

Modélisation 3D de MOSFETs multi-grilles

La diminution effrénée des dimensions des transistors conduit aujourd'hui à des MOSFETs ultimes SOI de taille nanométrique. Des composants présentant des longueurs de grille aussi petites que 5 et 6 nm [194, 195] ont d'ailleurs été fabriqués. Dans ce contexte, les MOSFETs multi-grilles constituent des candidats intéressants grâce à l'efficacité du contrôle des charges du canal par la tension de grille. Plusieurs groupes ont récemment réalisés des MOSFETs Tri-gate et Gate-all-around (cf § (1.3.3)) à partir de nano-fils de silicium à section rectangulaire [49,196]. Contrairement aux MOSFETs "bulk" et aux double-grilles plan, les multi-grilles ont des architectures 3D dans lesquelles longueur, hauteur et profondeur jouent un rôle principal. Jusqu'à présent, les dimensions transverses des nano-fils n'atteignent pas les 20 nm et la compréhension des composants ultra-minces (section < 10 nm) constitue un défi majeur.

Au chapitre 4 nous avons présenté une solution balistique d'un MOSFET double-grille simplifié dans lequel l'épaisseur du canal était réduite à une chaîne atomique : l'influence du confinement vertical, notamment pour obtenir l'électroneutralité des extensions des réservoirs, n'était pas prise en compte. L'objectif de ce chapitre est de modéliser des architectures innovantes 3D réalistes à l'aide du formalisme des fonctions de GREEN hors-équilibre. Afin de simplifier les étapes numériques, nous adopterons l'approche dite de l'"espace des modes" (pour l'anglais "mode-space"), qui consiste à séparer le traitement des effets du confinement transverse de celui du transport électronique le long de l'axe source-drain (section 5.1). Après avoir appliqué ce modèle au MOSFET Gate-All-Around en détaillant l'évolution des paramètres physiques internes (section 5.2), nous décrirons l'ensemble des architectures 3D (Tri-gate, Π -gate, Ω -gate) et étudierons leur comportement électrique (section 5.3). Enfin nous concluons.

5.1 Approximation de l'espace des modes

La résolution quantique exacte de systèmes 2D ou 3D nécessite un outil de calcul numérique très puissant. De plus en plus de laboratoires et industries s'équipent donc de *cluster*, c'est-à-dire d'une grappe d'ordinateurs connectés entre eux (jusqu'à 150 unités) et travaillant en parallèle sur la même tâche. Cette évolution a incité les chercheurs à proposer des alternatives physiques et/ou numériques susceptibles de réduire la taille des calculs. L'approximation de l'espace des modes est le fruit de cette volonté et permet de transformer, sous certaines conditions, un problème 2D ou 3D en plusieurs problèmes 1D indépendants, et traitables par un simple PC. La section suivante décrit l'approche "mode-space" et définit son domaine de validité.

5.1.1 Confinement quantique dans les architectures MOSFETs innovantes

Les prévisions des industriels indiquent que la géométrie SOI ("Silicon-On-Insulator") de par sa bonne immunité contre les effets canaux courts (cf § (1.2)), pourrait être le dispositif des futures technologies CMOS. Néanmoins, la diminution de la longueur de grille d'une génération de transistors à une autre, s'accompagne également de la réduction des dimensions transverses afin de contenir la progression des effets canaux courts. Le confinement des porteurs de charges dans la section transverse qui en résulte induit une quantification des états perpendiculaires et une distribution de ces derniers en sous-bandes. Ce comportement modifie ostensiblement les caractéristiques électriques des MOSFETs SOI à canaux courts et l'évaluation quantitative de son impact constitue une donnée essentielle.

