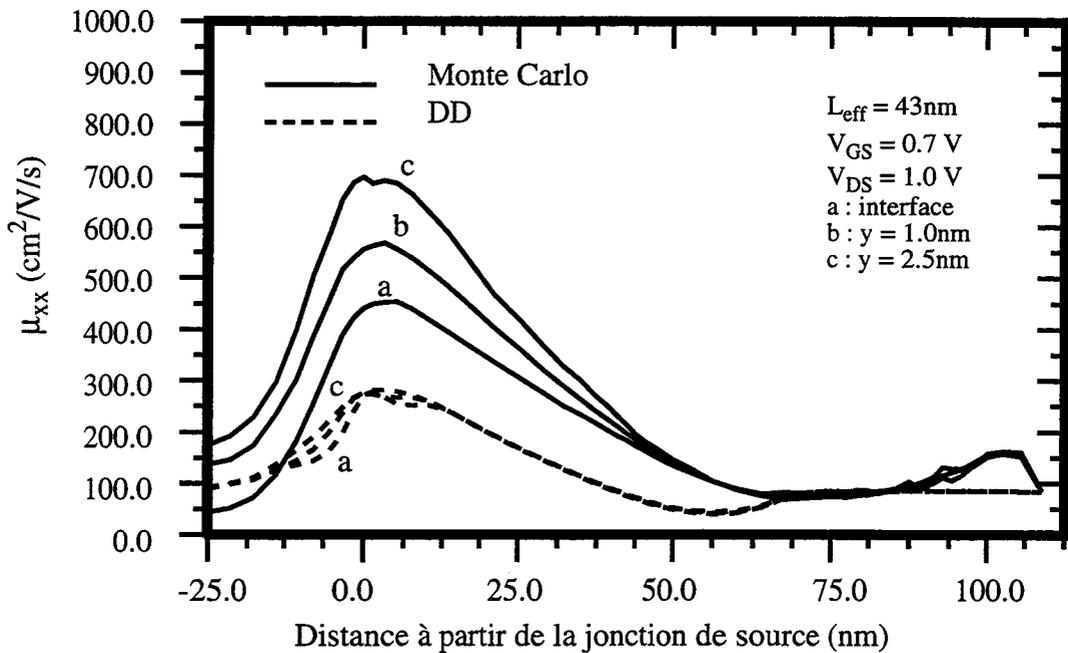


**Figure IV-18** : Moyennes du champ électrique et de la dérivée de la tension thermique subis par les électrons dans la direction du canal, calculés sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations

Dans la zone de drain, les composantes des deux champs s'ajoutent. Par contre, dans le canal, ces deux champs d'amplitudes comparables ont des effets antagonistes, ce qui devrait diminuer la vitesse des électrons par rapport à la modélisation classique de dérive-diffusion. Ce n'est cependant pas le cas, car la mobilité évaluée pour les mêmes simulations est plus grande que dans les modèles usuels, comme cela est présenté ensuite à la figure IV-19.

#### IV-5.2.3 Mobilité et diffusivité

La mobilité utilisée dans les équations de dérive-diffusion classiques est calculée à l'aide du modèle de Yamaguchi présenté au chapitre II. Par contre, lors des simulations Monte Carlo, elle est évaluée à partir de la vitesse moyenne des électrons et du taux de changement du moment. La comparaison des valeurs obtenues lors des deux types de simulation montre que la mobilité dans le canal est notablement plus grande lors des simulations Monte Carlo, ce qui correspond au comportement balistique des électrons. La forte mobilité pour le transport hors d'équilibre traduit le fait que les porteurs subissent peu de collisions. La mobilité est réduite en surface comme l'indiquent les courbes de la figure IV-19 correspondant à différentes profondeurs. Le phénomène de réduction de la mobilité par le champ électrique dans la direction du canal est également pris en compte. De ce point de vue, les mobilités des deux types de simulation ont le même comportement, même si les ordres de grandeur sont différents.

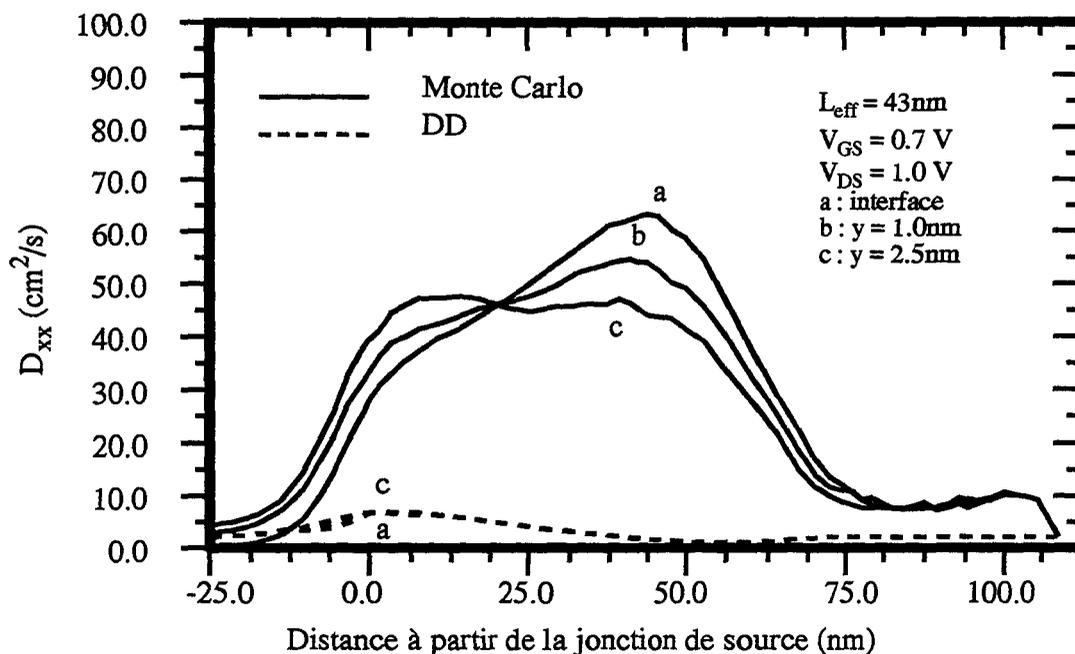


**Figure IV-19 :** Mobilité selon la direction du canal, le long de l'interface et à différentes profondeurs.

La mobilité dans la direction  $y$  n'est pas représentée ici. En effet, le courant est pour ce transistor quasiment parallèle à l'interface, au moins en surface. La valeur attribuée au terme de mobilité  $\mu_{yy}$  est donc identique à  $\mu_{xx}$ , conformément à la remarque du paragraphe III -5.2.

La figure IV-20 présente la diffusivité  $D_{xx}$  selon la direction du canal, pour différentes profondeurs. Ce terme de diffusivité est obtenu par la relation d'Einstein classique pour les simulations de dérive-diffusion, et par la relation généralisée utilisant la tension thermique  $U_{xx}$  dans le formalisme étendu:

$$D_{xx} = \mu_{xx} U_{xx} \quad (\text{IV-4})$$

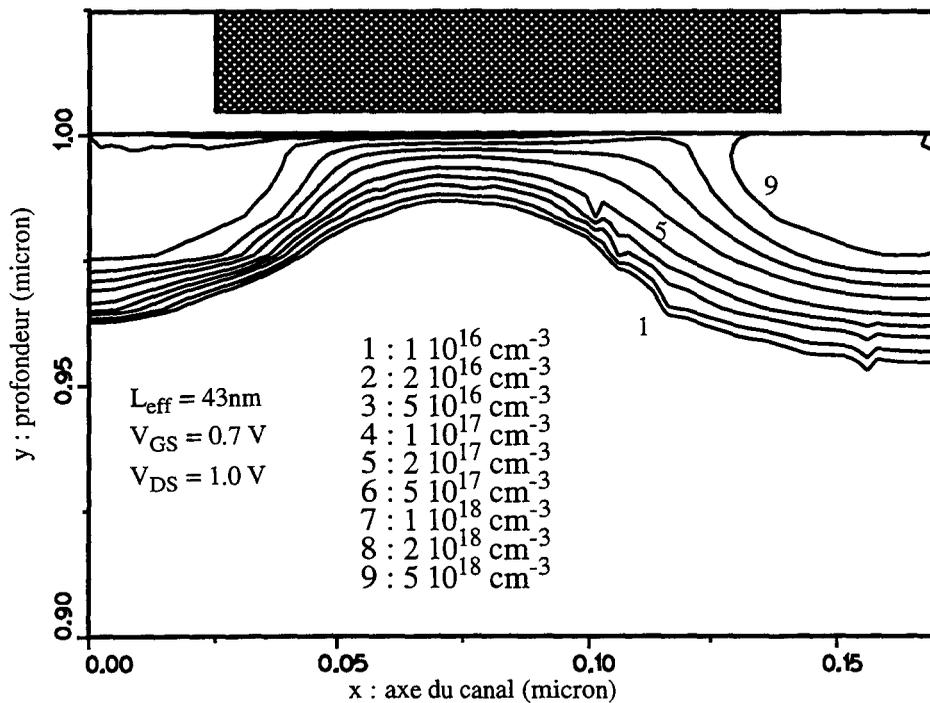


**Figure IV-20** : Diffusivité selon la direction du canal, le long de l'interface et à différentes profondeurs.

La diffusivité utilisée dans les équations de dérive-diffusion classiques est simplement proportionnelle à la mobilité, le coefficient de proportionnalité étant  $V_T = 25\text{ mV}$ . Comme cela a été noté à propos de la figure présentant les composantes de tension thermique, le transport hors d'équilibre se traduit lors des simulations Monte Carlo par des valeurs de  $U_{xx}$  dépassant largement  $V_T$  dans le canal et la zone de drain. La diffusivité, qui est évaluée en effectuant le produit de la mobilité par la tension thermique, est très fortement augmentée dans les zones hors d'équilibre, notamment en surface, où  $U_{xx}$  est maximale.

Le profil de la concentration en électrons à l'intérieur du transistor de longueur effective 43 nm est représenté sur la figure IV-21 sous forme d'isovaleurs, lors d'une simulation utilisant le formalisme étendu de dérive-diffusion. Ce profil est à comparer avec celui obtenu à l'aide des équations classiques (figure IV-22). Les isovaleurs sont pratiquement identiques dans la région de la source, à gauche sur ces représentations, mais elles diffèrent

notablement dans le canal et le drain, où elles sont plus espacées. Comme cela vient d'être remarqué, le maximum de tension thermique se traduit par une diffusivité très fortement augmentée par rapport au modèle de dérive-diffusion classique, ce qui entraîne un étalement plus important des électrons: le canal est plus profond que lors de la simulation classique. Un deuxième élément permettant d'interpréter cette différence de profil de concentration est le gradient de la tension thermique, qui s'oppose au champ électrique vertical et vient en atténuer les effets.



**Figure IV-21** : Profil bidimensionnel de la concentration en électrons obtenu par une simulation utilisant le formalisme étendu de dérive-diffusion.

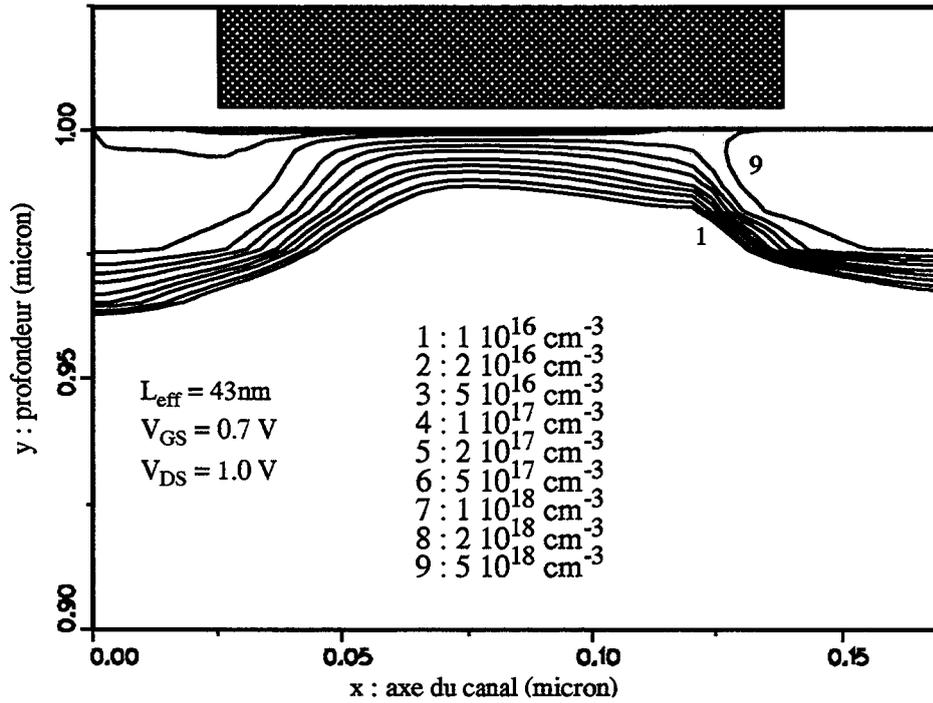


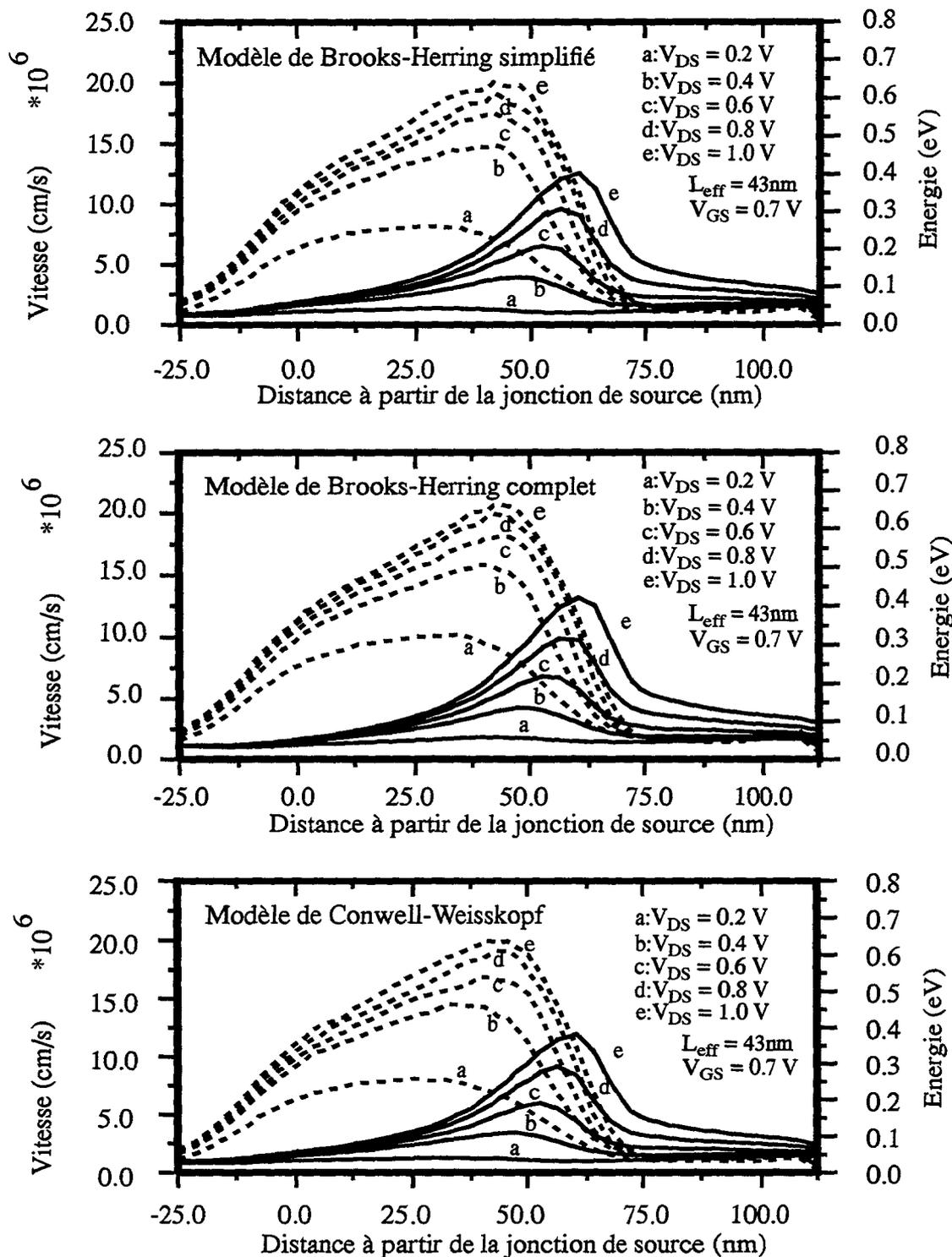
Figure IV-22 : Profil bidimensionnel de la concentration en électrons obtenu par une simulation classique de dérive-diffusion.

### IV-5.3 Influence du modèle d'interactions sur impuretés ionisées

Un des éléments les plus critiques pour les simulations Monte Carlo est l'évaluation de la probabilité des collisions sur impuretés ionisées, notamment à forte concentration. Ce point est important car ce type d'interaction est l'un des plus probable, voire le plus probable dans la plupart des cas. C'est donc la façon de modéliser ces collisions qui détermine en grande partie la validité des résultats de simulation Monte Carlo. Une première approche utilise la formulation de Brooks-Herring simplifiée, indépendante des concentrations, donnée par l'équation (II-49). Cette expression présente l'avantage d'éviter le calcul des concentrations en impuretés et en électrons aux positions occupées par chaque particule simulée, et donc de réduire un peu le temps de simulation. La modélisation complète de Brooks Herring (II-45) et celle de Conwell-Weisskopf (II-46) constituent les deux autres approches retenues. Ces trois façons de modéliser les interactions sur impuretés ionisées sont comparées ci-dessous dans le cas de simulations de transistors nMOS. Le traitement de ce type de collisions influence en effet la simulation du transport électronique à l'intérieur des dispositifs et l'estimation de leur comportement électrique général.

#### IV-5.3.1 Impact sur les paramètres de transport

Le dispositif de longueur effective 43 nm a été simulé à l'aide des trois modèles dans les conditions de polarisation suivantes:  $V_{GS} = 0.7$  V,  $V_{DS}$  variant de 0.2 à 1.0 V. Les trois jeux de courbes de la figure IV-23 présentent la vitesse et à l'énergie moyennes des électrons. Les résultats sont sensiblement les mêmes quel que soit le modèle utilisé. La seule différence notable concerne la vitesse à faible tension source drain, pour laquelle le modèle complet de Brooks-Herring donne des valeurs plus grandes.



**Figure IV-23 :** Vitesse et énergie moyennes des électrons (respectivement en pointillés et en traits pleins) calculées sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations.

Les courbes de la figure IV-24 représentent la mobilité évaluée le long du transistor à une profondeur de 1 nm, c'est-à-dire juste sous l'interface Si/SiO<sub>2</sub>. Pour chacun des trois modèles, elles donnent l'influence de la tension de drain. La comparaison des deux premières séries de courbes montre que lorsque les concentrations en électrons et en impuretés ionisées sont considérées comme égales, approximation qui est utilisée pour le modèle de Brooks-Herring simplifié, la mobilité est largement sous-estimée dans le canal: la concentration en électrons étant en fait plus élevée que celle en dopants dans la zone d'inversion, l'effet d'écrantage est sous-estimé.

L'étude de la mobilité à champ et concentration uniformes avait déjà permis de constater que le modèle de Brooks-Herring complet surestimait la mobilité par rapport aux autres formulations et par rapport aux valeurs expérimentales, à fortes concentrations. Cette lacune se retrouve lors de la simulation de dispositifs, comme le montre la comparaison du deuxième jeu de courbes de la figure IV-24 avec les résultats obtenus à l'aide de la formulation simplifiée de Brooks-Herring et du modèle de Conwell-Weisskopf.

Le modèle de Conwell-Weisskopf, qui fournit les résultats présentés dans la troisième partie de la figure, semble le plus cohérent du point de vue de l'évaluation de la mobilité des électrons, à la fois dans le canal et dans les zones de source et de drain à forte concentration.

Les différences d'évaluation de la mobilité à l'intérieur des dispositifs se traduisent également au niveau des courants de source et de drain, auxquels est consacré le paragraphe suivant.

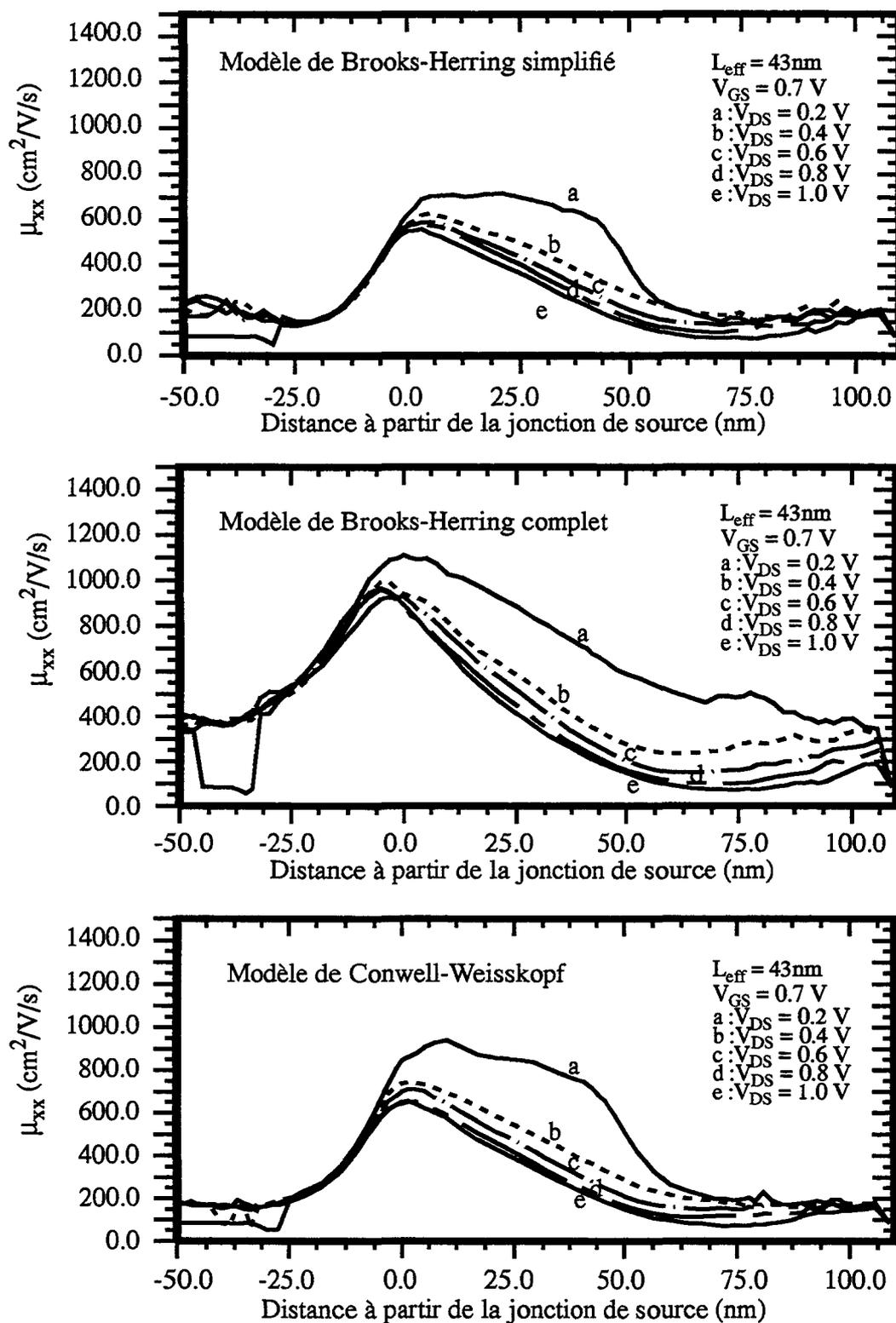


Figure IV-24 : Mobilité selon la direction du canal, calculée sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations.