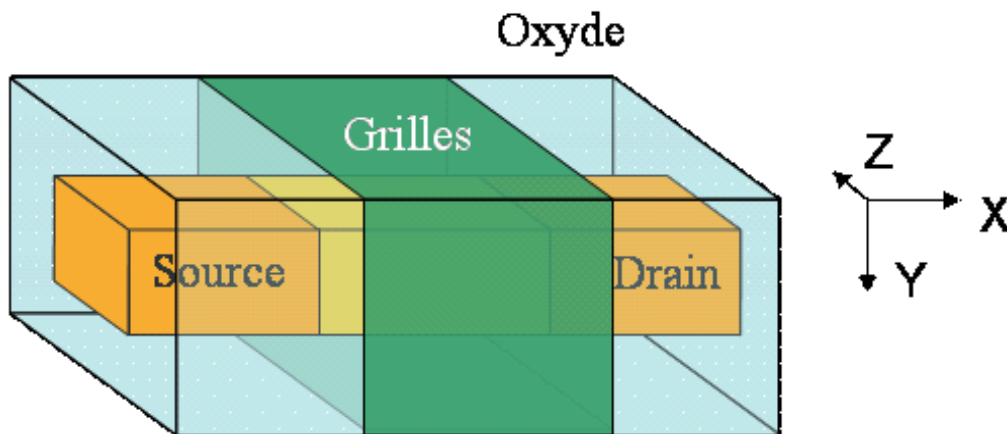


FIG. 5.1: Schéma d'un transistor MOSFET *Gate-All-Around* (GAA).

Considérons le transistor MOSFET *Gate-all-around* représenté sur la figure (5.1). Aux petites dimensions, les effets quantiques affectent sa tension de seuil (confinement suivant les axes y et z) et son courant *off* (effet tunnel entre source et drain). Pour traiter un tel système, les approches de modélisation quantique reposent sur une résolution auto-cohérente des équations

de SCHRÖDINGER et de POISSON. La solution de l'équation de SCHRÖDINGER peut être obtenue suivant plusieurs niveaux de complexité en fonction de la nature du dispositif étudié et du degré d'exactitude désiré. Dans le cas de transistors *bulk*, deux groupes ont développé des solutions 2D basées sur le formalisme des fonctions de GREEN hors-équilibre [197, 198, 200]. Leurs approches, qui reposent sur une discrétisation dans l'espace réel de l'hamiltonien 2D exprimé dans l'approximation de la masse effective, sont numériquement très coûteuses. La discrétisation de l'espace réel est néanmoins inévitable pour traiter correctement le transport quantique dans les MOSFETs massifs, car le confinement des porteurs évolue le long du canal et aucune symétrie ne permet de réduire la taille du problème. Il en est tout autrement pour les MOSFETs SOI ultimes, dans lesquels les charges mobiles restent confinées quantiquement tout le long du canal. Ces géométries permettent de contourner la difficulté numérique associée à la résolution dans l'espace réel (sans compromettre l'exactitude du résultat) en exprimant la partie de l'hamiltonien confiné dans l'espace des modes et en limitant les calculs aux seuls modes occupés (≈ 2 ou 3 sous-bandes par vallée du silicium). Comparée avec succès à une résolution complète de l'espace réel dans le cas du transistor SOI double-grille 2D [104, 199] l'approche "mode-space" a été récemment appliquée au transistor à nano-fil de silicium 3D [201, 202]. Afin de justifier intuitivement la solution de l'espace des modes, considérons un plan atomique vertical (\mathbf{y}, \mathbf{z}) du transistor de la figure (5.1) à \mathbf{x} fixé. Les porteurs de ce plan, fortement confinés suivant les deux directions \mathbf{y} et \mathbf{z} , sont distribués dans un réseau discret de sous-bandes. Les énergies des états transverses sont quantifiées et leur éloignement croît avec le confinement. On peut ainsi supposer qu'un transistor GAA ultime, présente un espacement entre sous-bandes suffisant pour ignorer le couplage inter-bande. Chaque sous-bande est ainsi traitée séparément, et le système initialement 3D devient un problème N fois 1D où N représente le nombre de sous-bandes occupées.