### IV-5.3.2 Dépendance des caractéristiques électriques

Un des objectifs de la simulation de dispositif, outre une compréhension détaillée du comportement électrique des composants, est d'aider à l'optimisation en étudiant entre autres l'influence des paramètres technologiques qui déterminent la géométrie et les profils de dopage. La validation des outils de simulations peut se faire par comparaison de grandeurs calculées et mesurées telles que les caractéristiques  $I_D(V_{DS})$  et la transconductance. L'objet de ce paragraphe est d'étudier l'impact des modèles d'interactions avec les impuretés ionisées sur de telles caractéristiques électriques.

Les phénomènes de génération-recombinaison par impact ne sont pas pris en compte lors des simulations concernant le transistor de longueur de canal  $L_{eff} = 43$  nm. Leur occurrence est en effet peu importante compte tenu des faibles polarisations utilisées pour ce type de transistor. Un électron ne peut en effet gagner suffisamment d'énergie pour entraîner la création d'une paire électron-trou. Le courant de substrat est donc nul pour les dispositifs considérés. D'autre part, les conditions imposées aux électrodes lors des simulations Monte Carlo garantissent la conservation de la quantité de charge totale, ce qui se traduit par l'égalité des courants de source et de drain, en valeur absolue (cf. paragraphe III - 4.4).

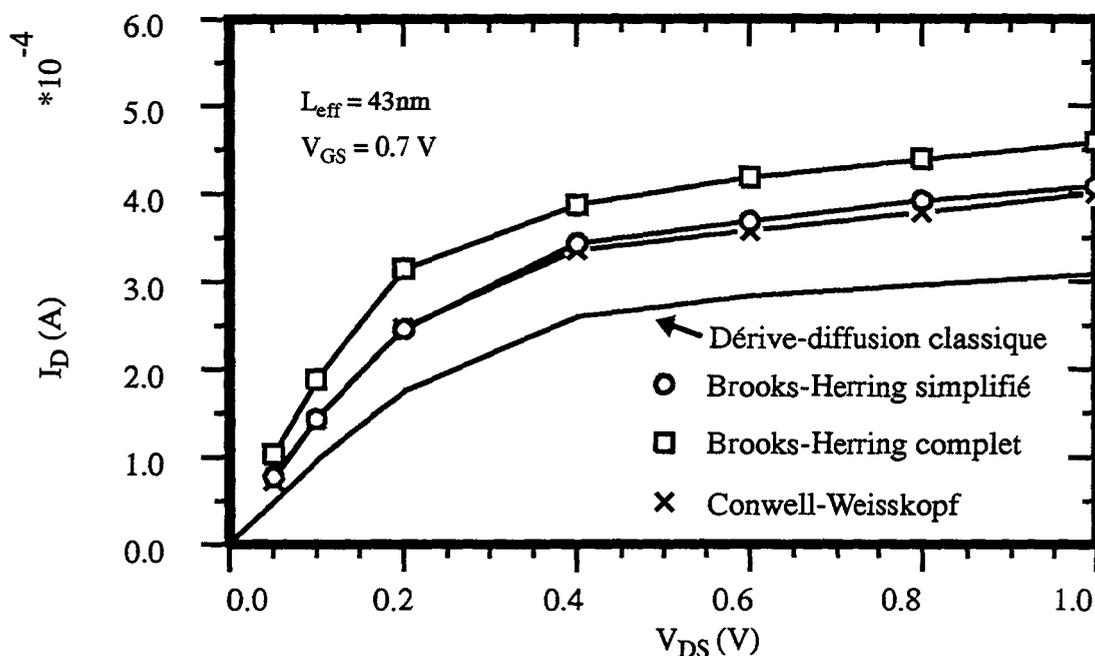
L'évaluation du courant source drain peut alors se faire de deux façons lorsque la méthode de Monte Carlo est utilisée: comptabiliser les particules injectées dans la source ou sortant du drain par unité de temps, ou bien calculer le courant moyen le long du transistor de la façon suivante:

$$I(x) = W \int_0^{y_{max}} J_x(x, y) \cdot dy \quad (IV-5)$$

où  $W$  est la largeur du transistor, prise ici égale à 1 micron,  $n$  la concentration en électrons,  $J_x$  la densité de courant dans la direction du canal et  $y$  la direction verticale. Ces deux méthodes donnent des résultats tout à fait cohérents, à quelques pour cents près, quand la durée de la simulation permet d'atteindre un régime stationnaire.

La figure IV-25 donne le courant en fonction de la tension  $V_{DS}$  pour une tension

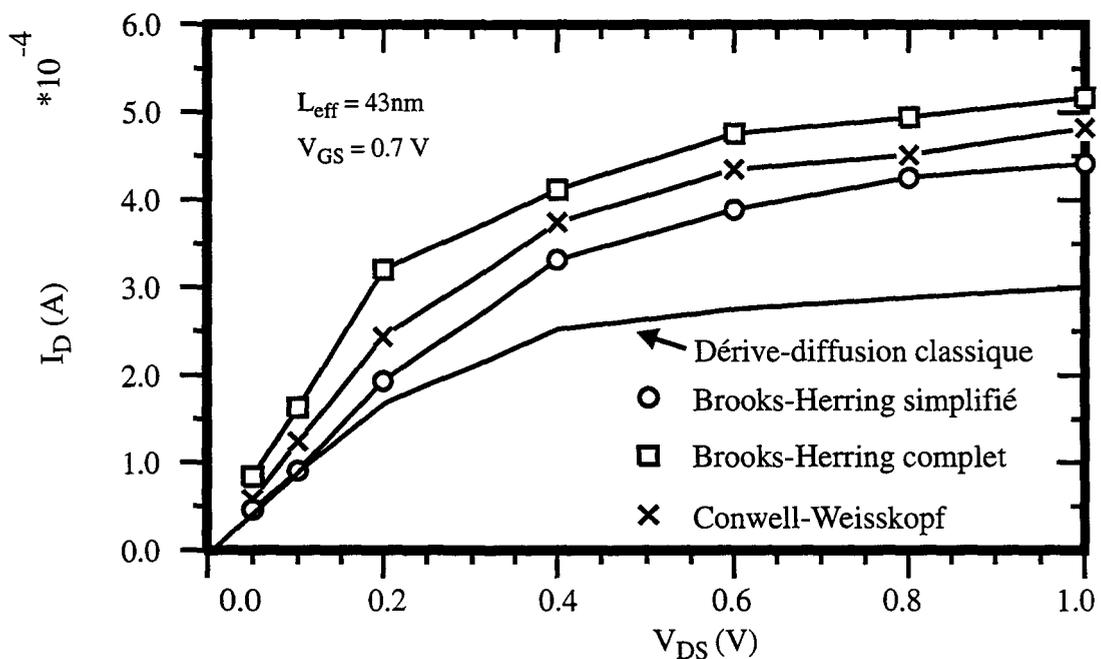
$V_{GS}$  fixée à 0.7V, tel qu'il est extrait de simulations Monte Carlo. De façon tout à fait cohérente avec les remarques précédentes, le courant est plus important lorsque le modèle complet de Brooks-Herring est utilisé, la mobilité étant surestimée, notamment dans les zones de source et de drain. Plus généralement, le courant est plus important que ce qui est calculé à l'aide du formalisme classique de dérive-diffusion. On retrouve là l'influence du phénomène de survitesse, qui provoque une augmentation du courant que les simulations classiques ne prennent pas en compte.



**Figure IV-25** : Courant calculé lors de simulations Monte Carlo utilisant différents modèles de collision sur impuretés ionisées. La courbe obtenue par une simulation utilisant le formalisme classique de dérive-diffusion est donnée pour comparaison.

Les résultats des simulations Monte Carlo ont été utilisés pour simuler le même dispositif en utilisant le formalisme de dérive-diffusion étendu. Les courbes de courant en fonction de la tension entre le drain et la source sont représentées à la figure IV-26. Là encore, les résultats sont cohérents avec les remarques faites sur les modèles de collisions sur impuretés ionisées sous-jacents. Le modèle complet de Brooks-Herring surestime la mobilité dans les zones fortement dopées, alors que sa version simplifiée la sous-estime

dans le canal. Les différentes estimations conduisent à des valeurs de courant plus élevées que le modèle de dérive-diffusion classique à forte tension de polarisation. A faible tension  $V_{DS}$ , la zone de transport hors d'équilibre diminue, les critères d'utilisation des résultats Monte Carlo sont tels que le modèle de dérive-diffusion classique est utilisé dans la plus grande partie du dispositif, et il n'est donc pas surprenant que les courbes se rapprochent de celles obtenues à l'aide de simulations de dérive-diffusion usuelles.

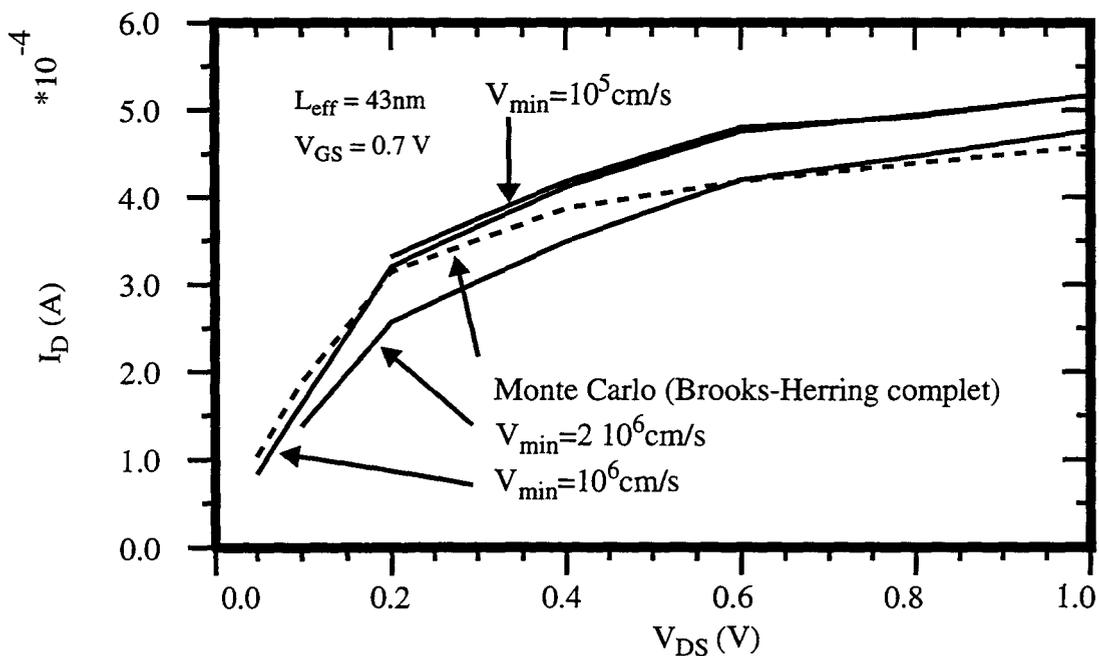


**Figure IV-26 :** Courant calculé lors de simulations utilisant le formalisme de dérive-diffusion étendu.

#### IV-5.4 Convergence du couplage Monte Carlo/dérive-diffusion

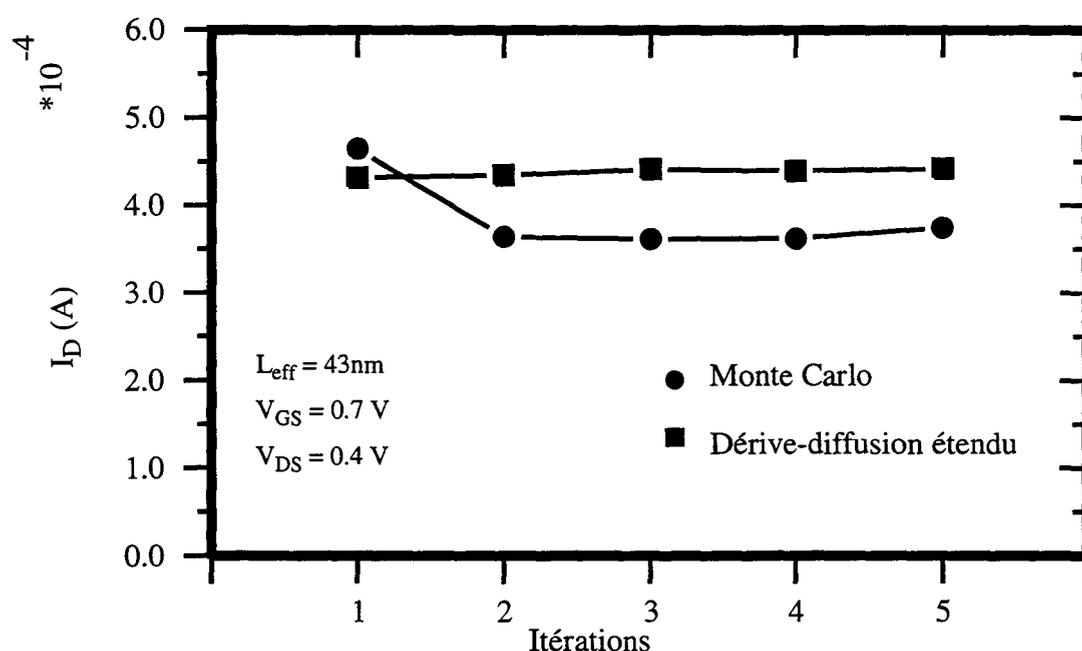
La cohérence entre les résultats Monte Carlo et ceux fournis par le formalisme de dérive-diffusion étendu est influencée par les critères de validation des valeurs de mobilité et de tension thermique extraites des simulations Monte Carlo, comme le montre la figure IV-27. Le nombre de particules minimal par noeud est fixé à une particule par pas de temps. Les termes de mobilité sont validés en utilisant le même seuil pour les composantes

de la vitesse moyenne. Diminuer ce seuil en dessous de  $10^6 \text{cm/s}$  n'a pratiquement aucune influence sur le courant calculé: les résultats Monte Carlo sont validés dans l'ensemble de la source, du drain et du canal, comme indiqué au paragraphe IV -5.1. Augmenter le seuil à  $2 \cdot 10^6 \text{cm/s}$  permet d'obtenir des résultats proches du courant obtenu lors des simulations Monte Carlo à forte tension  $V_{DS}$ . La zone où les valeurs Monte Carlo sont prises en compte est alors restreinte à la région hors d'équilibre du canal et du drain. Ce critère est cependant trop restrictif à faible tension  $V_{DS}$ , lorsque la vitesse est plus faible, et l'intensité du courant source drain est sous-estimée. Afin d'étudier la cohérence entre les simulations Monte Carlo et celles qui utilisent le formalisme de dérive-diffusion étendu, il faudrait examiner plus en détail la continuité des termes de mobilité et de tension thermique entre les régions à l'équilibre et celles qui sont hors d'équilibre.



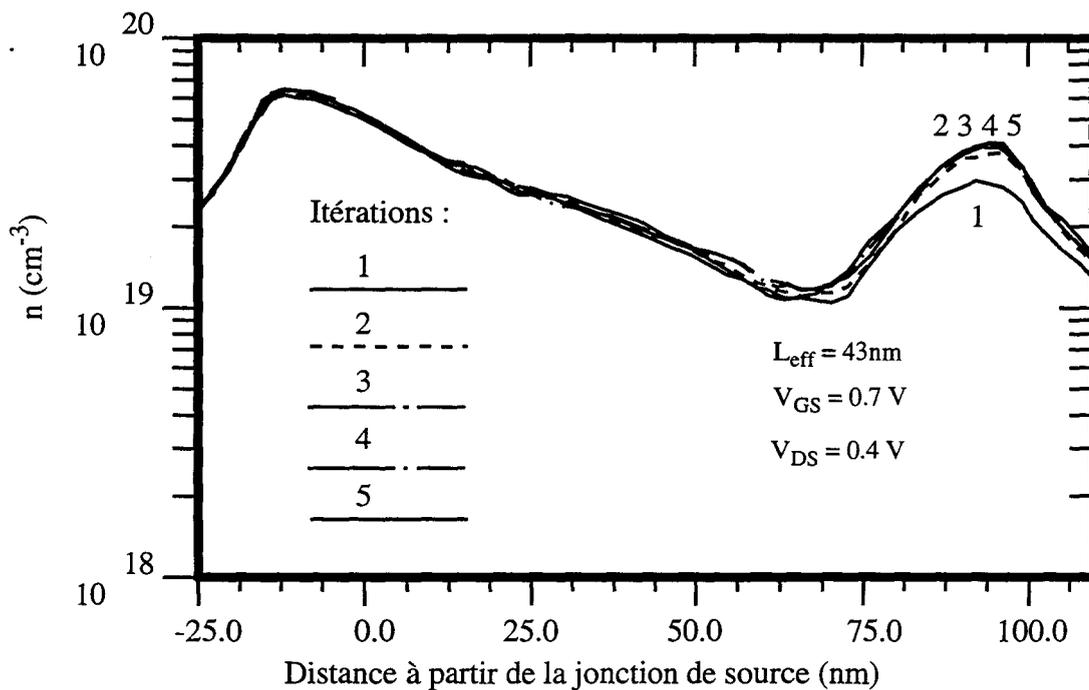
**Figure IV-27** : Courant calculé lors de simulations utilisant le formalisme de dérive-diffusion étendu, avec pour critère de validation différentes valeurs minimales de vitesse moyenne.

Un avantage important du couplage entre les techniques de Monte Carlo et de dérive-diffusion étendue est que la résolution des équations de dérive-diffusion permet une mise à jour rapide des concentrations en porteur, alors qu'une simulation Monte Carlo demande une durée de simulation de plusieurs picosecondes pour atteindre une situation stable [32] [33]. Pour permettre de réduire la durée des simulation Monte Carlo, les informations concernant les particules (position, vitesse et énergie) sont conservées pendant la résolution des équations étendues de dérive-diffusion. Les concentrations sont corrigées à l'itération suivante en supprimant ou en dupliquant des particules dans chaque maille du domaine de simulation. La convergence rapide de cette méthode est illustrée par la figure suivante, où une simulation Monte Carlo de 15000 particules pendant 1 picoseconde est réalisée à chaque itération.



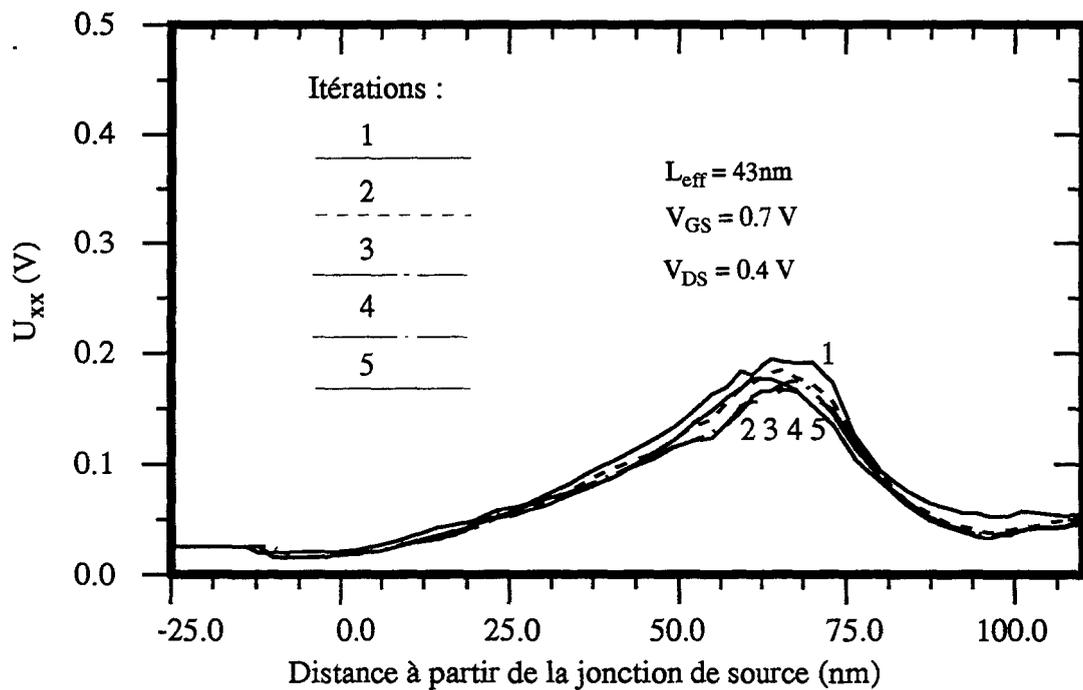
**Figure IV-28** : Courant calculé lors des cinq premières itérations Monte Carlo-formalisme de dérive-diffusion étendu. (Modèle de Conwell-Weisskopf pour les interactions sur impuretés ionisées)

La valeur du courant obtenu par la simulation Monte Carlo lors de la première itération est plus élevée que les suivantes. Il s'agit d'un effet transitoire: la simulation débute avec des particules en équilibre thermique, ce qui ne reflète pas le comportement des électrons dans le transistor à la polarisation considérée. Les résultats sont rapidement stables lors des itérations suivantes, le courant obtenu par les équations de dérive-diffusion restant pratiquement à la même valeur qu'à la première itération. Cette convergence rapide est due à la mise à jour des concentrations par la résolution des équations de dérive-diffusion et au fait que les paramètres utilisés pour le couplage, les termes de tension thermique et de mobilité, dépendent peu des concentrations en électrons et en trous. Ils sont faiblement modifiés suite aux corrections apportées aux concentrations, comme l'illustrent les courbes suivantes.

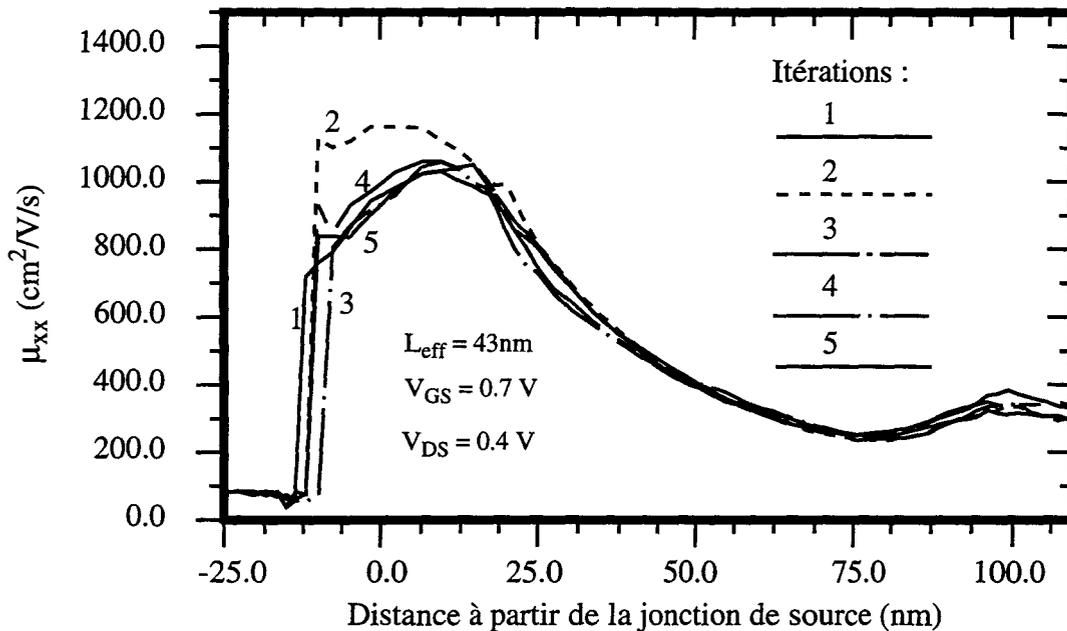


**Figure IV-29 :** Concentration en électrons le long de l'interface Si/SiO<sub>2</sub> calculée lors des cinq premières itérations Monte Carlo-formalisme de dérive-diffusion étendu.

La concentration en électrons évolue dans la région du drain, zone hors d'équilibre thermique. La résolution des équations de dérive-diffusion permet de trouver les valeurs de concentration en régime permanent plus rapidement que la seule méthode de Monte Carlo pour laquelle plusieurs picosecondes de simulation seraient nécessaires. De ce fait, le profil de concentration le long de l'interface ne varie plus notablement lors des itérations suivantes. La figure IV-30 indique l'évolution de la composante  $U_{xx}$  de la tension thermique le long de l'interface Si/SiO<sub>2</sub> lors des itérations. De la même façon, le terme de mobilité  $\mu_{xx}$  est présenté à la figure IV-31. Ces paramètres nécessaires au couplage sont stables dès la troisième itération. Il faut noter que l'on ne peut pas réduire la durée de la première simulation Monte Carlo en dessous de une picoseconde, qui est l'ordre de grandeur du temps de relaxation en énergie, et donc également celui de la période transitoire liée aux conditions initiales de simulation. Une fois passé ce délai, la convergence de la méthode de couplage est rapide.