5.1.2 Théorie

Considérons le MOSFET GAA représenté sur la figure (5.1). L'interface Si/SiO₂ est perpendiculaire à la direction cristalline $\langle 100 \rangle$. Le système, traité en liaisons fortes, est supposé en première approximation présenter une structure cubique simple (figure (5.2)) et la distance inter-atomique a est identique suivant toutes les directions ($a = 0.2$ nm). Chaque atome possède une orbitale et seules les interactions entre premiers voisins sont envisagées. Cette description est équivalente à basse énergie à l'approximation de la masse effective (cf § (4.1.2)), mais permet de faire apparaître la structure atomique du composant (bien que de façon simpliste) et de modéliser l'influence de défauts ponctuels (cf chapitre 6). La fonction de GREEN est maintenant calculée afin d'obtenir la densité électronique dans la zone active et le courant de drain. En régime balistique la méthode des fonctions de GREEN est équivalente à la résolution directe de l'équation de SCHRÖDINGER avec des conditions aux limites ouvertes. Une description quantique 3D exacte étant numériquement peu envisageable, nous adoptons la représentation de l'espace des modes dans les directions du confinement des grilles. Les différentes étapes de cette approche, dont les conditions de validité seront explicitées par la suite, sont ici décrites. Nous résolvons tout d'abord l'hamiltonien 2D H_{conf} dans chaque plan atomique n perpendiculaire à

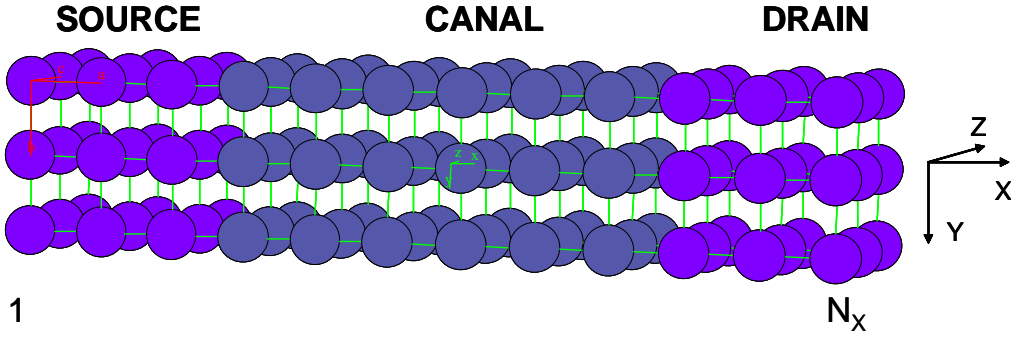


FIG. 5.2: *Le silicium et son oxyde sont supposés, en première approximation, présenter une structure cubique simple. Cette approche, équivalente à l'approximation de la masse effective, nous permettra de mesurer l'impact de défauts ponctuels (cf chapitre 6). Par volonté de clarté la structure des oxydes de grille n'est pas représentée.*

l'axe source-drain :

$$H_{conf} |\psi_{n,i}\rangle = \varepsilon_{n,i} |\psi_{n,i}\rangle, \quad (5.1)$$

avec :

$$\begin{aligned} H_{conf} &= \sum_{l,m} V_{lmn,lmn} |l, m, n\rangle \langle l, m, n| \\ &+ \beta_y \sum_{l,m} |l, m, n\rangle \langle l, m+1, n| \\ &+ \beta_z \sum_{l,m} |l, m, n\rangle \langle l+1, m, n|, \end{aligned} \quad (5.2)$$

où $|l, m, n\rangle$ est le ket représentant l'orbitale de l'atome de coordonnées $z = l \times a$, $y = m \times a$ du plan vertical n . $V_{lmn,lmn}$ représente l'élément de matrice du potentiel électrostatique supposé diagonal dans la base atomique conformément à la discussion de la section (4.1.2) sur la méthode des liaisons fortes. $|\psi_{n,i}\rangle$ et $\varepsilon_{n,i}$ sont respectivement les vecteurs et les valeurs propres des plans transverses. Enfin β_y et β_z sont les couplages entre premiers voisins le long des directions correspondantes. Elles sont reliées aux masses effectives par les expressions suivantes (cf § (2.3.2)) :

$$\beta_y = -\frac{\hbar^2}{2m_y^* a^2}, \text{ et } \beta_z = -\frac{\hbar^2}{2m_z^* a^2}. \quad (5.3)$$