**Figure IV-30 :** Composante de tension thermique  $U_{xx}$  le long de l'interface Si/SiO<sub>2</sub> calculée lors des cinq premières itérations Monte Carlo-formalisme de dérive-diffusion étendu.



**Figure IV-31 :** Composante de mobilité  $\mu_{xx}$  le long de l'interface Si/SiO<sub>2</sub> calculée lors des cinq premières itérations Monte Carlo-formalisme de dérive-diffusion étendu.

Les deux figures suivantes, IV-32 et IV-33, présentent le courant en fonction de la tension drain source obtenu lors de trois itérations, avec des modèles différents pour les collisions sur impuretés ionisées. Comme cela a déjà été précisé, le modèle choisi pour les interactions sur les impuretés ionisées a une influence sur la mobilité des électrons et donc sur le courant simulé. Les deux modèles permettent cependant une convergence rapide pour l'estimation du courant source-drain, et ce pour toutes les valeurs de  $V_{\text{DS}}$  considérées. En effet, pour ce qui concerne les simulations utilisant le formalisme étendu de dérive-diffusion, la deuxième et la troisième itération donnent des intensités de courant source drain quasiment identiques alors que des écarts sont constatés entre les deux premières itérations, dus à la mise à jour du profil de concentration en électrons.

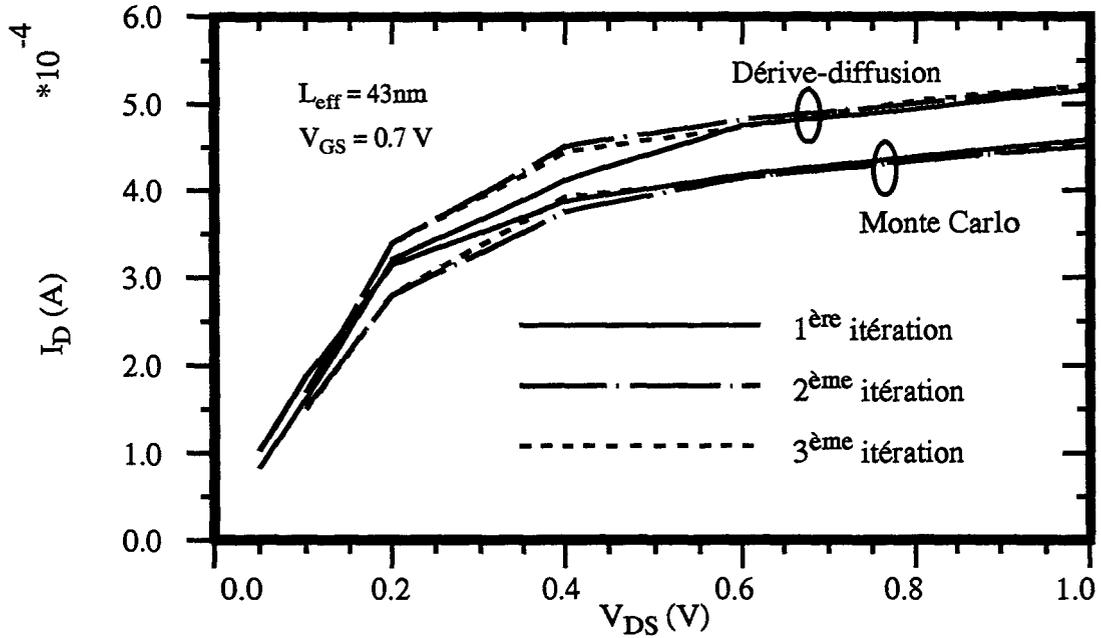


Figure IV-32 : Courant calculé lors des trois premières itérations Monte Carlo-formalisme de dérive-diffusion étendu. (Modèle de Brooks Herring complet pour les interactions sur impuretés ionisées)

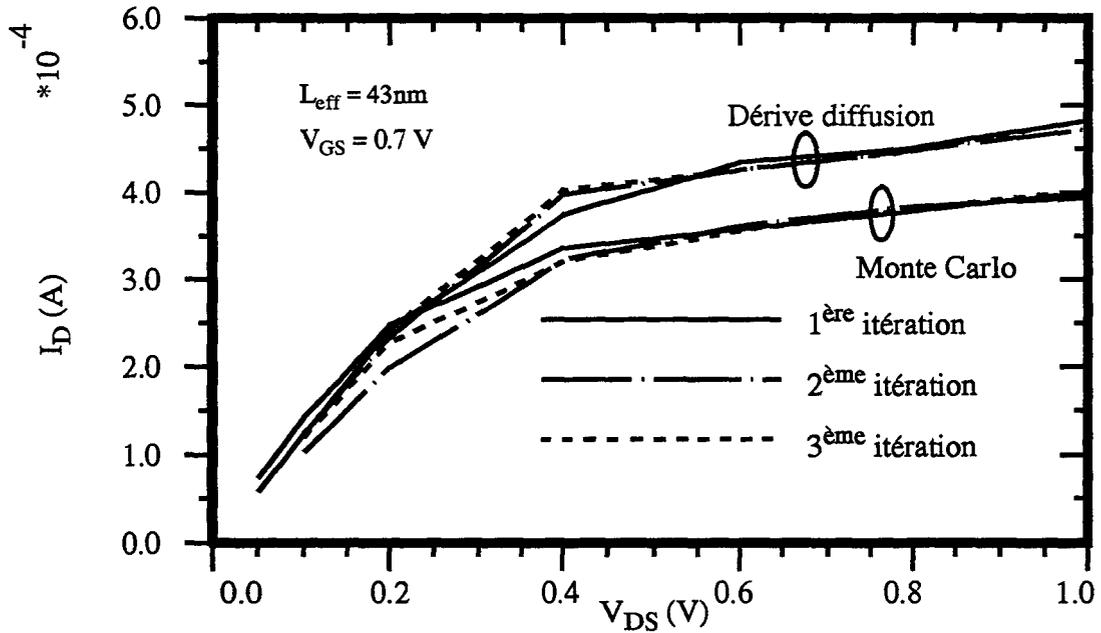


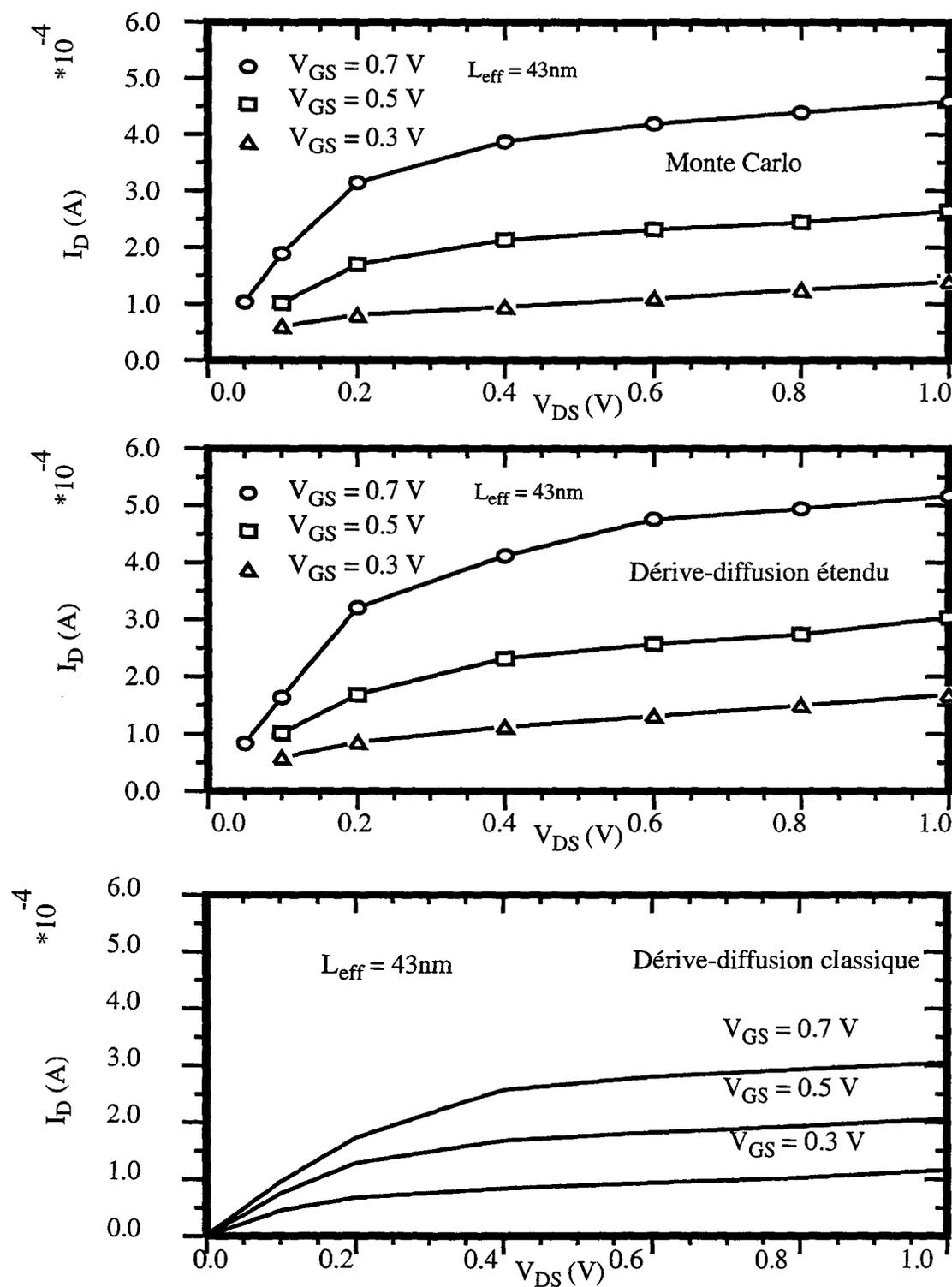
Figure IV-33 : Courant calculé lors des trois premières itérations Monte Carlo-formalisme de dérive-diffusion étendu. (Modèle de Conwell-Weisskopf pour les interactions sur impuretés ionisées)

#### IV-5.5 Caractéristiques électriques obtenues par la méthode de couplage

Le couplage entre les simulations Monte Carlo et les équations étendues de dérive-diffusion permet de simuler les phénomènes de surtension lorsqu'ils interviennent dans le fonctionnement du dispositif. A faible tension  $V_{DS}$ , l'utilisation du formalisme classique de dérive-diffusion fournit des informations suffisamment précises. A plus forte tension  $V_{DS}$ , la méthode de couplage donne de bons résultats, à condition que le courant soit assez élevé, ce qui est le cas pour les transistors passants. La méthode de Monte Carlo semblerait également la mieux adaptée pour simuler le comportement sous le seuil à fort  $V_{DS}$  puisque le champ électrique longitudinal élevé introduit des effets balistiques. Le problème est que pour de tels dispositifs, le nombre de particules, qui s'élève à quelques dizaines de milliers, est du même ordre de grandeur que le nombre d'électrons en bande de conduction. Un courant de  $1.6 \mu\text{A}$  correspond à une dizaine d'électrons, et donc de particules, traversant le dispositif lors d'une simulation de une picoseconde. Or, les résultats obtenus par la technique de Monte Carlo sont d'autant plus précis que le nombre de particules mis en jeu est grand. Même en se contentant d'une précision médiocre, il n'est donc pas possible de simuler les caractéristiques à très faible courant par cette méthode. Par contre, pour un dispositif passant et donc pour des courants non négligeables, la méthode de couplage permet d'obtenir des informations tout à fait cohérentes sur le comportement électrique global, ainsi que le montrent les résultats présentés dans ce paragraphe.

#### IV-5.5.1 Réseaux de caractéristiques $I_D(V_{DS})$

Le couplage entre les simulations Monte Carlo et les équations étendues de dérive-diffusion permet le calcul du réseau de caractéristiques  $I_D(V_{DS})$  d'un dispositif. La première partie de la figure IV-34 est le résultat de simulations Monte Carlo du dispositif de longueur effective de canal 43nm alors que la partie suivante présente le réseau de caractéristiques calculé à l'aide du formalisme étendu de dérive-diffusion. La mobilité et la tension thermique nécessaires à la résolution des équations étendues de dérive-diffusion sont évalués lors des simulations Monte Carlo. Il est à noter que les courants obtenus à  $V_{GS}$  élevé sont notablement supérieurs à ceux calculés par le modèle classique de dérive-diffusion, comme l'indique la comparaison avec le troisième réseau de courbes. Ceci est dû au transport hors d'équilibre caractérisé par le phénomène de survitesses. L'écart est d'autant plus grand que les tensions de polarisation sont élevées. Pour  $V_{GS} = 0.3V$ , c'est-à-dire pour des conditions où l'influence de la survitesses est faible, les courants issus des simulations Monte Carlo et de dérive-diffusion classique sont pratiquement identiques. Ces résultats montrent que le couplage de la méthode de Monte Carlo et du formalisme de dérive-diffusion étendu fournit un outil de simulation de dispositif efficace permettant la prise en compte de phénomènes de transport hors d'équilibre.



**Figure IV-34 :** Caractéristiques  $I_D(V_{DS})$  obtenues par la méthode de Monte Carlo, par le formalisme de dérive-diffusion étendu et par le formalisme de dérive-diffusion classique.

#### IV-5.5.2 Transconductance

La transconductance a été calculée d'après les simulations Monte Carlo et de dérive-diffusion classique, à  $V_{DS} = 1$  V et à  $V_{GS} = 0.7$  V, pour quatre longueurs de canal (figure IV-35). La transconductance obtenue par les simulations Monte Carlo est visiblement plus élevée que ce que la méthode de dérive-diffusion permet de calculer, ce qui constitue une illustration supplémentaire du phénomène de survitesse [35]. L'estimation de la transconductance à partir des équations de transport classiques ne peut dépasser  $C_{Ox}v_{Sat}$  [3] [15], ce qui pour un transistor dont l'épaisseur d'oxyde est 45 Å est inférieur à 770  $\mu\text{S}/\mu\text{m}$ , valeur calculée en prenant la vitesse de saturation égale à  $10^7$  cm/s. Or cette valeur de transconductance est nettement dépassée pour les faibles longueurs de canal. Cela signifie que le phénomène de survitesse peut être mis à profit pour obtenir des transistors de transconductance plus élevée que la limite théorique donnée par la vitesse de saturation [14]. Les performances en vitesse de ces circuits sont alors plus élevées que les prédictions des modèles classiques [15].

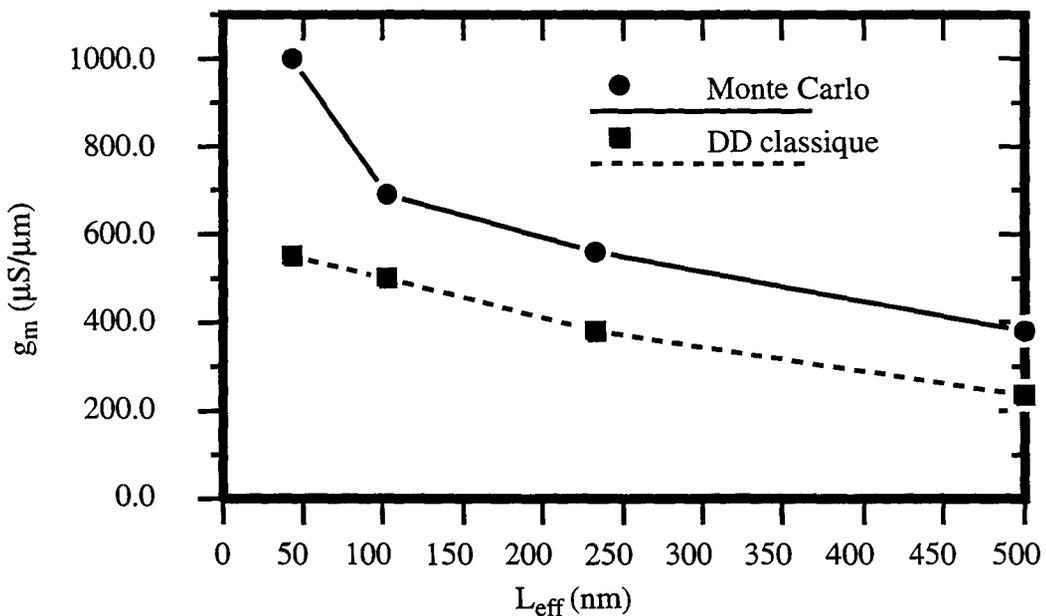


Figure IV-35 : Transconductance  $g_m$  en fonction de la longueur du canal.

### IV-5.5.3 Influence des zones résistives

Les courbes présentées ci-dessus sont relatives à des transistors pour lesquels les zones de source et de drain sont tronquées de façon à limiter le nombre de particules nécessaires lors des simulations Monte Carlo. Une meilleure prise en compte des résistances d'accès corrigerait les valeurs de transconductance à la baisse, comme l'indiquent les courbes suivantes, sans toutefois supprimer l'influence du phénomène de survitesse.

La figure IV-36 donne les courbes  $I_D(V_{DS})$  calculées par la méthode de Monte Carlo pour un transistor de longueur effective 43 nm où les électrodes de source et de drain sont placées à 0.12 micron de la jonction au lieu de 0.07 micron lors des simulations réalisées jusqu'ici. On retrouve globalement le même comportement que pour le dispositif précédent, mis à part une diminution du courant due à la chute de potentiel dans les zones résistives de source et de drain et une baisse de la transconductance à environ  $700 \mu S/\mu m$ . Les caractéristiques calculées à l'aide du formalisme étendu de dérive-diffusion sont reportées à la figure IV-37. Elles montrent une plus grande cohérence avec les résultats de simulations Monte Carlo que celles du paragraphe IV -5.5.1, illustrant l'importance de la prise en compte d'une partie suffisante des zones fortement dopées afin de simuler correctement les contacts ohmiques.

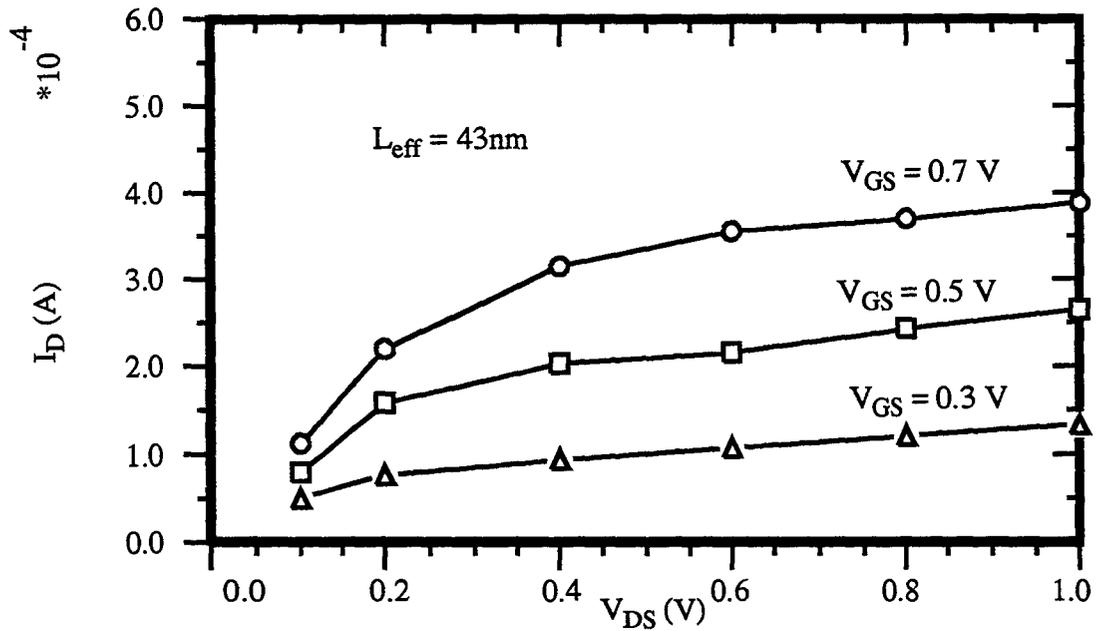


Figure IV-36 : Caractéristique  $I_D(V_{DS})$  issue de simulations Monte Carlo. La longueur des zones de source et de drain est augmentée de 0.05 micron.

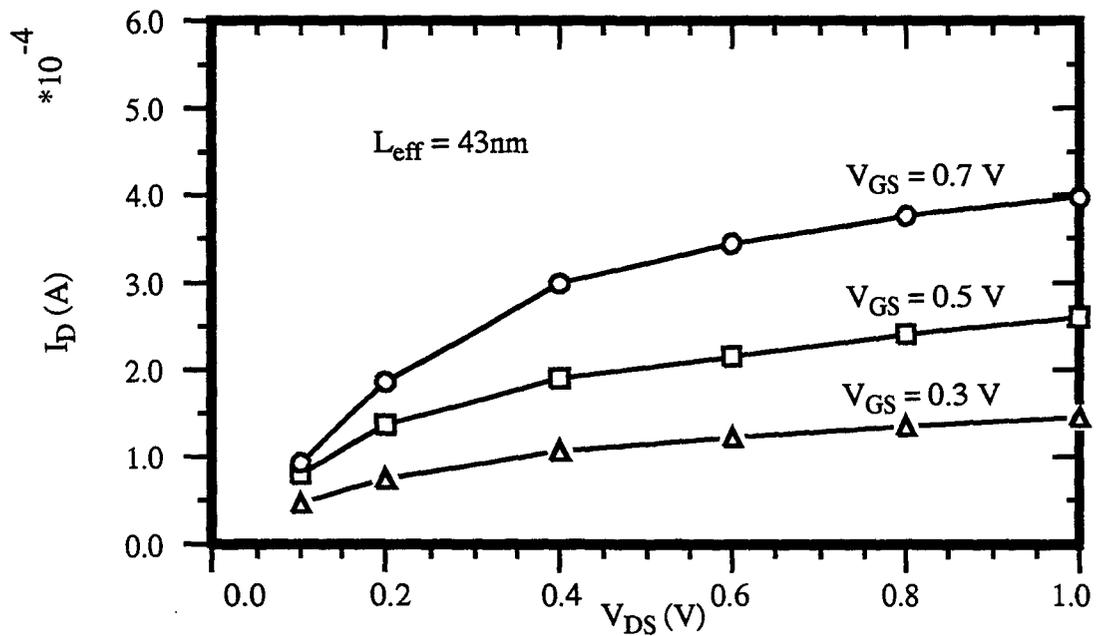


Figure IV-37 : Caractéristique  $I_D(V_{DS})$  issue de simulations utilisant le formalisme de dérive-diffusion étendu. La mobilité et la tension thermique sont évalués lors des simulations Monte Carlo dont les résultats sont donnés à la figure IV-36.

#### IV-5.6 Conclusion

La méthode de Monte Carlo prend en compte les effets balistiques survenant dans le fonctionnement des dispositifs submicroniques. La technique de couplage entre les simulations Monte Carlo et le formalisme étendu de dérive-diffusion permet de simuler les phénomènes de survitesse lorsqu'ils interviennent dans le fonctionnement d'un dispositif, ainsi que le montrent les résultats présentés dans ce chapitre. Le couplage a également l'avantage de permettre une convergence rapide des simulations, la résolution des équations de dérive-diffusion réalisant une mise à jour des concentrations en porteurs plus rapide que la simple méthode de Monte Carlo. Outre des informations détaillées sur les phénomènes de transport à l'intérieur des dispositifs, il est également possible de simuler leurs caractéristiques électriques, ce qui est d'un grand intérêt pour la compréhension du fonctionnement des transistors fortement submicroniques et leur optimisation. Les simulations présentées ci-dessus montrent ainsi que le phénomène de survitesse peut être mis à profit pour améliorer certaines caractéristiques électriques et notamment augmenter la transconductance des transistors. Ces différents atouts font du couplage de la méthode de Monte Carlo et du formalisme de dérive-diffusion étendu un outil efficace de simulation de dispositif.



## REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE IV

- [1] : J. L. Prince, *Very large scale integration*, Ed. D. F. Barbe, Springer-Verlag
- [2] : C. Jacoboni, P. Lugli, *The Monte Carlo method for semiconductor device simulation*, Ed. S. Selberherr, Springer-Verlag
- [3] : S. M. Sze, *Physics of semiconductor devices*, John Wiley & Sons, 1981
- [4] : J. P. Uyemura, *Fundamentals of MOS digital integrated circuits*, Addison-Wesley, 1988
- [5] : P. K. Chatterjee, W. R. Hunter, I. C. Holloway, Y. T. Lin, *IEEE Elec. Dev. Let.*, vol.1, p220, 1980
- [6] : M. R. Wordeman, A. M. Schweighart, R. H. Dennard, G. Sai-Halasz, W. W. Molzen, *A fully scaled submicrometer NMOS technology using direct-write e-beam lithography*, *IEEE Trans. Elec. Dev.*, vol.32, No.11, p2214, novembre 1985
- [7] : R. H. Yan, A. Ourmazd, K. F. Lee, D. Y. Jeon, C. S. Rafferty, M. R. Pinto, *Scaling the Si metal-oxide-semiconductor field-effect transistor into the 0.1- $\mu$ m regime using vertical doping engineering*, *Appl. Phys. Lett.* vol.59, No.25, p3315, 16 décembre 1991
- [8] : D. J. Frank, S. E. Laux, M. V. Fischetti, *Monte Carlo simulation of 30nm dual-gate MOSFET: how short can Si go?*, *IEDM 92 technical digest*, p553, 1992
- [9] : S. Y. Chou, D. A. Antoniadis, H. I. Smith, *Observation of electron velocity overshoot in sub-100-nm-channel MOSFET's in silicon*, *IEEE Elec. Dev. Let.*, vol.6, No.12, décembre 1985
- [10] : G. G. Shahidi, D. A. Antoniadis, H. I. Smith, *Electron velocity overshoot at room and liquid nitrogen temperatures in silicon inversion layers*, *IEEE Elec. Dev. Let.*, vol.9, No.2, février 1988
- [11] : D. A. Antoniadis, *Physics and technology of ultra short channel MOSFET devices*, *IEDM 91 technical digest*, p21, 1991

- [12] : D. A. Antoniadis, *Physics and technology for MOSFETs at 0.1 micron and below*, ESSDERC'93, p3-9, Ed. Frontières, 1993
- [13] : G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, E. Ganin, S. Rishton, D. S. Zicherman, H. Schmid, M. R. Polcari, H. Y. Ng, P. J. Restle, T. H. P. Chang, R. H. Denard, *Design and experimental technology for 0.1- $\mu\text{m}$  gate-length low-temperature operation FET's*, IEEE Elec. Dev. Let., vol.8, No.10, octobre 1987
- [14] : G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, S. Rishton, E. Ganin, *High transconductance and velocity overshoot in NMOS devices at the 0.1- $\mu\text{m}$  gate-length level*, IEEE Elec. Dev. Let., vol.9, No.19, septembre 1988
- [15] : G. A. Sai-Halasz, *Experimental technology and performance of 0.1- $\mu\text{m}$  gate-length low-temperature operation MOSFET's*, Extended abstracts of the 1988 International Conference on Solid State Devices and Materials, Tokyo, 1988
- [16] : R. H. Yan, K. F. Lee, D. Y. Jeon, Y. O. Kim, B. G. Park, M. R. Pinto, C. S. Rafferty, D.M. Tennant, *High performance 0.1- $\mu\text{m}$  room temperature Si MOSFETs*, 1992 Symposium on VLSI Technology Digest of technical papers, p86, 1992
- [17] : K. F. Lee, R. H. Yan, D. Y. Jeon, G.M. Chin, Y. O. Kim, D.M. Tennant, B. Razavi, H.D. Lin, Y.G. Wey, *Room temperature 0.1- $\mu\text{m}$  CMOS technology with 11.8 ps gate delay*, IEDM 93 technical digest, p131, 1993
- [18] : M. Iwase, T. Mizuno, M. Takahashi, H. Niiyama, M. Fukumoto, K. Ishida, S. Inaba, Y. Takigami, A. Sanda, A. Toriumi, M. Yoshimi, *High-performance 0.10- $\mu\text{m}$  CMOS devices operating at room temperature*, IEEE Elec. Dev. Let., vol.14, No.2, février 1993
- [19] : M. Saito, T. Yoshitomi, M. Ono, Y. Akasaka, H. Nii, S. Matsuda, H. S. Momose, Y. Katsumata, Y. Ushiku, H. Iwai, *An SPDD P-MOSFET structure suitable for 0.1 and sub 0.1 micron channel length and its electrical characteristics*, IEDM 92 technical digest, p897, 1992
- [20] : M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, H. Iwai, *Sub-50nm gate length N-MOSFETs with 10nm phosphorus source and drain junctions*, IEDM 93 technical digest, p119, 1993

- [21] : M. Aoki, T. Ishii, T. Yoshimura, Y. Kiyota, S. Iijima, T. Yamanaka, T. Kure, K. Ohyu, T. Nishida, S. Okazaki, K. Seki, K. Shimohigashi, *0.1- $\mu\text{m}$  CMOS devices using low-impurity-channel transistors*, IEDM 90 technical digest, p939, 1990
- [22] : M. Aoki, T. Ishii, T. Yoshimura, Y. Kiyota, S. Iijima, T. Yamanaka, T. Kure, K. Ohyu, T. Nishida, S. Okazaki, K. Seki, K. Shimohigashi, *Design and performance of 0.1- $\mu\text{m}$  CMOS devices using low-impurity-channel transistors*, IEEE Elec. Dev. Let., vol.13, No.1, janvier 1992
- [23] : H. Noda, F. Murai, S. Kimura, *Threshold voltage controlled 0.1- $\mu\text{m}$  MOSFET utilizing inversion layer as extreme shallow source/drain*, IEDM 93 technical digest, p123, 1993
- [24] : J. Tanaka, T. Toyabe, S. Ihara, S. Kimura, H. Noda, K. Itoch, *Simulation of sub-0.1- $\mu\text{m}$  MOSFET's with completely suppressed short-channel effect*, IEEE Elec. Dev. Let., vol.14, No.8, août 1993
- [25] : J. Tanaka, S. Kimura, H. Noda, T. Toyabe, S. Ihara, *A sub-0.1- $\mu\text{m}$  grooved gate MOSFET with high immunity to short-channel effect*, IEDM 93 technical digest, p537, 1993
- [26] : E. Dubois, P.H. Bricout, *Efficient scaling of MOSFETs in the sub-50 nm regime using a recessed channel*, 1993 International Semiconductor Device Research Symposium, Engineering Academic Outreach, University of Virginia, Charlottesville.
- [27] : W.-H. Lee, Y. J. Park, J. D. Lee, *Gate recessed (GR) MOSFET with selectively halo-doped channel and deep graded source/drain for deep submicron CMOS*, IEDM 93 technical digest, p135, 1993
- [28] : E. Vandenbossche, D.E.A. d'électronique, Université des Sciences et Techniques de Lille Flandres-Artois, 1990
- [29] : M. V. Fischetti, S. E. Laux, D. J. Frank, *Monte Carlo analysis of semiconductor devices: the Damocles program*, IBM J. Res. Develop. vol.34, No.4, p466, juillet 1990

- [30] : E. Dubois, P.-H. Bricout, R. Fauquembergue, D. Collard, *Monte Carlo/drift-diffusion coupling: non-equilibrium transport in deep submicron MOSFETs*, IEDM 91 technical digest, p519, 1991
- [31] : S. Bandyopadhyay, M. E. Klausmeier-Brown, C. M. Maziar, S. Datta, M. S. Lundstrom, *A rigorous technique to couple Monte Carlo and drift-diffusion models for computationnally efficient device simulation*, IEEE Trans. Elec. Dev., vol.34, No.2, p392, février 1987
- [32] : P.-H. Bricout, E. Dubois, R. Fauquembergue, *Transport simulation in submicron MOSFETs by coupling Monte Carlo and drift-diffusion*, NASECODE IX conf., p24, 1993
- [33] : P.-H. Bricout, E. Dubois, R. Fauquembergue, *Simulation of non-equilibrium transport in deep submicron MOSFETs*, ESSDERC'93, Editions Frontières, p609, 1993
- [34] : M. V. Fischetti, *Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures-Part I: homogeneous transport*, IEEE Trans. Elec. Dev., vol.38, No.3, p634, mars 1991
- [35] : M. V. Fischetti, S. E. Laux, *Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures-Part II: sub-micrometer MOSFET's*, IEEE Trans. Elec. Dev., vol.38, No.3, p650, mars 1991

## CHAPITRE V

### TRANSISTORS SUBMICRONIQUES NON PLANS

La miniaturisation des dispositifs de base des circuits intégrés s'accompagne de la recherche de compromis liés aux différents effets de la réduction des dimensions [1] [2] [3] [4]. Ainsi, le dopage de substrat des transistors MOS est augmenté afin de restreindre la taille de la zone de déplétion du canal, ce qui par contre dégrade le pouvoir de coupure, diminue la tension de claquage source-drain et accentue les effets canaux courts. L'épaisseur de la couche d'oxyde tend à diminuer comme les autres dimensions, au prix d'un contrôle plus difficile de sa qualité et de son uniformité et d'une plus grande sensibilité au claquage qui interdisent de descendre en dessous de 3 nm [5]. La profondeur des jonctions est réduite, ce qui tend à augmenter les résistances d'accès et pose le problème technologique de la réalisation de jonctions ultra-fines. Selon la plupart des stratégies de miniaturisation, les tensions de polarisation sont diminuées et la tension de seuil doit alors être mieux maîtrisée afin de garantir des marges de bruit suffisantes. Globalement, la miniaturisation entraîne souvent une plus grande sensibilité aux paramètres technologiques tels que dose et énergie d'implantation, température et durée des diffusions et oxydations, qui se répercutent sur les performances des dispositifs. Le contrôle des effets canaux courts devient alors plus difficile pour les transistors MOS classiques au dessous de 0.1 micron. Au prix de techniques de fabrication plus complexes, il est cependant possible d'obtenir des dispositifs peu sensibles à ces effets canaux courts en adoptant une géométrie différente [5]-[16], tout en conservant un comportement satisfaisant à fort courant, comme le montrent les résultats présentés dans ce chapitre.

## V-1 TRANSISTORS NON PLANS

La fabrication de dispositifs MOS à grille enterrée est bien sûr différente de celle des transistors plans conventionnels. Les articles de Tanaka, Kimura [12] [13], et de Lee [9] présentent deux techniques permettant d'obtenir des structures auto-alignées après gravure de la zone de canal. L'utilisation de gravure fine permet d'obtenir une longueur de canal inférieure à 0.1 micron.

Les dispositifs simulés sur lesquels porte la suite de cette première partie sont représentés schématiquement en figure V-1. Pour le transistor plan conventionnel, la profondeur des jonctions notée  $x_j$  est de l'ordre de 10 nm, ce qui représente la limite technologique actuelle. La longueur effective du canal est simplement définie par la distance entre les deux jonctions. Pour le dispositif non plan,  $t_{ent}$  désigne la profondeur à laquelle la grille est enterrée. L'épaisseur d'oxyde est  $t_{ox}$ , de même que pour le transistor plan. La profondeur de jonction  $x_j$  peut par contre être positive ou négative, selon que l'interface entre oxyde et silicium est plus haute ou plus basse que la jonction: quand la profondeur de gravure est plus grande que l'épaisseur des zones de source et de drain,  $x_j$  est négatif, comme indiqué sur le deuxième schéma. Enfin, la topologie est caractérisée par deux paramètres essentiels:  $\alpha$  qui est l'angle entre l'horizontale et la partie oblique de l'interface Si/SiO<sub>2</sub>, et  $L_{eff}$  qui est la longueur du canal prise entre les deux jonctions de source et de drain, le long de l'oxyde de grille.

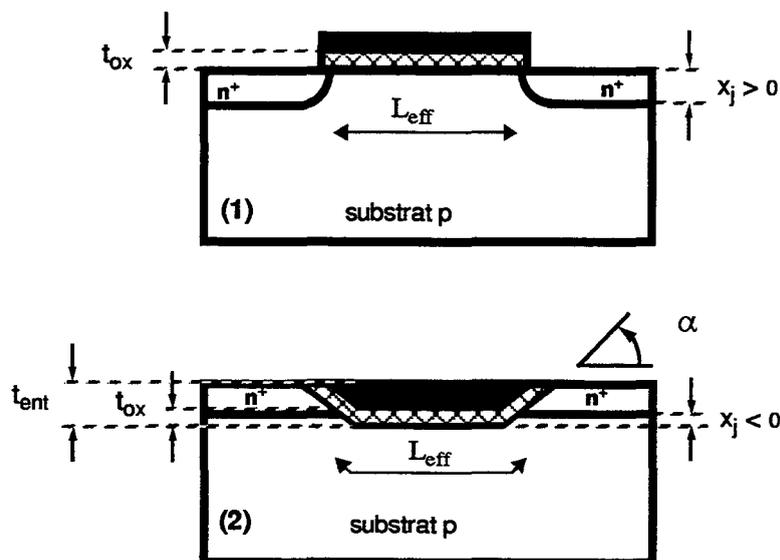


Figure V-1 : Représentation schématique des transistors plans (1) et à grille enterrée (2).

## V-2 ETUDE A FAIBLE POLARISATION

La tension de seuil est définie pour les transistors longs comme la tension à appliquer sur la grille pour amorcer l'inversion du canal, ce qui permet la mise en conduction du dispositif [1]. Pour les transistors courts, la zone de déplétion du canal subit l'effet de champ induit par la grille, mais aussi l'influence des jonctions source-substrat et drain-substrat. La tension de seuil devient alors une notion relative, dépendant des paramètres géométriques tels que la longueur effective et la forme des jonctions [2]. Par simplicité, on appelle ici tension de seuil, notée  $V_{TH}$ , la valeur de  $V_{GS}$  qui permet d'obtenir un courant de 10 nA par micron de largeur.

### V-2.1 Effet de coin

La forme du canal, dans le cas d'une grille enterrée, permet de générer aux angles une barrière de potentiel qui limite le courant source-drain [15]. Cet effet est illustré par la figure V-2 qui présente le potentiel à l'interface, le long du canal, pour les deux types de dispositifs. Les distributions de potentiel correspondent ici à  $V_{GS} = V_{TH}$ , la tension de drain augmentant de 0.2 V à 2 V. On observe clairement la présence des barrières de potentiel aux coins de la grille du dispositif non plan. De plus, l'augmentation de la tension de drain se répercute plus dans la zone de drain que dans le canal, contrairement à ce qui se produit pour le transistor plan. L'effet de champ, principe de fonctionnement du transistor MOS, est donc commandé plus efficacement par la grille, ce qui constitue une première explication à la diminution des effets liés à la réduction de la longueur du canal.

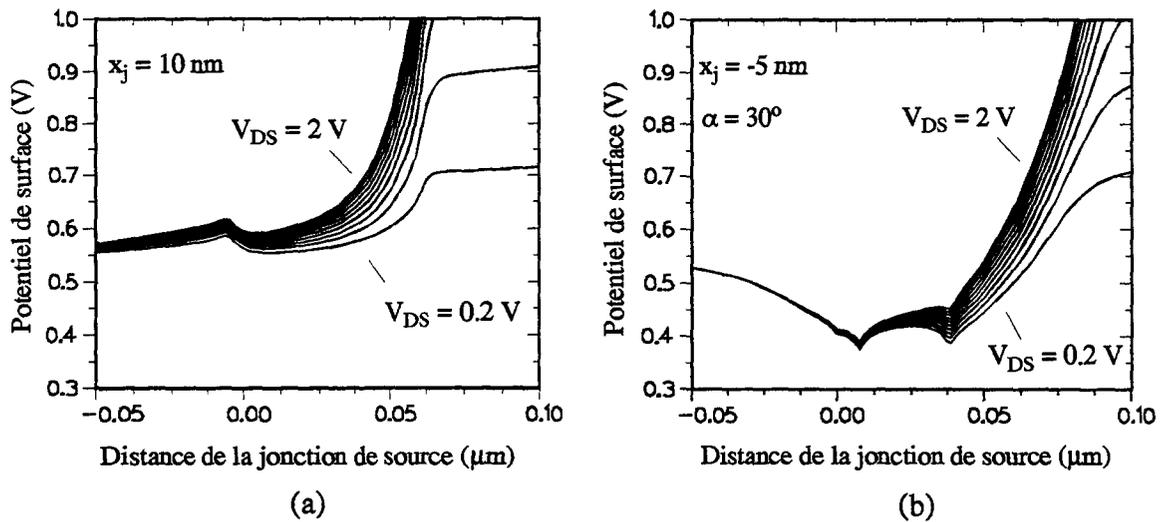
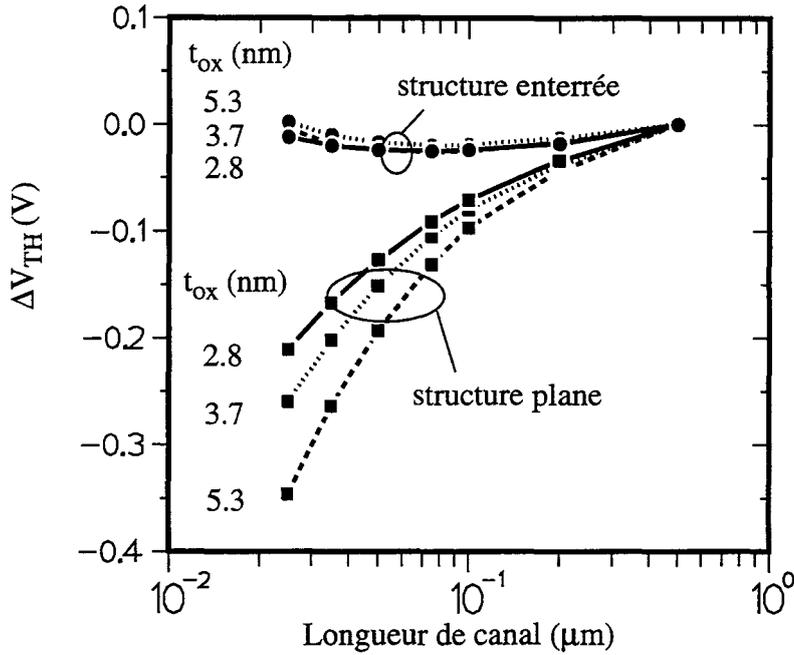


Figure V-2 : Potentiel de surface le long du canal à  $V_{GS} = V_{TH}$  pour les transistors plan (a) et à grille enterrée (b). ( $t_{ox} = 2.8 \text{ nm}$ ,  $N_{sub} = 2 \cdot 10^{17} \text{ cm}^{-3}$ )

### V-2.2 Réduction de la tension de seuil et pouvoir de coupure

La figure V-3 représente pour plusieurs épaisseurs d'oxyde de grille la variation de la tension de seuil  $\Delta V_{TH}$  lorsque la longueur du canal diminue, avec pour référence le dispositif de 0.25 micron. Les courbes se rapportant à la structure plane indiquent une diminution de la tension de seuil, ce qui s'explique par le contrôle plus important exercé par la zone de charge d'espace de la jonction drain-substrat sur la zone du canal. La structure à grille enterrée, quant à elle, garde pratiquement la même tension de seuil, sans influence notable de l'épaisseur d'oxyde.



**Figure V-3 :** Variation de la tension de seuil en fonction de la longueur effective du canal.

$$V_{DS} = 0.1 \text{ V}, N_{\text{sub}} = 2 \cdot 10^{17} \text{ cm}^{-3}; \text{ dispositif non plan: } x_j = -5 \text{ nm et } \alpha = 30^\circ. [15]$$

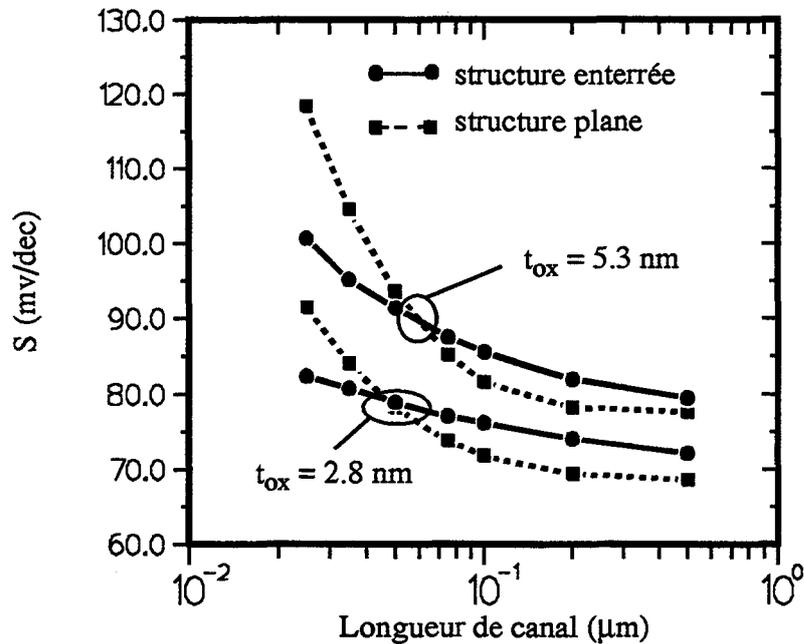
Un autre paramètre important pour décrire le fonctionnement des dispositifs MOS est l'inverse de la pente sous le seuil  $S$ , qui est la différence de tension de grille à appliquer pour accroître le courant d'une décade. Un bon pouvoir de coupure correspond donc à une valeur de  $S$  la plus faible possible, la limite à température ambiante étant  $S_{\text{min}} = \frac{k_B T}{q} \ln(10) \approx 60 \text{ mV/dec}$ . L'expression de  $S$  donnée ci-dessous, dérivée du comportement "canal long", permet d'interpréter l'influence de l'épaisseur d'oxyde:

$$S \approx \frac{k_B T}{q} \ln(10) \left( 1 + \frac{C_D}{C_{ox}} \right) \quad (\text{V-1})$$

où  $C_D$  et  $C_{ox}$  sont respectivement les capacités de déplétion et d'oxyde. Lorsque l'épaisseur d'oxyde  $t_{ox}$  est réduite, la capacité  $C_{ox}$  augmente et  $S$  diminue. Le pouvoir de coupure est amélioré par la réduction de l'épaisseur d'oxyde, comme le montre la figure V-4 qui donne également l'évolution de  $S$  avec la longueur de canal pour les deux types de transistors.

La structure plane dévie clairement du comportement "canal long". En effet, selon l'équa-

tion (III-1), la longueur ne devrait pas influencer le pouvoir de coupure. Dans le cas des transistors non plans,  $S$  est globalement moins affecté par la longueur du canal. Pour les très faibles longueurs, inférieures à 50 nm, la structure enterrée a ainsi un meilleur pouvoir de coupure, pour les différentes épaisseurs d'oxyde considérées.



**Figure V-4 :** Evolution du facteur  $S$  en fonction de la longueur effective du canal.

$$V_{DS} = 0.1 \text{ V}, N_{\text{sub}} = 2 \cdot 10^{17} \text{ cm}^{-3}; \text{ dispositif non plan: } x_j = -5 \text{ nm et } \alpha = 30^\circ. [15]$$

### V-2.3 Influence de la forme du canal

Les deux figures suivantes traduisent l'influence de la profondeur de jonction  $x_j$  et de la pente de l'interface  $\alpha$  sur la tension de seuil et sur le pouvoir de coupure. Alors que l'angle  $\alpha$  n'a aucune influence sur les variations de la tension de seuil, l'augmentation de l'effet de coin entraîne une dégradation du pouvoir de coupure:  $S$  augmente quand  $\alpha$  se rapproche de  $90^\circ$ , c'est-à-dire lorsque la grille devient rectangulaire.

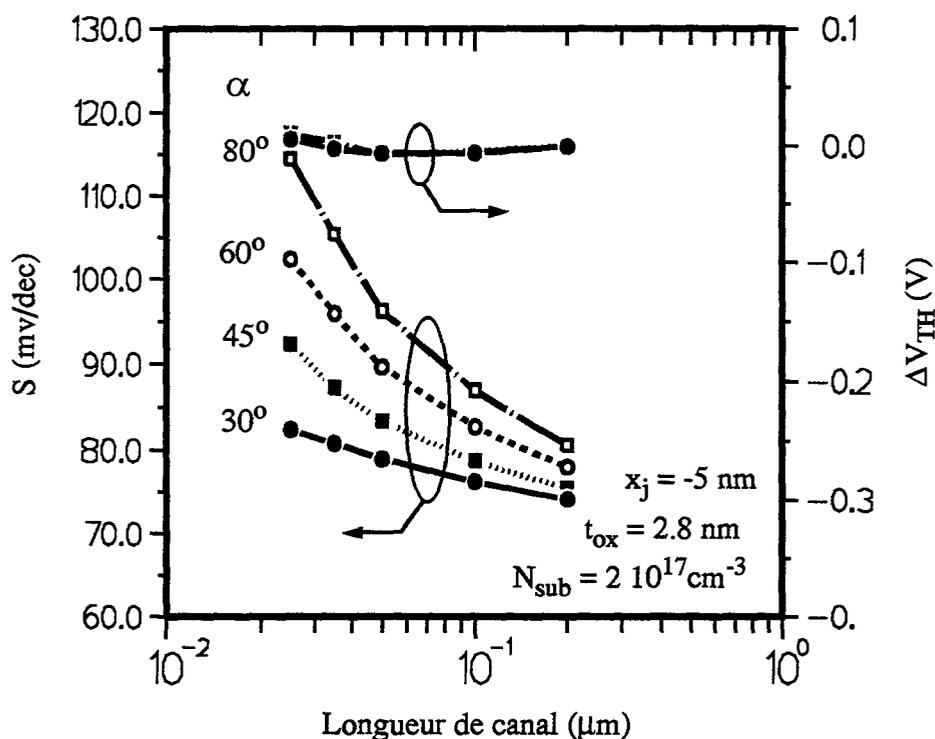
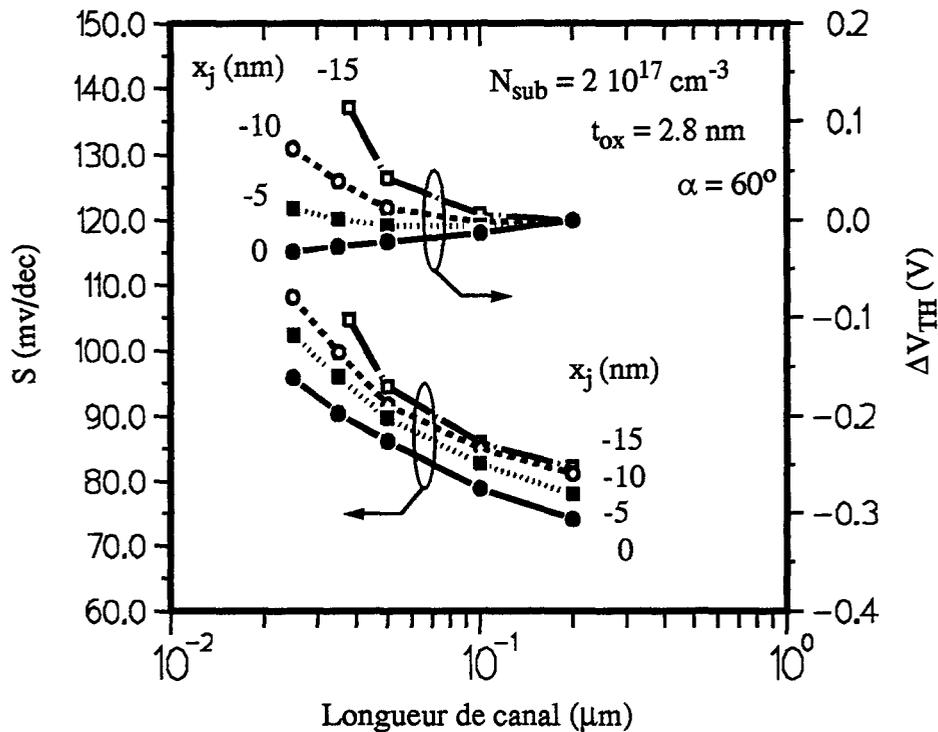


Figure V-5 : Influence de l'inclinaison de l'interface Si/SiO<sub>2</sub> sur les variations du pouvoir de coupure et de la tension de seuil en fonction de la longueur de canal.

Comme le montre la figure V-6, l'effet de coin peut aussi être renforcé en entrant plus profondément la grille, ce qui se traduit par une augmentation de  $|x_j|$ . On retrouve alors un décalage plus prononcé de la tension de seuil, mais cette fois dans le sens d'une augmentation. Pour  $x_j$  faible, entre 0 et -5 nm, la tension de seuil est peu sensible aux variations de la longueur du canal. D'autre part, le pouvoir de coupure se dégrade lui aussi plus rapidement lorsque  $|x_j|$  augmente.



**Figure V-6 :** Influence de la profondeur de jonction sur les variations du pouvoir de coupure et de la tension de seuil en fonction de la longueur de canal.

Pour garantir une bonne immunité de la tension de seuil par rapport aux fluctuations de la longueur du canal ainsi qu'un pouvoir de coupure suffisant, la forme de canal la plus adaptée doit être caractérisée par une pente faible de l'interface ( $\alpha = 30^\circ$ ) qui limite l'effet de coin, et une jonction au même niveau que le bas de l'interface ( $x_j$  négatif, proche de 0)

#### V-2.4 Influence de la concentration du substrat

Pour les transistors plans conventionnels, la concentration en dopants du substrat influence à la fois la tension de seuil et les caractéristiques sous le seuil [1]. Ceci est d'ailleurs mis à profit pour ajuster la tension de seuil en réalisant éventuellement un implant supplémentaire dans la zone de canal. Le fonctionnement sous le seuil est habituellement

modifié dans le sens d'une augmentation de  $S$  et donc d'une détérioration du pouvoir de coupure quand le dopage augmente [1]. En effet, une plus grande concentration du substrat contribue à augmenter la capacité de déplétion  $C_D$ . La figure V-7 présente ce type de dépendance pour les transistors plans les plus longs. Lorsque la longueur du canal est inférieure à 50 nm, la tendance s'inverse et les résultats ne peuvent plus être interprétés à l'aide des équations dérivées des modèles de type "canal long".

Pour les dispositifs non plans, la figure V-8 indique que le facteur  $S$  augmente avec la concentration pour toutes les longueurs de canal considérées. L'évolution est donc la même que celle suivie par les transistors classiques longs, ce qui dénote un excellent contrôle de l'effet de champ par la grille.

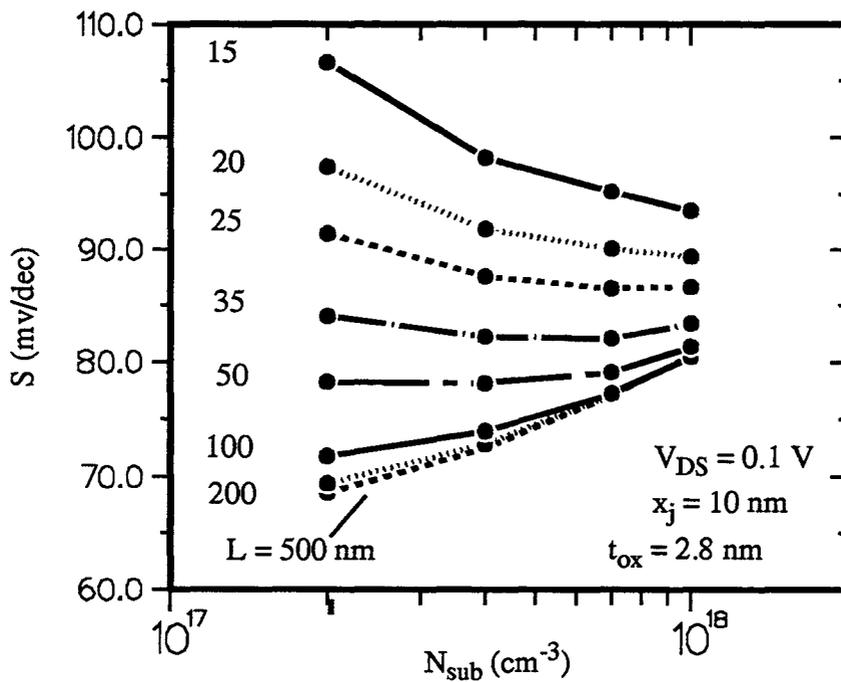
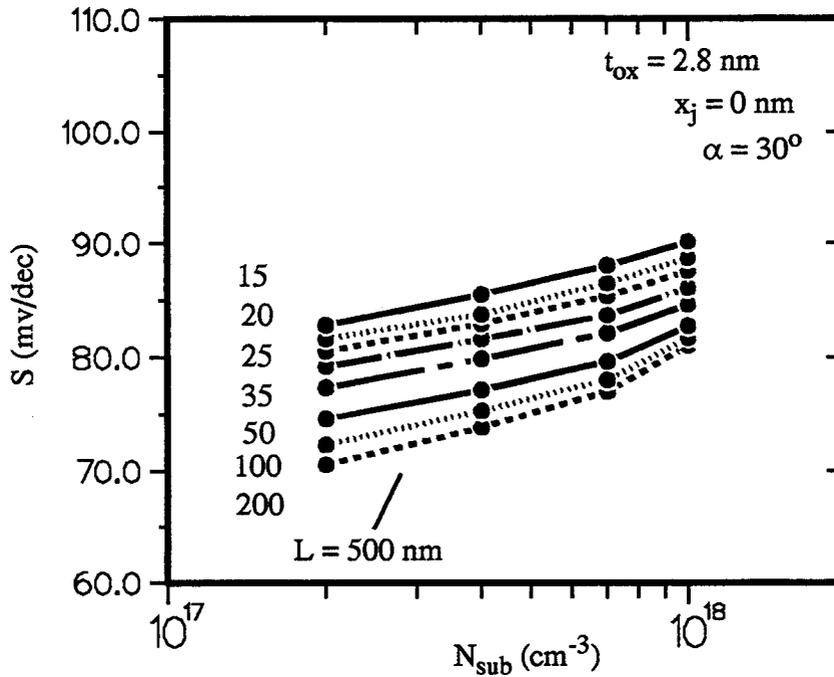


Figure V-7 : Evolution du facteur  $S$  en fonction de la concentration du substrat, pour les dispositifs plans [15].



**Figure V-8 :** Evolution du facteur  $S$  en fonction de la concentration du substrat pour les transistors non plans.

La figure V-9 ci-dessous donne l'influence de la concentration de substrat et de la longueur de canal sur la tension de seuil. Pour les deux types de dispositifs, la variation de la concentration permet l'ajustement de la tension de seuil à la valeur souhaitée. Les courbes de cette figure confirment également que la tension de seuil des transistors non plans est peu influencée par les fluctuations de longueur du canal, ce qui est loin d'être le cas pour les structures conventionnelles.

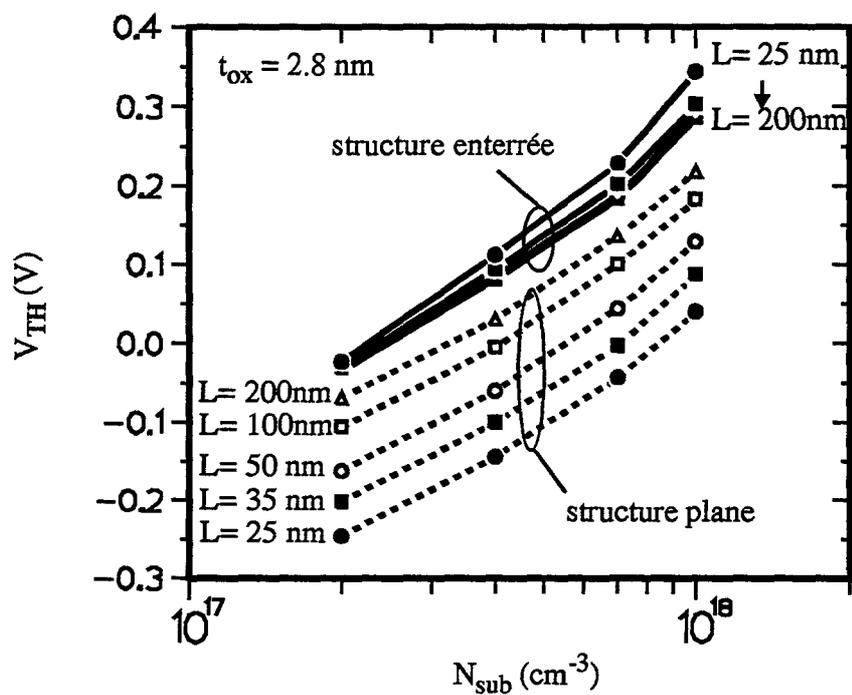


Figure V-9 : Evolution de la tension de seuil en fonction de la concentration du substrat.

$V_{DS} = 0.1$  V ; dispositif non plan:  $x_j = 0$  nm et  $\alpha = 30^\circ$  .

### V-3 ETUDE A FORTE POLARISATION

Afin de limiter l'effet des résistances de source et de drain, le dispositif plan est modifié et les simulations sont réalisées avec une structure LDD<sup>1</sup> pour laquelle la profondeur de la jonction est  $x_j = 10$  nm dans la zone d'extension. Les transistors à grille enterrée ne sont par contre pas modifiés, la profondeur de diffusion de la source et du drain étant suffisante.

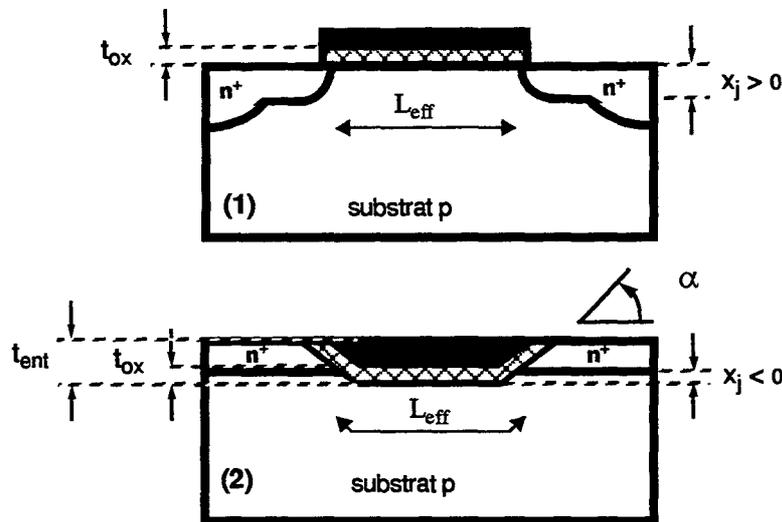


Figure V-10 : Représentation schématique des dispositifs simulés à forte polarisation.

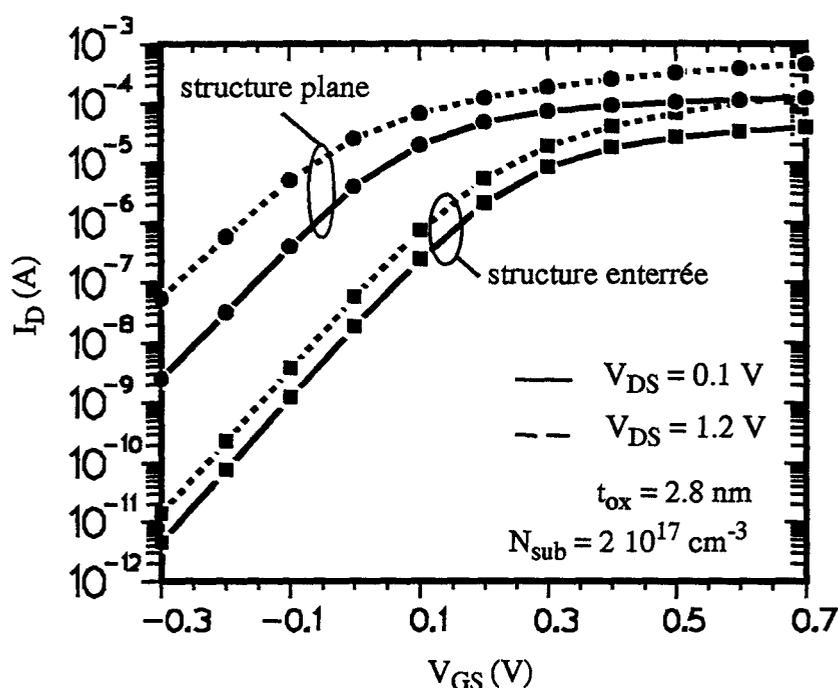
#### V-3.1 Simulations classiques de dérive-diffusion

Ces simulations réalisées à l'aide du formalisme de dérive-diffusion classique donnent quelques indications précieuses sur le fonctionnement des dispositifs fortement submicroniques, au moins du point de vue qualitatif, même si les modèles utilisés sont mis en défaut à forte polarisation.

Une des qualités recherchées pour ces transistors est un courant résiduel suffisamment faible dans l'état bloqué, c'est-à-dire à faible  $V_{GS}$ , afin de limiter la puissance consommée [17]. La figure V-11 donne des caractéristiques  $I_D(V_{GS})$  pour les deux types de structures, à

1. LDD: Lightly Doped Drain. La source et le drain sont réalisés à l'aide de deux implantations (voir le paragraphe IV-3 et la référence [3]).

faible  $V_{DS}$  (0.1 V) et à fort  $V_{DS}$  (1.2 V). Le courant à  $V_{GS}$  élevé, dans la zone de saturation, est plus élevé dans le transistor plan. A  $V_{GS}$  faible, le courant dépend moins de la tension de drain pour le dispositif non plan. Ceci confirme ce qui a déjà été observé lors de l'étude à faible polarisation, à savoir que le canal subit moins l'influence du drain que dans les structures planes, y compris pour des tensions  $V_{DS}$  importantes.



**Figure V-11** : Caractéristiques  $I_D(V_{GS})$  pour les deux types de dispositifs. Transistor non plan:  $x_j = -5$  nm et  $\alpha = 30^\circ$ .

Cette tendance est également illustrée par les courbes suivantes [16] qui donnent la tension de seuil (tension de grille permettant d'obtenir un courant de 10 nA dans un transistor de largeur  $1\mu\text{m}$ ) en fonction de la longueur du canal, à  $V_{DS} = 1.2$  V. D'après les résultats de simulations classiques de dérive-diffusion, néanmoins soumis à réserve à forte polarisation, la tension de seuil des dispositifs à grille enterrée dépend moins de la longueur du dispositif.

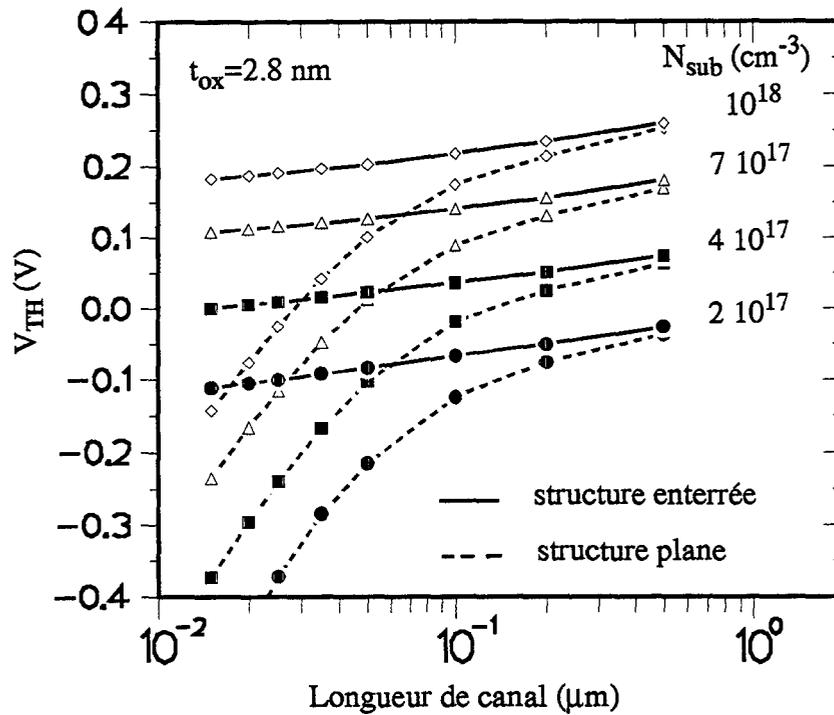


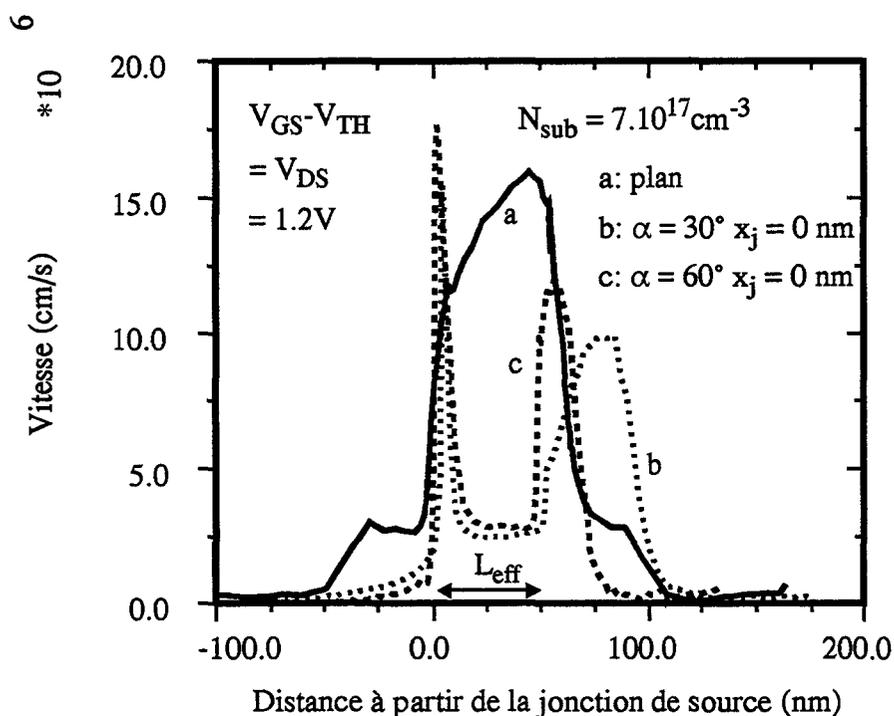
Figure V-12 : Evolution de la tension de seuil en fonction de la longueur effective du canal.

$V_{DS} = 1.2 \text{ V}$  .Dispositif non plan:  $x_j = 0 \text{ nm}$  et  $\alpha = 30^\circ$ .

### V-3.2 Effets balistiques: étude par la méthode de Monte Carlo

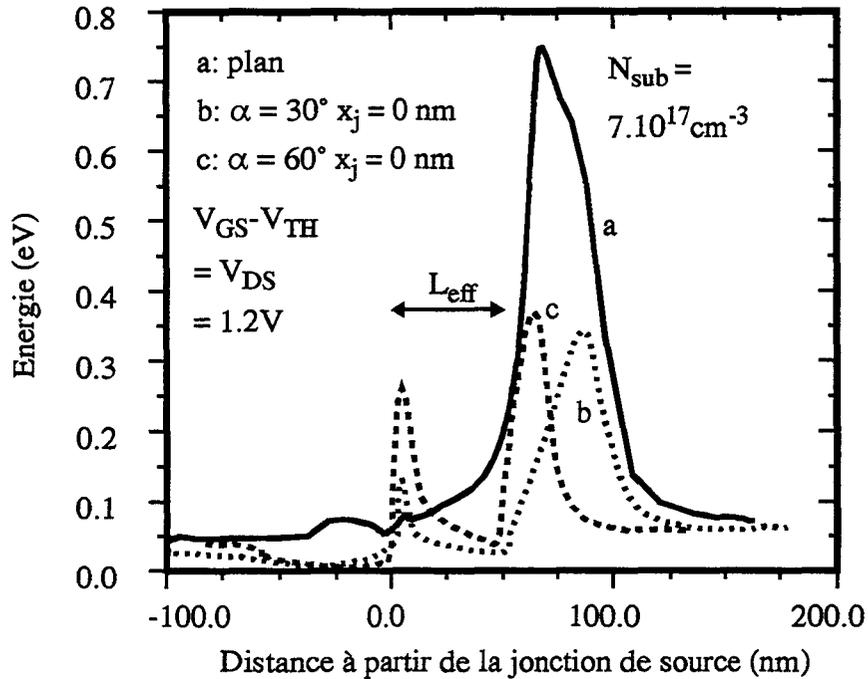
Les structures non planes et les dispositifs classiques plans diffèrent en termes de contrôle des effets canaux courts et d'amplitude des courants maximaux de saturation. Ces points peuvent être interprétés par une étude plus fine du transport électronique dans le canal des différents dispositifs, à l'aide de la méthode de Monte Carlo.