La structure cubique simple est étendue aux oxydes de grille dont la description est incluse dans l'hamiltonien H_{conf} autorisant ainsi la pénétration de la fonction d'onde $|\psi_{n,i}\rangle$. Les mêmes couplages β_y et β_z sont considérés puisqu'il a été vérifié que la différence de gap entre le Si et le SiO₂ déterminait pour l'essentiel la pénétration de l'état transverse dans l'oxyde. Les éléments de matrice du potentiel électrostatique représentant l'oxyde sont donc augmentés de 3 eV par

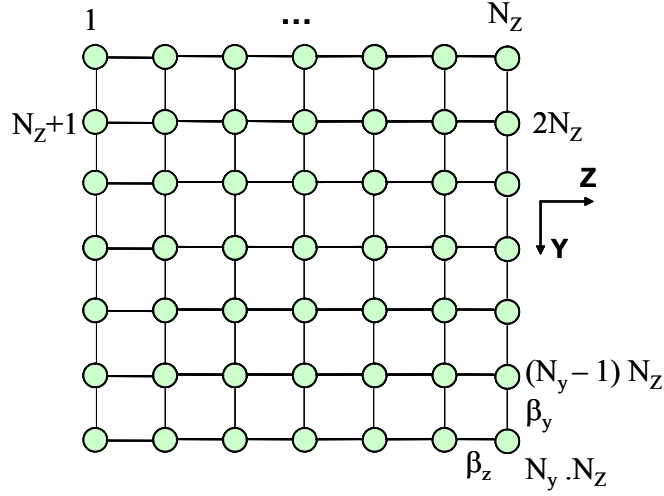


FIG. 5.3: Représentation schématique et numérotation des atomes de la section transverse du MOSFET GAA possédant N_Y atomes suivant la direction \mathbf{y} et N_Z suivant la direction \mathbf{z} . Les énergies de couplage des oxydes ont une influence négligeable face à la barrière de potentiel d'interface et sont donc choisies égales à celles du silicium [203].

rapport à ceux du silicium et l'expression matricielle de l'équation (5.1) s'écrit :

$$H_{conf} = \begin{bmatrix} V_{1,n} & \beta_z & 0 & \cdots & 0 & \beta_y & 0 & \cdots & 0 \\ \beta_z & V_{2,n} & \beta_z & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \beta_z & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & V_{i,n} & \ddots & \ddots & \ddots & 0 & \beta_y \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \beta_y & \ddots & \ddots & \ddots & \ddots & V_{N_Z+1,n} & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \beta_z & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \beta_z & V_{(N_Y \cdot N_Z)-1,n} & \beta_z \\ 0 & \cdots & 0 & \beta_y & 0 & \cdots & 0 & \beta_z & V_{N_Y \cdot N_Z,n} \end{bmatrix}, \quad (5.4)$$

en utilisant les notations de la figure (5.3). L'approximation de l'espace des modes consiste ensuite à représenter l'hamiltonien 3D en considérant que les sous-bandes sont indépendantes. En d'autres termes, l'interaction entre $|\psi_{n,i}\rangle$ d'indice i différent est négligée :

$$H_{3D} = \begin{bmatrix} h_1 & 0 & \cdots & \cdots & \cdots \\ 0 & h_2 & 0 & \cdots & \cdots \\ \cdots & 0 & \ddots & 0 & \cdots \\ \cdots & \cdots & 0 & h_i & 0 \\ \cdots & \cdots & \cdots & 0 & \ddots \end{bmatrix}, \quad (5.5)$$

où

$$h_i = \begin{bmatrix} \varepsilon_{1,i} & \beta_x & 0 & \cdots \\ \beta_x & \varepsilon_{2,i} & \beta_x & 0 \\ 0 & \beta_x & \ddots & \beta_x \\ \cdots & 0 & \beta_x & \varepsilon_{N_X,i} \end{bmatrix} \quad (5.6)$$

est l'hamiltonien de la sous-bande i , N_X est le nombre d'atomes le long de l'axe source-drain et β_x est l'énergie de couplage suivant \mathbf{x} reliée à la masse effective par (cf § (4.1.2)) :

$$\beta_x = -\frac{\hbar^2}{2m_x^* a^2}. \quad (5.7)$$

En théorie, les hamiltoniens h_i de toutes les sous-bandes devraient être calculés, mais concrètement seules les premières sous-bandes sont occupées et participent à la conduction électronique. Il en résulte un gain numérique substantiel puisque le problème 3D de dimensions $(N_X \times N_Y \times N_Z)^2$ se transforme en un problème N^1 fois 1D dont la dimension est définie par le nombre d'atomes suivant l'axe \mathbf{x} ($= N_X$)².

Connaissant les hamiltoniens des sous-bandes peuplées (h_i), le transport 1D à travers chaque sous-bande est calculé. La fonction de GREEN retardée de la sous-bande i s'exprime comme :

$$G_i(\varepsilon) = \lim_{\eta \rightarrow 0^+} [(\varepsilon + i\eta)I - h_i - \Sigma^i], \quad (5.8)$$

où Σ^i est la matrice de la self-énergie des contacts qui modélise l'élargissement des niveaux d'énergie du système fini lors du couplage avec les réservoirs d'électrons. Le transport étant toujours 1D, leur expression reste identique à celle du § (4.2.1) au décalage en énergie près (confinement transverse). On obtient donc :

$$\Sigma^i = \Sigma_S^i + \Sigma_D^i. \quad (5.9)$$

Σ_S^i et Σ_D^i s'expriment en fonction de l'énergie de couplage réservoir-zone active ($= \beta_x$) et de la fonction de GREEN de surface des réservoirs. En reprenant les équations (4.34) et (4.36) la self-énergie de la source s'exprime de la façon suivante :

$$\Sigma_S^i = \begin{bmatrix} \beta_x g_S^i \beta_x & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}, \quad (5.10)$$

où g_S^i est la fonction de GREEN du contact de la source qui représente une chaîne semi-infinie

¹ N représente le nombre de sous-bandes.

dont l'énergie potentielle est ajustée à celle de la sous-bande i :

$$g_S^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{1,i}}{2\beta_x^2} \left(1 - \sqrt{1 - \frac{4\beta_x^2}{(\varepsilon + \varepsilon_{1,i})^2}} \right) \quad \text{si } |\varepsilon + \varepsilon_{1,i}| > 2|\beta_x|, \quad (5.11)$$

$$g_S^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{1,i}}{2\beta_x^2} - \frac{i}{|\beta_x|} \sqrt{1 - \left| \frac{\varepsilon + \varepsilon_{1,i}}{2\beta_x} \right|^2} \quad \text{si } |\varepsilon + \varepsilon_{1,i}| < 2|\beta_x|. \quad (5.12)$$

De façon similaire la self-énergie du drain s'écrit :

$$\Sigma_D^i = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & \beta_x g_D^i \beta_x \end{bmatrix}, \quad (5.13)$$

où g_D^i est la fonction de GREEN du contact de drain :

$$g_D^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{N_X,i}}{2\beta_x^2} \left(1 - \sqrt{1 - \frac{4\beta_x^2}{(\varepsilon + \varepsilon_{N_X,i})^2}} \right) \quad \text{si } |\varepsilon + \varepsilon_{N_X,i}| > 2|\beta_x|, \quad (5.14)$$

$$g_D^i(\varepsilon) = \frac{\varepsilon + \varepsilon_{N_X,i}}{2\beta_x^2} - \frac{i}{|\beta_x|} \sqrt{1 - \left| \frac{\varepsilon + \varepsilon_{N_X,i}}{2\beta_x} \right|^2} \quad \text{si } |\varepsilon + \varepsilon_{N_X,i}| < 2|\beta_x|. \quad (5.15)$$

De telles self-énergies permettent de réaliser des contacts parfaits exempts de réflexion.