La figure V-13 représente la vitesse moyenne des électrons le long des transistors. La courbe de vitesse du dispositif plan est classique, montrant clairement un phénomène de survitesse prononcé. La figure V-14 concernant l'énergie indique d'ailleurs que l'on approche des limites de validité des modèles, l'énergie moyenne culminant à environ 0.75 eV.



**Figure V-13** : Composante de la vitesse parallèle à l'interface Si/SiO<sub>2</sub>. Moyenne calculée sur une profondeur de 10 nm. (Les créneaux de vitesse observés dans les zones de source et de drain correspondent aux extensions de la structure LDD.)

Pour les dispositifs non plans, les barrières de potentiel présentes aux angles de l'interface s'opposent au passage des électrons. Ceux qui franchissent l'obstacle en contournant la grille voient leur vitesse se réorienter dans la direction du champ électrique, ce qui se traduit par un fort pic de la vitesse moyenne. Les électrons sont ensuite accélérés dans le canal et gagnent de l'énergie cinétique (figure V-14) avant de subir l'influence de la deuxième barrière de potentiel.



**Figure V-14 :** Energie moyenne des électrons le long de l'interface Si/SiO<sub>2</sub>. Moyenne calculée sur une profondeur de 10 nm.

Quelques remarques s'imposent quant à l'interprétation de ces résultats:

\* D'abord, la réalisation de dispositifs à grille enterrée ne peut donner de profil géométrique vraiment anguleux. Les pics ne seront donc pas aussi marqués que ce qu'indiquent les résultats de simulation.

\* En second lieu, l'amplitude des phénomènes de transport balistique est fortement limitée par rapport aux transistors plans. Ce point doit cependant être relativisé, dans la mesure où les collisions des électrons sur les parties obliques de l'interface sont traitées comme des interactions diffuses lors de ces simulations, ce qui contribue à diminuer les effets balistiques.

\* Enfin, contrairement à ce qui se produit dans les transistors plans, les électrons perdent de l'énergie dans le canal. En effet, le champ parallèle à l'interface est extrêmement faible dans la zone du canal, comme le montre la figure V-15, ce qui confirme le peu d'influence exercée par le drain sur les électrons de la zone du canal, y compris à forte polarisation.

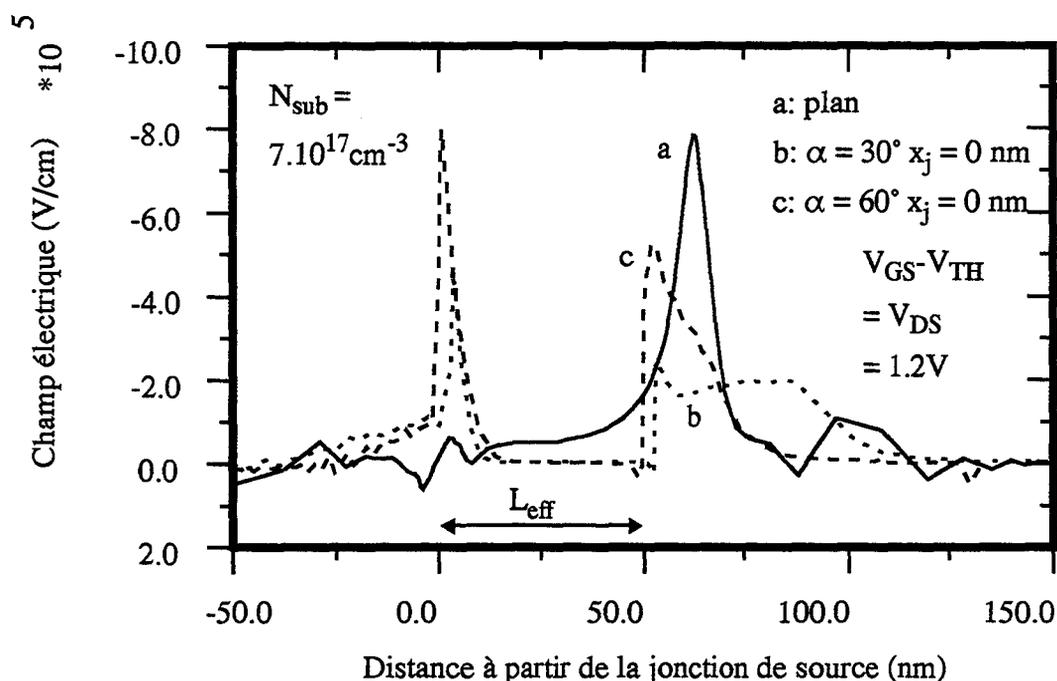


Figure V-15 : Champ électrique moyen subi par les électrons le long de l'interface Si/SiO<sub>2</sub>, calculé sur une profondeur de 10 nm.

Le transport balistique est mis en évidence dans le formalisme étendu de dérivediffusion par la valeur que prend la tension thermique. Ses différentes composantes évaluées lors d'une simulation Monte Carlo sont représentées à la figure V-16 le long de l'interface. Comme sur les courbes de vitesse et d'énergie, on retrouve deux zones hors d'équilibre situées près des angles de la grille. Dans la région de drain où les lignes de courant ne sont pas dirigées selon les axes principaux  $x$  et  $y$  mais obliques, les termes non diagonaux du tenseur de tension thermique ne sont pas nuls, même si leur valeur reste bien inférieure à celle des composantes diagonales  $U_{xx}$  et  $U_{yy}$ . Le signe négatif des termes  $U_{xy}$  et  $U_{yx}$  s'explique par le fait que le courant est dirigé vers le drain ( $x > 0$ ) et vers la surface ( $y < 0$ , par convention)

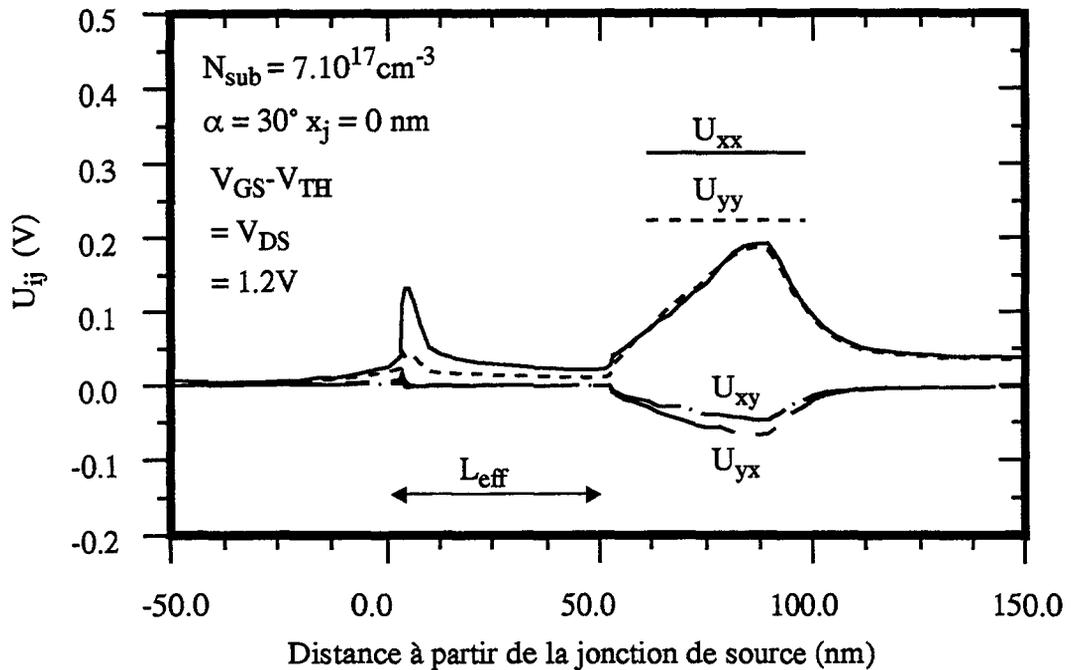
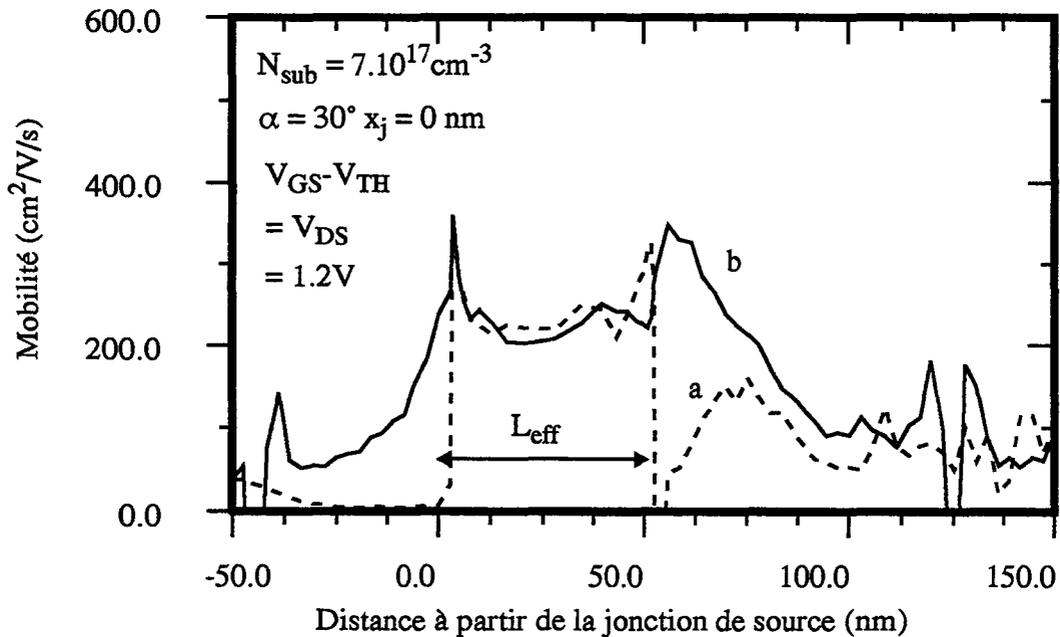


Figure V-16 : Composantes du tenseur de tension thermique le long de l'interface Si/SiO<sub>2</sub>: moyennes calculées sur une profondeur de 10nm.

Le terme de mobilité  $\mu_{xx}$  est tracé le long du canal en figure V-17. La courbe en trait plein correspond à la mobilité calculée sans tenir compte des collisions à l'interface pour l'évaluation du taux de perte du moment (cf. Chapitre III) alors qu'elles sont comptabilisées pour la courbe en pointillés. Pour les deux évaluations, la mobilité est plus élevée dans la zone du canal, où la concentration en dopants est plus faible. Contrairement à ce que l'on observe pour les dispositifs plans (IV-5.2.3) [18], il n'y a pas de réduction de mobilité à l'approche du drain due au champ électrique longitudinal, cette composante du champ étant nulle dans le canal. La différence entre les deux courbes dans les zones de source et de drain illustre la difficulté rencontrée pour modéliser les interactions des électrons avec l'interface. Lorsque ces collisions sont traitées comme des réflexions lors des simulations Monte Carlo, elles n'ont pas d'influence sur la mobilité, car elles ne "freinent" pas les particules. Si au contraire elles sont simulées comme des diffusions entraînant un changement aléatoire de direction, elles contribuent à la diminution de la mobilité, comme l'indique la comparaison des courbes dans les régions de source et de drain. Le traitement des collisions sur l'inter-

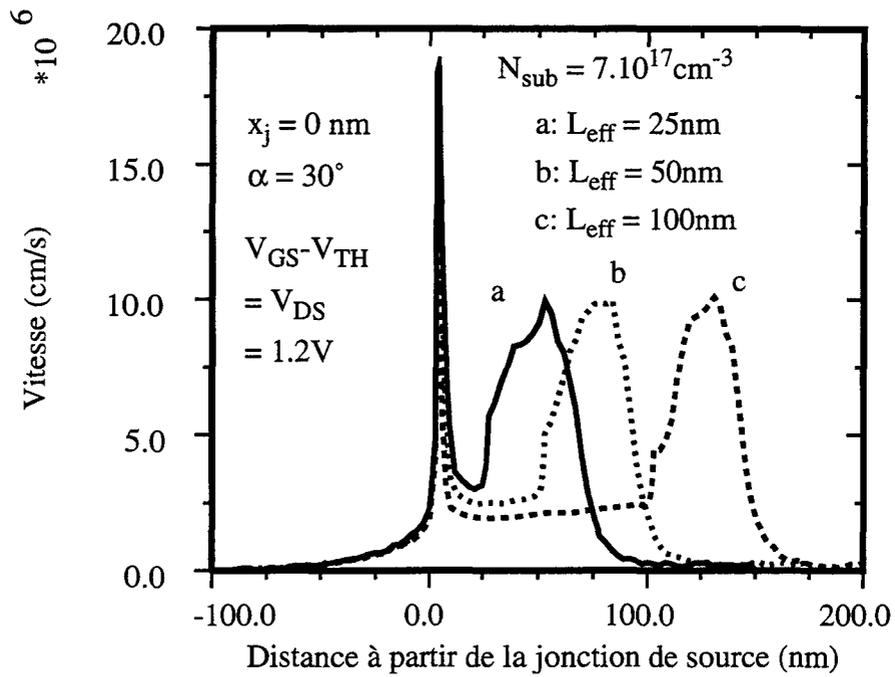
face oblique par redistribution aléatoire de la direction de la vitesse tel qu'il a été réalisé s'accompagne d'une réduction de mobilité sans aucun doute surestimée pour les électrons proches de l'interface. Un travail de modélisation reste à effectuer sur ce point précis afin de mettre au point un traitement plus réaliste de l'interface.



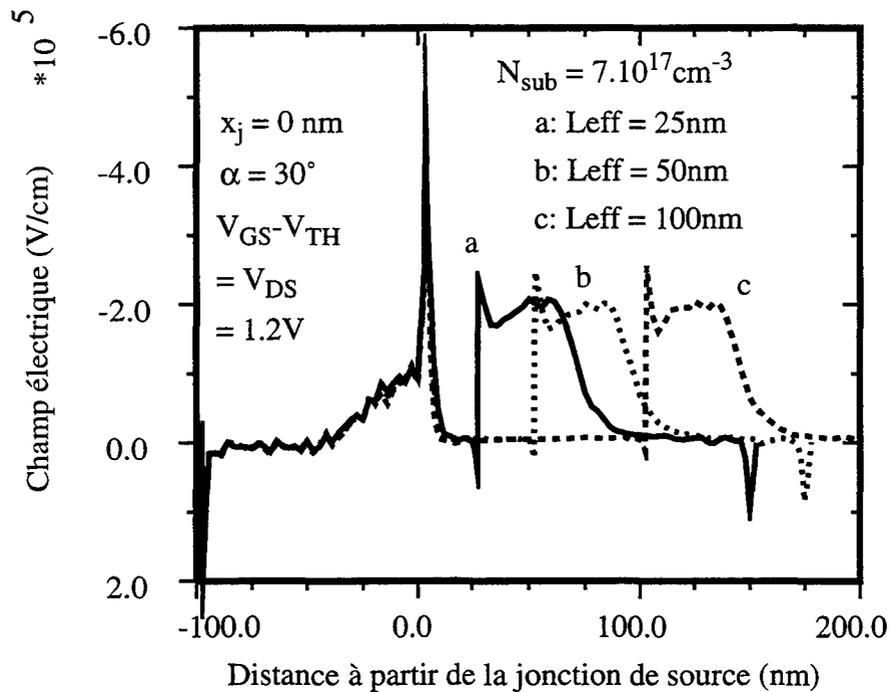
**Figure V-17 :** Mobilité des électrons sous l'interface Si/SiO<sub>2</sub> calculée en tenant compte (a) ou pas (b) des collisions à l'interface.

### V-3.3 Influence de la longueur du canal

D'après l'étude réalisée à forte polarisation à l'aide du formalisme classique de dérive-diffusion, au paragraphe V-3.1, les caractéristiques électriques des dispositifs non plans dépendent peu de la longueur du canal. Des transistors ayant les mêmes paramètres géométriques ( $\alpha = 30^\circ$ ,  $x_j = 0$  nm) mais des longueurs différentes ont été simulés par la méthode de Monte Carlo pour des polarisations semblables. Les profils de vitesse moyenne obtenus (figure V-18) sont peu différents: on observe simplement une translation du pic de vitesse coïncidant avec le deuxième angle de la grille.



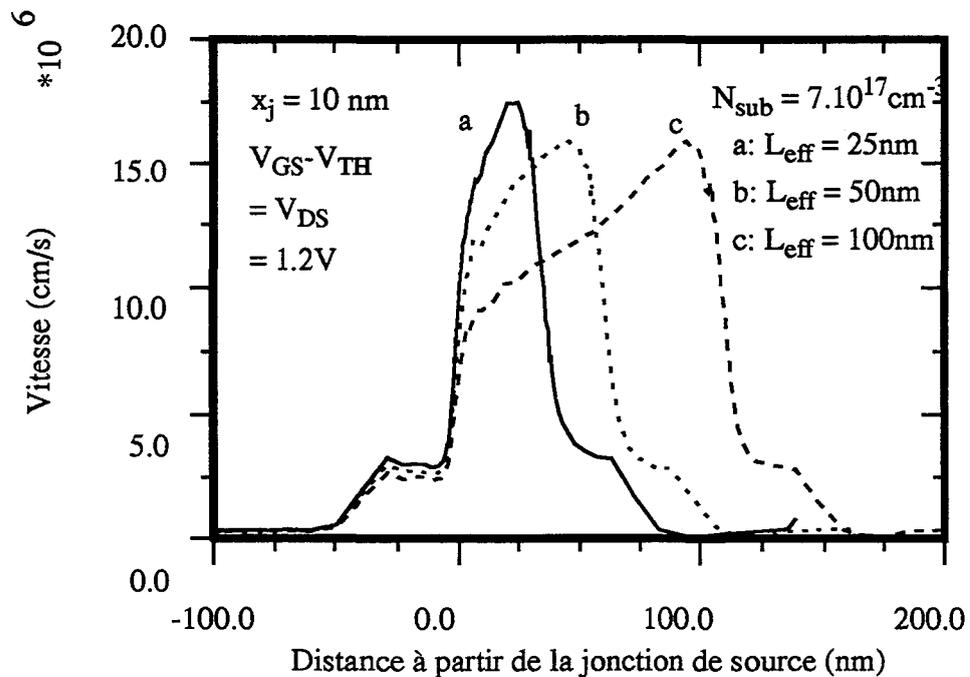
**Figure V-18** : Composante de la vitesse parallèle à l'interface Si/SiO<sub>2</sub>. Moyenne calculée sur une profondeur de 10 nm.



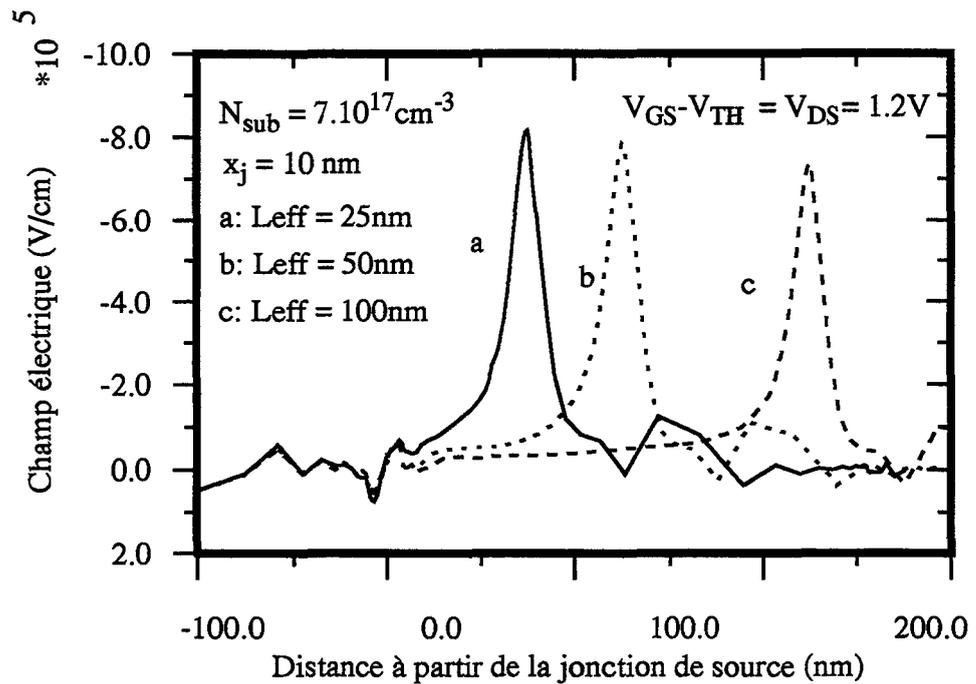
**Figure V-19** : Champ électrique moyen subi par les électrons le long de l'interface Si/SiO<sub>2</sub>, calculé sur une profondeur de 10 nm.

Ceci peut s'expliquer par le profil du champ électrique présenté à la figure V-19: le champ longitudinal retombe rapidement à une valeur quasiment nulle dans le canal, dont la longueur a par conséquent peu d'influence. Globalement, même à forte polarisation, le courant source-drain est totalement contrôlé par l'effet de coin.

Dans les transistors plans, au contraire, le drain exerce une influence croissante sur le transport dans le canal à mesure que la longueur effective diminue. Le phénomène de survitesse, localisé à proximité de la jonction de drain pour les dispositifs les plus longs, se produit dans l'ensemble du canal pour les très faibles longueurs. Ceci amène une augmentation de la vitesse moyenne des électrons, ce qu'indique la figure V-20. Cette évolution est liée au champ électrique longitudinal croissant dans le canal, représenté à la figure V-21. Ce phénomène de survitesse conduit ainsi à une augmentation du courant maximal et de la transconductance lorsque la longueur du canal est réduite.



**Figure V-20 :** Transistors plans: composante de la vitesse parallèle à l'interface Si/SiO<sub>2</sub>. Moyenne calculée sur une profondeur de 10 nm.



**Figure V-21** : Transistors plans: champ électrique moyen subi par les électrons le long de l'interface Si/SiO<sub>2</sub>, calculé sur une profondeur de 10 nm.