La fonction de GREEN de la sous-bande i obtenue, nous définissons deux nouvelles quantités :

$$\Gamma_S^i = i \left(\Sigma_S^i - (\Sigma_S^i)^\dagger \right) \quad \text{et} \quad \Gamma_D^i = i \left(\Sigma_D^i - (\Sigma_D^i)^\dagger \right), \quad (5.16)$$

qui ont la dimension d'une énergie et représentent la force d'interaction entre les réservoirs et la zone active. Les fonctions de la densité spectrale des contacts s'expriment sous la forme :

$$A_S^i = G_i \Gamma_S^i G_i \quad \text{et} \quad A_D^i = G_i \Gamma_D^i G_i. \quad (5.17)$$

Ce sont des matrices de dimensions (N_X^2) dont les termes diagonaux représentent la densité locale d'états de chaque atome. Bien que le système soit dans un état hors-équilibre, les électrons injectés depuis les réservoirs de la source et du drain conservent leur distribution initiale (transport balistique). La fonction spectrale de la source est donc occupée suivant la distribution de FERMÍ-DIRAC de la source (f_S) alors que la fonction spectrale du drain est occupée suivant la distribution de FERMÍ-DIRAC du drain (f_D).

La matrice de la densité d'électrons 1D associée à la sous-bande i (le long de l'axe source-drain) est définie par :

$$n_i(\varepsilon) = \frac{1}{\pi a} \times [f_S(\varepsilon) A_S^i + f_D(\varepsilon) A_D^i]. \quad (5.18)$$

Pour obtenir la densité électronique totale 3D, il suffit d'intégrer l'équation (5.18) sur l'énergie ε , de sommer les contributions de chaque vallée et sous-bande, de les multiplier par le module au carré de l'état transverse associé $|\psi_{n,i}(y,z)|^2$ et de répéter l'opération pour tous les plans n :

$$n_{l,m,n}^{3D} = \sum_{l,m,n} \sum_i |\langle l, m, n | \psi_{n,i} \rangle|^2 \int_0^{+\infty} n_i(\varepsilon) d\varepsilon, \quad (5.19)$$

où i représente la somme sur les premières sous-bandes de chaque vallée. La densité électronique rétro-agit ensuite sur l'équation de POISSON 3D² afin d'obtenir un système auto-cohérent.

L'auto-cohérence atteinte, le courant de drain est calculé en sommant les courants associés à tous les modes peuplés. Pour ce faire, le système est clivé en deux régions et le courant traversant le plan séparant ces deux régions est déterminé. La matrice de la fonction de GREEN retardée de la sous-bande i du système non-couplé s'écrit :

$$G_i^0(\varepsilon) = \lim_{\eta \rightarrow 0^+} \begin{bmatrix} (\varepsilon + i\eta)I_{11} - (h_i)_{11} - \Sigma_S^i & 0 \\ 0 & (\varepsilon + i\eta)I_{22} - (h_i)_{22} - \Sigma_D^i \end{bmatrix} \quad (5.20)$$

où $(h_i)_{11}$ (respectivement $(h_i)_{22}$) est la matrice de l'hamiltonien dans la région 1 (respectivement dans la région 2) du mode i . Soit C_i la matrice de couplage assurant la connexion pour la sous-bande i entre les régions 1 et 2 :

$$C_i = \begin{bmatrix} 0 & (h_i)_{12} \\ (h_i)_{21} & 0 \end{bmatrix}, \quad (5.21)$$

où $(h_i)_{12}$ et $(h_i)_{21}$ sont les matrices des blocs non-diagonaux de l'hamiltonien du mode i . En se reportant au § (3.3.3), l'expression finale du courant³ traversant la sous-bande i est :