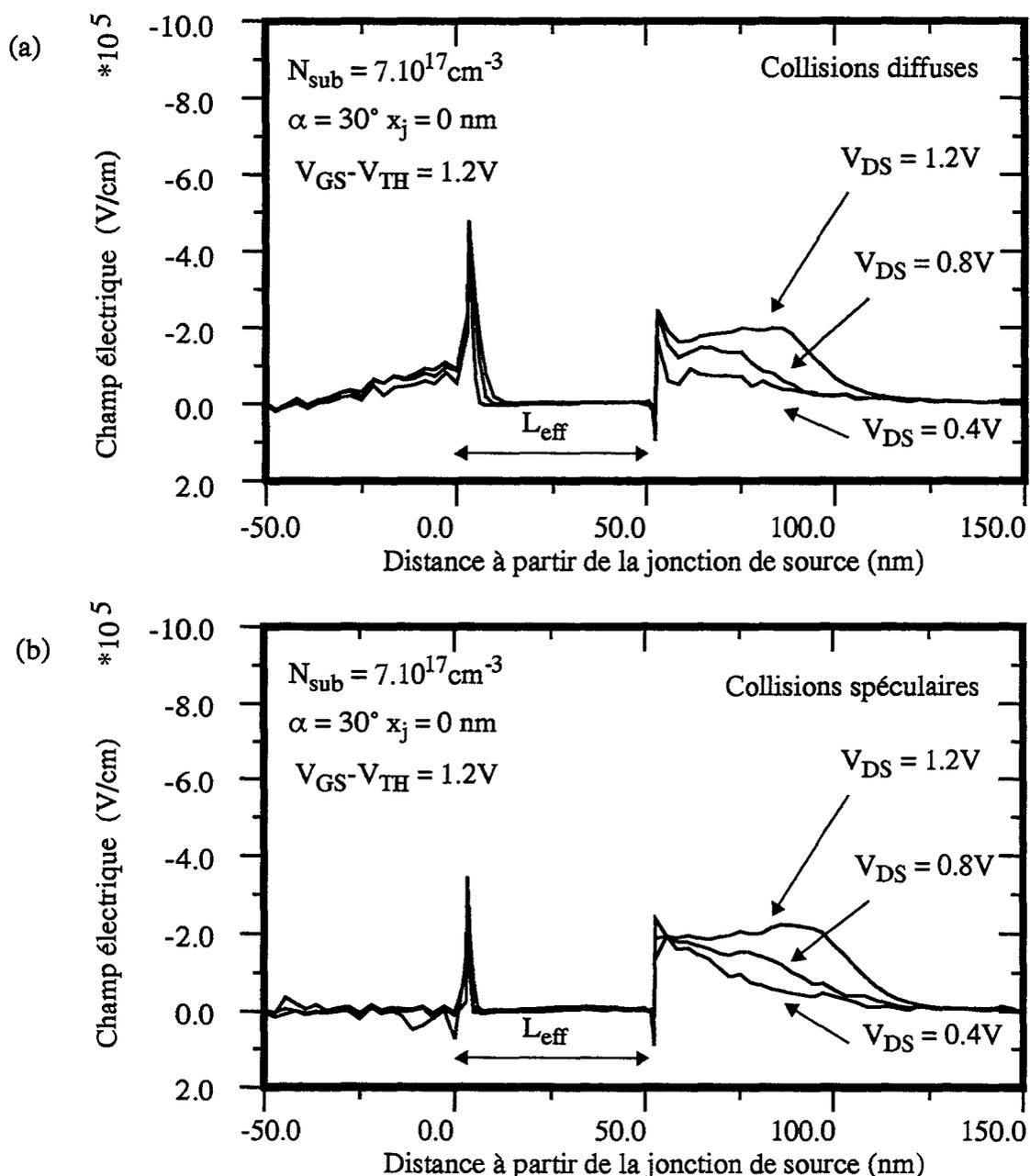
### V-3.4 Influence de la polarisation

Ce paragraphe est consacré à l'influence des tensions de polarisation  $V_{DS}$  et  $V_{GS}$  sur le transport hors d'équilibre dans les dispositifs à grille enterrée. L'ensemble des résultats Monte Carlo est fourni pour deux types de traitement des collisions avec les parties obliques de l'interface. Le premier les assimile à des collisions diffuses: la direction de la vitesse après le choc contre l'interface est tirée au sort, et toutes les directions possibles dirigées vers le silicium sont équiprobables. Le second traitement permet de conserver aux collisions un caractère spéculaire: il consiste en une réflexion effectuée sur la direction du vecteur vitesse et en une correction de la norme du vecteur d'onde et de la vitesse garantissant la conservation de l'énergie.

#### V-3.4.1 Influence de la tension drain-source $V_{DS}$

Les effets balistiques sont limités aux angles de l'interface et à la zone de drain, comme cela a déjà été mentionné. L'augmentation de la tension drain-source se traduit par

de plus grandes valeurs de champ électrique longitudinal, mais presque uniquement dans la région du drain, le champ dans le canal demeurant pratiquement nul, ainsi que l'illustrent les courbes suivantes.



**Figure V-22 :** Moyenne du champ électrique subi par les électrons dans la direction du canal, calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations. Les deux séries de courbes correspondent aux deux traitements des parties obliques de l'interface.

Ceci se répercute évidemment sur la vitesse des électrons, présentée à la figure V-23, qui diminue quand la tension de drain décroît, ainsi que sur l'énergie moyenne. Il y a cependant peu d'influence sur la zone du canal, contrairement à ce qui est observé pour les transistors plans. Ces différentes observations illustrent une fois de plus le contrôle efficace réalisé par l'effet de coin.

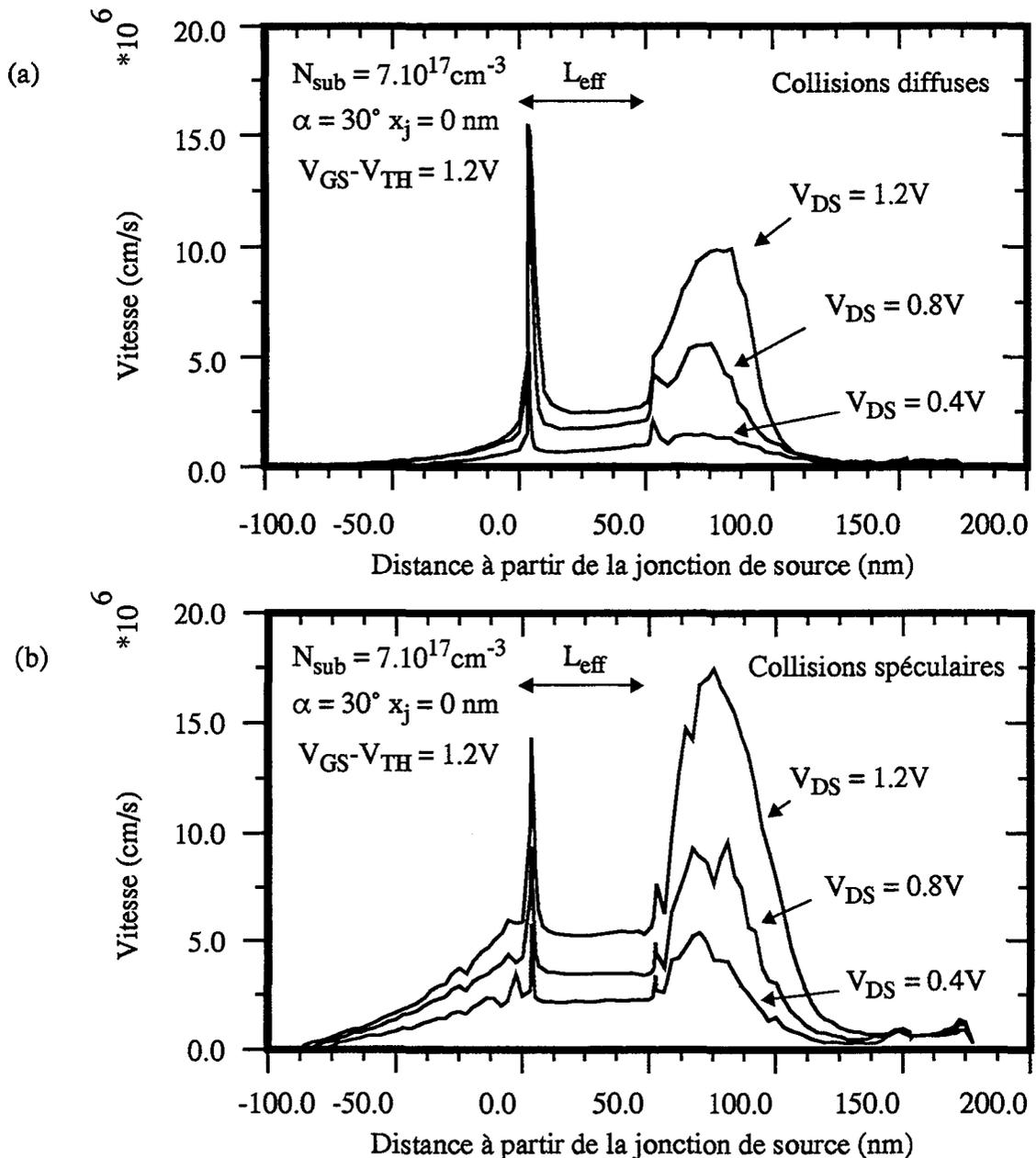


Figure V-23 : Vitesse moyenne calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations du drain

La comparaison des réseaux de courbes (a) et (b) de la figure V-23 montre également l'influence du type de traitement retenu pour l'interface sur l'amplitude du phénomène de survitesse dans le drain. La même observation peut être faite à propos de l'énergie moyenne des électrons représentée en figure V-24: les effets balistiques sont moins importants lorsque l'on considère que les électrons sont diffusés par les parties obliques de l'interface Si/SiO<sub>2</sub>.

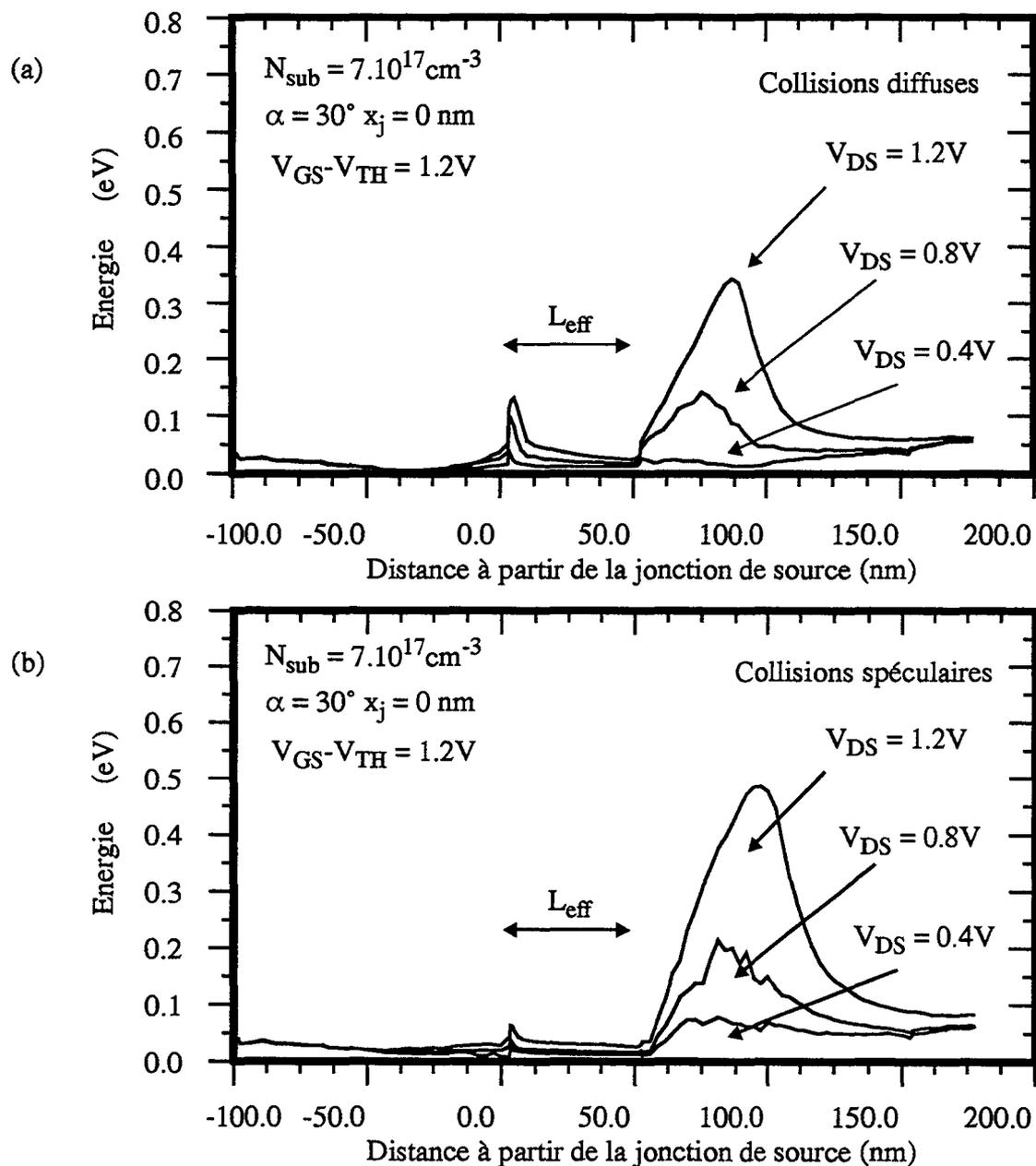


Figure V-24 : Energie moyenne calculée sur une profondeur de 10 nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations de drain

Les deux méthodes de traitement des parties obliques de l'interface permettent de cerner un phénomène physique réel, qui est la réduction de mobilité due à la rugosité. Alors que le traitement spéculaire ne les prend pas en compte, le recours aux collisions diffuses exagère sans doute l'importance. Quelque soit la méthode retenue, les conclusions concernant l'influence de la polarisation  $V_{DS}$  sont identiques: les effets balistiques sont localisés dans la région du drain, mettant en évidence l'effet de coin généré par la grille enterrée.

Les courbes d'énergie moyenne tracées à la figure V-24 présentent une anomalie dans la zone de source, où l'énergie est plus faible que la valeur à l'équilibre thermique. C'est une région où des électrons s'accumulent sous l'effet de la barrière de potentiel et pour laquelle la description des mécanismes d'interaction devrait être complétée pour tenir compte des fortes concentrations.

#### V-3.4.2 Influence de la tension grille-source $V_{GS}$

Les trois séries de courbes présentées ci-dessous illustrent l'influence de la tension de grille  $V_{GS}$ , à tension  $V_{DS}$  constante. Une réduction de la tension de grille correspond à une augmentation de la tension drain-grille, et donc à un champ électrique plus intense, comme l'indiquent les courbes de la figure V-25. Ceci se répercute sur la vitesse (figure V-26) et l'énergie des électrons par une extension de la zone de drain où interviennent des effets balistiques. Le pic d'énergie moyenne observé à la figure V-27 est ainsi de plus en plus important à mesure que la tension  $V_{GS}$  diminue, le traitement spéculaire de l'interface donnant ici aussi les valeurs les plus élevées. C'est donc à forte tension de drain et faible tension de grille que sont générés les électrons les plus chauds. Cette remarque est très importante du point de vue du vieillissement des dispositifs, d'autant plus que la réalisation d'une structure enterrée soulève le problème du claquage de l'oxyde aux angles de la grille.

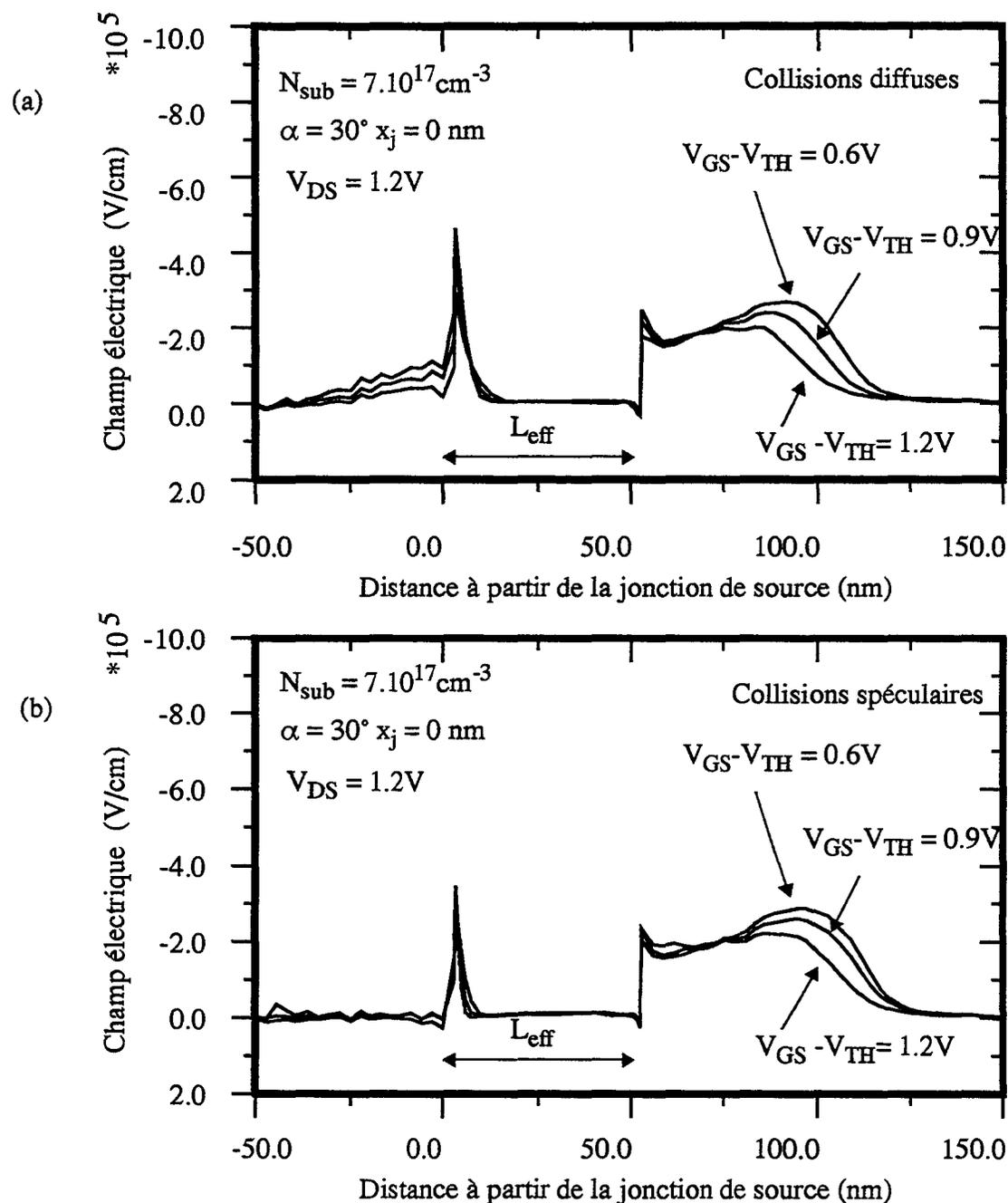


Figure V-25 : Moyenne du champ électrique subi par les électrons dans la direction du canal, calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations. Il est à noter que le traitement de l'interface a peu d'influence sur le profil du champ électrique.

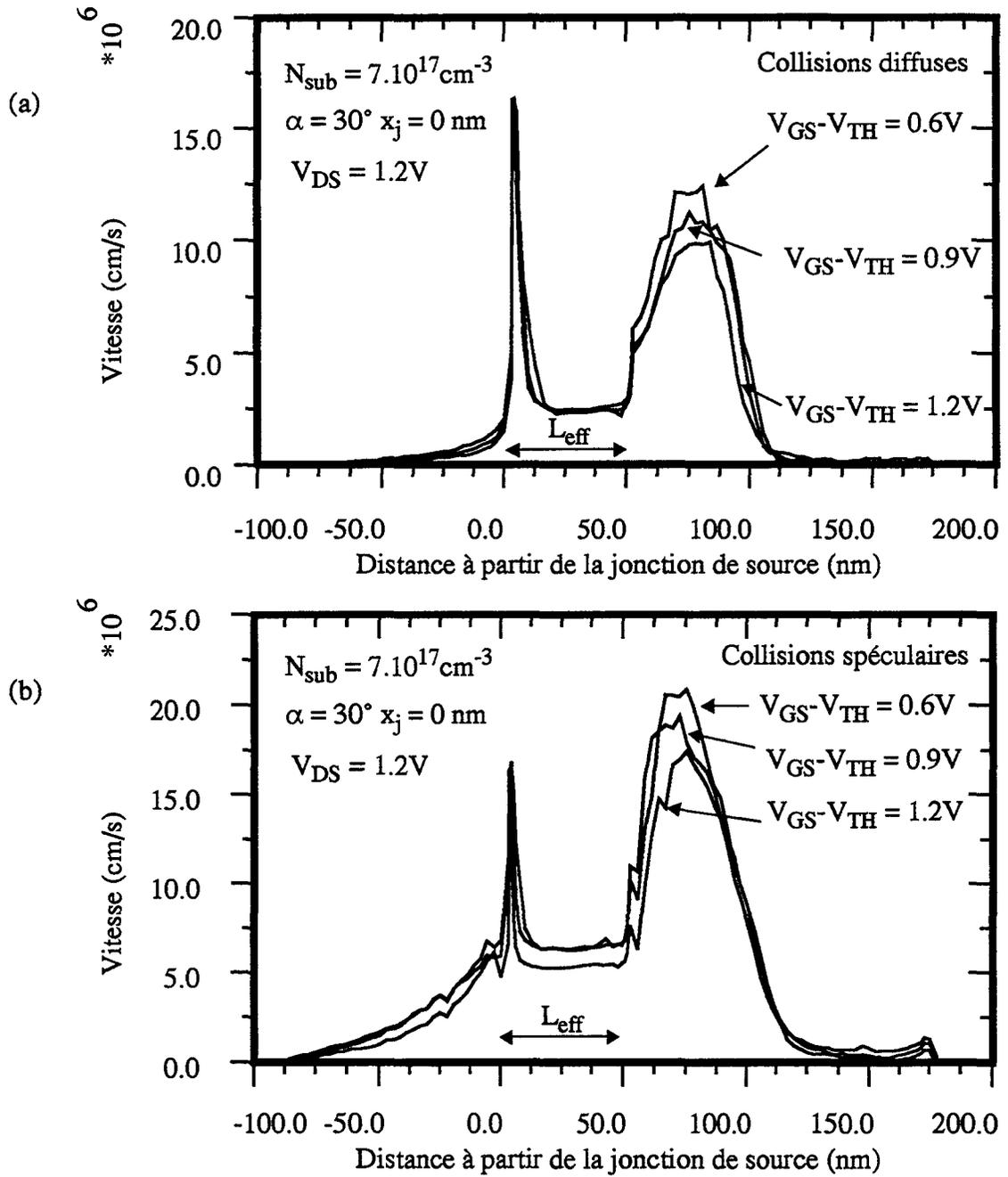


Figure V-26 : Vitesse moyenne calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations

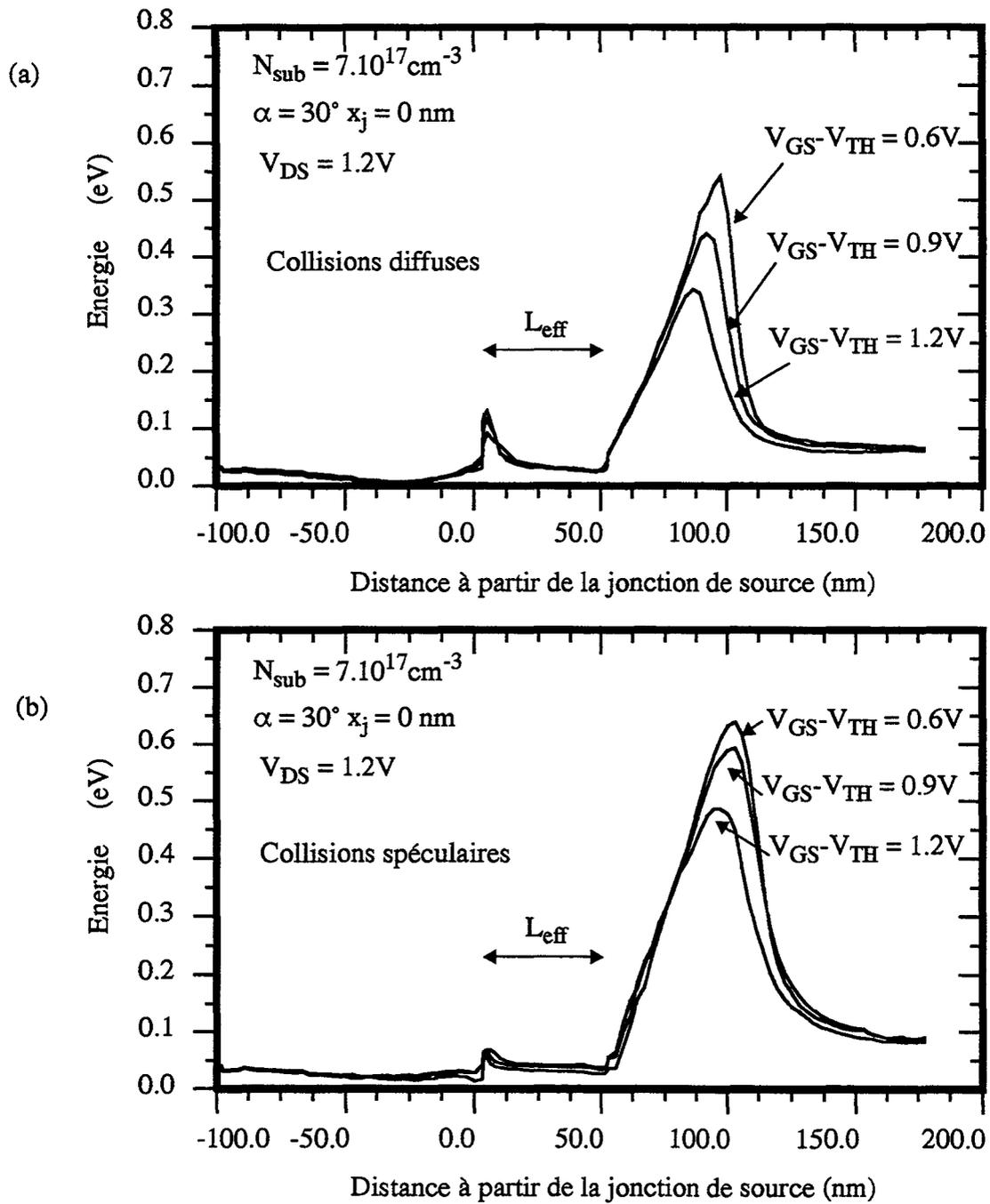
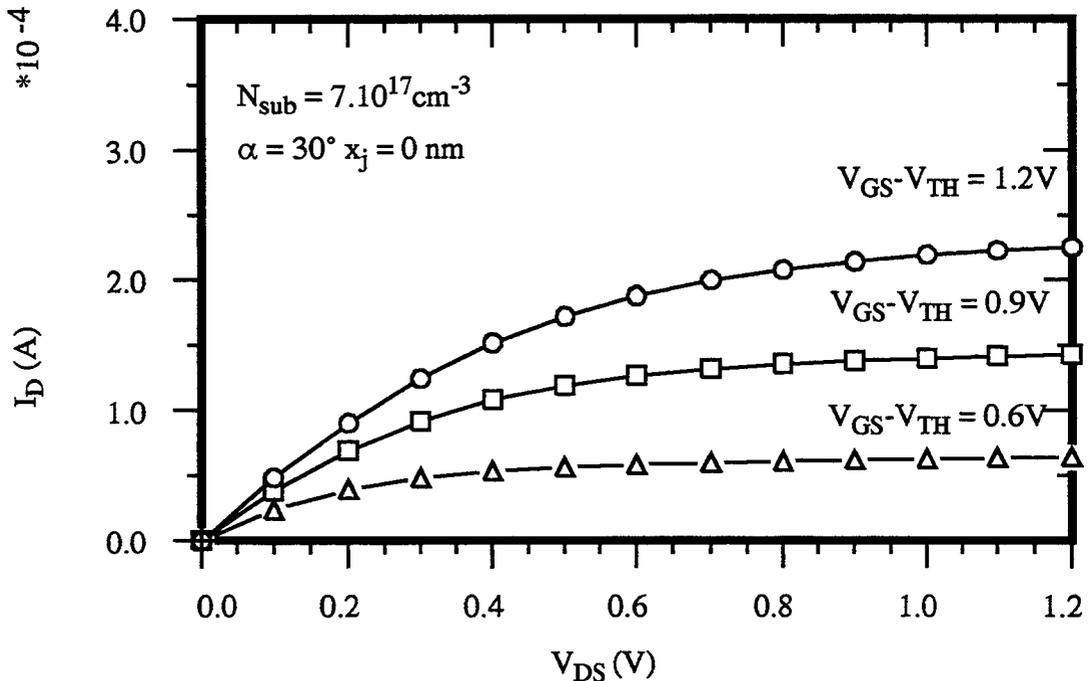


Figure V-27 : Energie moyenne calculée sur une profondeur de 10nm sous l'interface Si/SiO<sub>2</sub> pour différentes polarisations

### V-3.5 Performances en courant des dispositifs non plans

Pour le dispositif à grille enterrée caractérisé par une inclinaison de l'interface  $\alpha = 30^\circ$  et une profondeur de jonction  $x_j = 0$  nm, le réseau de caractéristiques  $I_D(V_{DS})$  a été calculé de deux façons, à l'aide du formalisme classique de dérive-diffusion et par la méthode de Monte Carlo.



**Figure V-28 :** Caractéristiques  $I_D(V_{DS})$  calculées par la méthode classique de dérive-diffusion.

La figure V-28 correspondant au formalisme de dérive-diffusion indique un comportement classique de la structure non plane. Les courbes de la figure V-29 sont issues de simulations Monte Carlo où les interactions avec les parties obliques de l'interface sont assimilées à des collisions diffuses. Elles gardent la même allure, avec toutefois des intensités de courant réduites de plus d'un facteur deux. Ceci est à relier à la remarque sur la mobilité du paragraphe V-3.2. La simulation de l'interface à l'aide de collisions diffuses entraîne une réduction dramatique de la mobilité le long des parties obliques, et une forte diminution du courant. Le traitement spéculaire de l'interface conduit aux caractéristiques présentées en figure V-30.

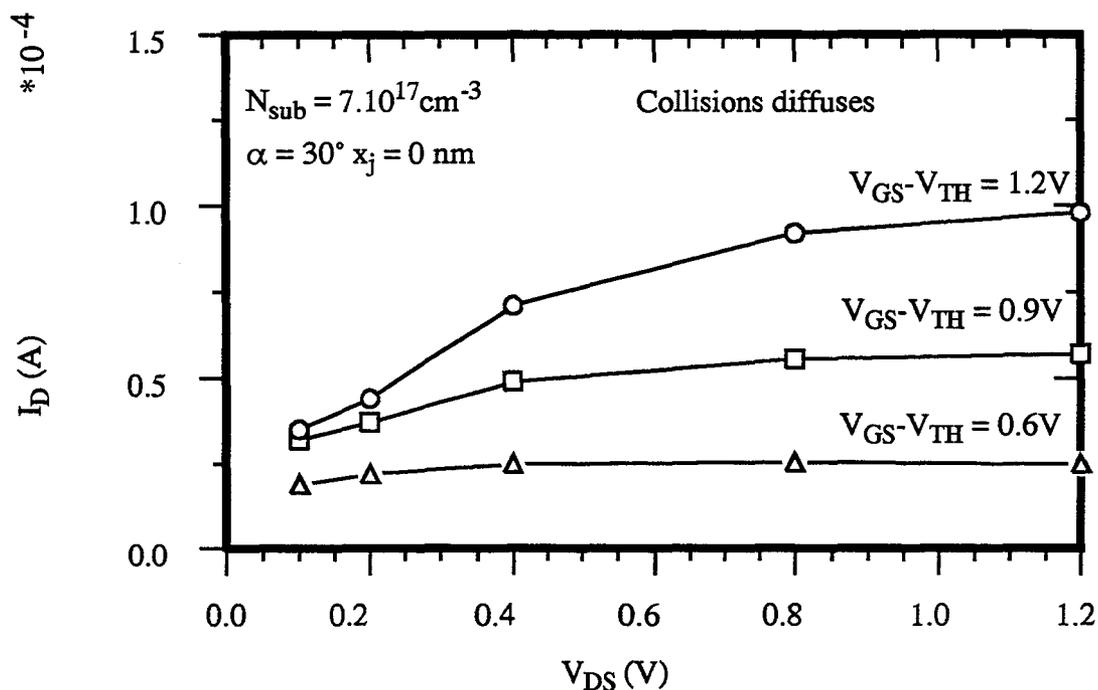


Figure V-29 : Caractéristiques  $I_D(V_{DS})$  calculées par la méthode de Monte Carlo.  $I_D$  est la valeur moyenne du courant dans le canal.

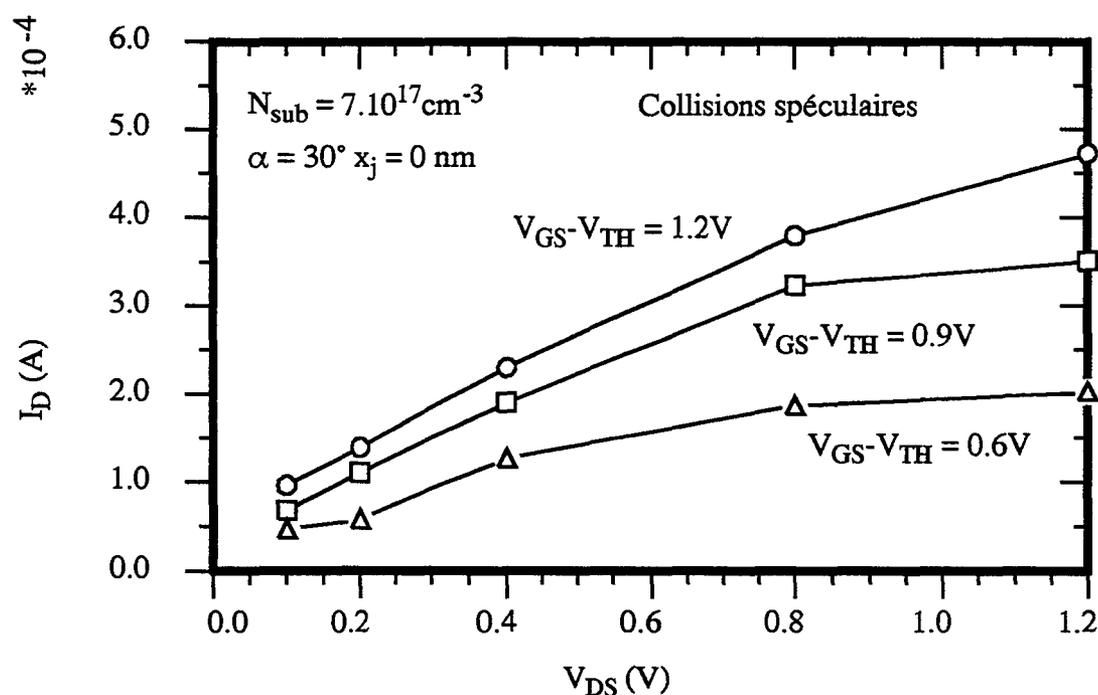


Figure V-30 : Caractéristiques  $I_D(V_{DS})$  calculées par la méthode de Monte Carlo.  $I_D$  est la valeur moyenne du courant dans le canal.

Les valeurs de courant sont dans ce cas plus grandes que celles obtenues par le formalisme classique de dérive-diffusion, environ d'un facteur deux.

L'obtention d'informations quantitatives plus précises sur le courant passe donc par une meilleure description de l'interface Si/SiO<sub>2</sub> et notamment de ses parties obliques, de manière à prendre en compte l'influence des rugosités de la surface sur la mobilité des électrons.

#### V -4 CONCLUSION

L'étude à faible polarisation montre que la mise à profit de l'effet de coin généré par la grille enterrée permet de supprimer ou d'atténuer fortement les effets indésirables de la réduction de la longueur de canal: peu de variation de la tension de seuil due au changement de longueur, pouvoir de coupure variant moins que pour les structures planes avec la taille des dispositifs. Le pouvoir de coupure est d'ailleurs meilleur que celui des dispositifs classiques pour les très faibles longueurs. L'utilisation d'un dopage modéré pour le substrat, qui a un effet bénéfique sur les caractéristiques sous le seuil, permet de plus de limiter la réduction de mobilité liée aux fortes concentrations.

L'effet de coin permet un contrôle plus efficace du canal par la grille, ce qui est confirmé également à forte polarisation. Les tendances indiquées par les simulations de dérive-diffusion classiques en terme de réduction de tension de seuil et de pouvoir de coupure sont meilleures que pour les transistors plans. Les simulations Monte Carlo montrent que les effets balistiques sont peu influencés par la longueur du canal. Ils sont pratiquement absents dans le canal, le champ électrique longitudinal ne prenant de valeur importante que dans la zone de drain.

Ces résultats font du transistor non plan une structure très intéressante pour la miniaturisation du MOS sous l'échelle de 0.1 micron.

## REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE V

- [1] : S. M. Sze, *Physics of semiconductor devices*, John Wiley & Sons, 1981
- [2] : J. P. Uyemura, *Fundamentals of MOS digital integrated circuits*, Addison-Wesley, 1988
- [3] : M. R. Wordeman, A. M. Schweighart, R. H. Dennard, G. Sai-Halasz, W. W. Molzen, *A fully scaled submicrometer NMOS technology using direct-write e-beam lithography*, IEEE Trans. Elec. Dev., vol.32, No.11, p2214, novembre 1985
- [4] : R. H. Yan, A. Ourmazd, K. F. Lee, D. Y. Jeon, C. S. Rafferty, M. R. Pinto, *Scaling the Si metal-oxide-semiconductor field-effect transistor into the 0.1- $\mu\text{m}$  regime using vertical doping engineering*, Appl. Phys. Lett. vol.59, No.25, p3315, 16 décembre 1991
- [5] : C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, B. Riccò, *A new scaling methodology for the 0.1-0.025  $\mu\text{m}$  MOSFET*, 1993 Symposium on VLSI Technology Digest of technical papers, p33, 1993
- [6] : C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, B. Riccò, *Scaling the MOS transistor below 0.1  $\mu\text{m}$ : methodology, device structures, and technology requirements*, IEEE Trans. Elec. Dev., vol.41, No.6, p941, juin 1994
- [7] : D. J. Frank, S. E. Laux, M. V. Fischetti, *Monte Carlo simulation of 30nm dual-gate MOSFET: how short can Si go?*, IEDM 92 technical digest, p553, 1992
- [8] : H. Shin, A.F. Tasch, T.J. Bordelon, C.M. Maziar, *MOSFET drain engineering analysis for deep-submicrometer dimensions: a new structural approach*, IEEE Trans. Elec. Dev., vol.39, No.8, p1922, août 1992
- [9] : W.-H. Lee, Y. J. Park, J. D. Lee, *Gate recessed (GR) MOSFET with selectively halo-doped channel and deep graded source/drain for deep submicron CMOS*, IEDM 93 technical digest, p135, 1993

- [10] : H. Noda, F. Murai, S. Kimura, *Threshold voltage controlled 0.1- $\mu\text{m}$  MOSFET utilizing inversion layer as extreme shallow source/drain*, IEDM 93 technical digest, p123, 1993
- [11] : J. Tanaka, T. Toyabe, S. Ihara, S. Kimura, H. Noda, K. Itoch, *Simulation of sub-0.1- $\mu\text{m}$  MOSFET's with completely suppressed short-channel effect*, IEEE Elec. Dev. Let., vol.14, No.8, août 1993
- [12] : J. Tanaka, S. Kimura, H. Noda, T. Toyabe, S. Ihara, *A sub-0.1- $\mu\text{m}$  grooved gate MOSFET with high immunity to short-channel effect*, IEDM 93 technical digest, p537, 1993
- [13] : S. Kimura, H. Noda, D. Hisamoto, E. Takeda, *A 0.1- $\mu\text{m}$  elevated source and drain MOSFET fabricated by phase-shifted lithography*, IEDM 91 technical digest, p950, 1991
- [14] : K. Hieda, K. Sunouchi, H. Takato, A. Nitayama, F. Horiguchi, F. Masuoka, *Sub-half-micrometer concave MOSFET with double LDD structure*, IEEE Trans. Elec. Dev., vol.39, No.3, p671, mars1992
- [15] : E. Dubois, P.H. Bricout, *Efficient scaling of MOSFETs in the sub-50 nm regime using a recessed channel*, 1993 International Semiconductor Device Research Symposium, University of Virginia, Charlottesville.
- [16] : E. Dubois, P.H. Bricout, *Short channel immunity and current drive capabilities of recessed MOSFETs in the sub-50 nm regime*, Device Reseach Conference, 1994
- [17] : H. Oyamatsu, M. Kinugawa, M. Kakumu, *Design methodology of deep submicron CMOS devices for IV operation*, 1993 Symposium on VLSI Technology Digest of technical papers, p89, 1993
- [18] : S. Selberherr, *Analysis and simulation of semiconductor devices*, Ed. Springer Verlag, 1984.

## CHAPITRE VI

### CONCLUSIONS

L'évolution des composants microélectroniques est liée à une progression régulière en terme de miniaturisation, de densité d'intégration et de vitesse. Ces progrès dépendent du développement rapide des techniques de fabrication, et notamment de la lithographie haute résolution qui permet de réaliser des structures géométriques dont les motifs peuvent descendre en dessous de  $0.1 \mu\text{m}$ .

La miniaturisation des dispositifs est également liée à une connaissance approfondie de leur fonctionnement. Cette compréhension du comportement électrique des dispositifs et des phénomènes de transport qui s'y produisent, ainsi que leur optimisation, s'appuient sur l'utilisation d'outils de simulation.

Les principaux points abordés dans le cadre de ce travail sont résumés ici:

\* Une première partie porte sur la description du transport électronique dans le silicium par l'équation de Boltzmann et sur les méthodes de résolution couramment utilisées en simulation de dispositifs. Suite à la comparaison de ces différentes méthodes, un simulateur de dispositifs nMOS par la technique de Monte Carlo a été élaboré. En effet, la réduction des dimensions fait apparaître des champs électriques intenses, présentant de brusques variations spatiales et temporelles, ce qui introduit des effets balistiques que la méthode de Monte Carlo permet d'appréhender de façon satisfaisante.

\* Le simulateur Monte Carlo a été validé d'une part en étudiant le phénomène de survitesse dans un matériau homogène où les électrons sont soumis à un échelon de champ électrique,

et d'autre part en comparant les résultats de simulations de dispositifs nMOS submicroniques avec des données publiées [1].

\* L'approche de type Monte Carlo a été complétée par la mise au point d'une technique de couplage avec la résolution d'équations étendues de dérive-diffusion. Outre une interprétation du transport en termes usuels de champs subis par les électrons, de mobilité, de tension thermique et de diffusivité, le couplage permet une mise à jour rapide des concentrations de porteurs et la prise en compte de phénomènes intervenant à des échelles de temps différentes.

\* La méthode de couplage a été appliquée à la simulation de dispositifs fortement submicroniques, montrant que le phénomène de survitesse peut être mis à profit pour améliorer certaines caractéristiques électriques et notamment augmenter la transconductance. Un autre point important pour les effets liés aux porteurs chauds est que l'énergie maximale atteinte par les électrons diminue lorsque les dispositifs sont miniaturisés, dans la mesure où les tensions de polarisation sont réduites.

\* La mise au point d'une stratégie de miniaturisation en deçà de 0.1 micron se heurte à des limites technologiques. L'épaisseur d'oxyde peut difficilement être réduite en dessous de 3 nm afin d'éviter le claquage par effet tunnel. La garantie de marges de bruit suffisantes pour les portes logiques interdit de réduire les tensions d'alimentation en dessous de 1.5-1.0 V [2] [3]. Le recours à une structure de transistor non plane, à grille enterrée, permet de réduire de façon importante les effets canaux courts par rapport aux dispositifs plans classiques. Ceci est rendu possible par un meilleur contrôle du canal par la grille, dû aux barrières de potentiel générées aux angles de l'interface Si/SiO<sub>2</sub>. Un autre avantage de cette structure est que le dopage de substrat peut conserver des valeurs modérées, ce qui évite la réduction de mobilité causée par les fortes concentrations.

Quelques constats effectués au cours des études présentées ci-dessus permettent de dégager des perspectives à moyen terme:

\* Les modèles de collisions sur impuretés ionisées utilisés lors des simulations Monte Carlo demandent à être affinés pour les fortes concentrations. Ce point est important pour l'obtention d'informations quantitatives sur les caractéristiques électriques des dispositifs submicroniques, où de forts dopages sont employés.

\* De même, la description des rugosités et des défauts de l'interface Si/SiO<sub>2</sub> et la prise en compte de leur influence sur la mobilité peuvent être améliorés. Ces éléments permettront de poursuivre l'étude du dispositif à grille enterrée comme structure susceptible d'être miniaturisée au dessous de 0.1 micron de longueur effective.

\* Il est possible d'appliquer aux transistors pMOS la méthode de couplage Monte Carlo/dérive-diffusion étendu. Il est nécessaire pour cela d'utiliser une modélisation de la structure de bande de valence du silicium ainsi que des probabilités d'interaction correspondantes [4]. La transposition des équations étendues de dérive-diffusion au transport par les trous est par contre immédiate.

\* Dans la cadre d'une utilisation intensive des outils de simulation pour l'optimisation de composants, il est intéressant d'automatiser au maximum la procédure de couplage entre les programmes Monte Carlo et de dérive-diffusion.

D'autres remarques constituent des pistes à plus long terme:

\* La limite à laquelle se heurte la réduction de l'épaisseur d'oxyde de grille pousse à envisager d'autres matériaux isolants, notamment des molécules organiques [5].

\* Les forts champs électriques qui apparaissent sous la grille des transistors MOS peuvent se traduire par des effets quantiques non négligeables, qui devront être considérés dans le fonctionnement des dispositifs.

\* La prise en compte des effets latéraux sur les caractéristiques électriques des composants submicroniques devra passer par le développement d'outils de simulation tridimensionnels, avec les difficultés de modélisation, de maillage et de discrétisation que cela suppose. Des

travaux dans ce sens ont déjà été effectués [6], et le recours aux simulateurs 3D devrait se répandre lors des prochaines années.

\* La miniaturisation des composants à l'échelle nanométrique s'accompagne d'une réduction du nombre de porteurs présents dans la zone active des dispositifs. Le nombre des électrons dans la bande de conduction d'un transistor MOS de 0.1 micron de longueur effective et de 1 micron de largeur est ainsi de quelques dizaines de milliers. L'équation de Boltzmann, qui fournit une description statistique du transport pourrait ainsi se trouver en défaut. D'autres limites de validité de la formulation de l'équation de Boltzmann utilisée entre autres par la méthode de Monte Carlo sont atteintes sous l'échelle nanométrique. Les interactions subies par les électrons ne peuvent plus être considérées comme instantanées. De plus, la séparation entre forces macroscopiques et interactions n'est plus valide si le champ électrique présente de fortes variations sur une distance de l'ordre de la longueur d'onde de de Broglie de l'électron. C'est là le fondement même des méthodes usuelles de simulation de dispositifs qui est mis en cause. La modélisation des composants futurs, qui auront des principes de fonctionnement vraisemblablement différents des transistors actuels, devra intégrer des éléments de physique encore plus fondamentaux.

## REFERENCES BIBLIOGRAPHIQUES DU CHAPITRE VI

- [1] : M. V. Fischetti, S. E. Laux, D. J. Frank, *Monte Carlo analysis of semiconductor devices: the Damocles program*, IBM J. Res. Develop. vol.34, No.4, p466, juillet 1990
- [2] : C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, B. Riccò, *A new scaling methodology for the 0.1-0.025  $\mu\text{m}$  MOSFET*, 1993 Symposium on VLSI Technology Digest of technical papers, p33, 1993
- [3] : C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, B. Riccò, *Scaling the MOS transistor below 0.1  $\mu\text{m}$ : methodology, device structures, and technology requirements*, IEEE Trans. Elec. Dev., vol.41, No.6, p941, juin 1994
- [4] : C. Jacoboni, P. Lugli, *The Monte Carlo method for semiconductor device simulation*, Ed. S. Selberherr, Springer-Verlag, 1989
- [5] : P. Fontaine, D. Goguenheim, D. Deresmes, D. Vuillaume, M. Garet, F. Rondelez, *Octadecyltrichlorosilane monolayers as ultrathin gate insulating films in metal-insulator-semiconductor devices*, Appl. Phys. Lett. vol.62, No.18, p2256, mai 1993
- [6] : D.C. Cole, E.M. Buturla, S.S. Furkay, K. Varahramyan, J. Slinkman, J.A. Mandelman, D.P. Foty, O. Bula, A.W. Strong, J.W. Park, T.D. Linton Jr, J.B. Johnson, M.V. Fischetti, S.E. Laux, P.E. Cottrell, H.G. Lustig, F. Pileggi, D.Katcoff, *The use of simulation in semiconductor technology development*, Solid State Elec., vol.33, No.6, p591-623, 1990