$$I_i = -\frac{4\pi e}{\hbar} \int_{-\infty}^{+\infty} d\varepsilon Tr_1 \left[N_0^i \Lambda_i^\dagger C_i N_0^i C_i \Lambda_i \right] (f_S(\varepsilon) - f_D(\varepsilon)), \quad (5.22)$$

où Tr_1 désigne la trace restreinte à la base des orbitales de la région 1, $N_0^i (= -(1/\pi) \text{Im}(G_0^i))$ est la matrice densité à l'équilibre du système non-couplé, et $\Lambda_i = (I - G_0^i C_i G_0^i C_i)^{-1}$. Enfin, le courant total s'obtient en sommant sur les premiers modes de chaque vallée :

$$I = \sum_i I_i. \quad (5.23)$$

Nous modélisons le cas du Si par six bandes de liaisons fortes dont la partie basse énergie doit reproduire l'approximation de la masse effective. La figure (5.4) représente donc les six ellipsoïdes équivalents du silicium volumique. Le confinement suivant \mathbf{y} et \mathbf{z} , sépare les six vallées en trois

²Les conditions aux limites de l'équation de POISSON 3D seront discutées au § (5.1.4).

³L'équation (5.22) est la généralisation de la formule de LANDAUER (cf § (3.3.1)) qui relie le courant au coefficient de transmission $T(\varepsilon) = 4\pi^2 Tr_1 [N_0 \Lambda^\dagger C N_0 C \Lambda]$.

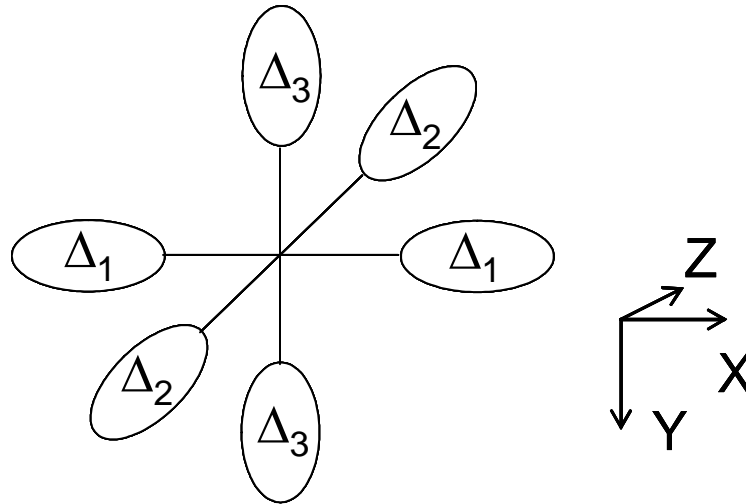


FIG. 5.4: Représentation des six ellipsoïdes équivalents du silicium volumique. Sous l'influence du confinement, trois paires prennent des énergies différentes.

paires dont les masses effectives dans les trois directions sont :

· Vallées $\Delta 1$:

$$- m_x^* = m_l, \text{ et } m_y^* = m_z^* = m_t,$$

· Vallées $\Delta 2$:

$$- m_x^* = m_z^* = m_t, \text{ et } m_y^* = m_l,$$

· Vallées $\Delta 3$:

$$- m_x^* = m_y^* = m_t, \text{ et } m_z^* = m_l.$$

Les masses effectives des vallées $\Delta 1$ dans les directions du confinement sont égales à la masse transverse m_t , qui est environ cinq fois plus faible que la masse longitudinale m_l . Les niveaux d'énergie longitudinale (suivant \mathbf{x}) associés à ces vallées sont plus élevés et ne participent donc pas à la conduction des électrons⁴. En conséquence seules les vallées $\Delta 2$ et $\Delta 3$ seront considérées dans la suite de l'étude.

5.1.3 Justification de l'approche "mode-space"

Jusqu'à présent, nous avons détaillé la méthode de l'espace des modes sans justifier sa validité. Il est cependant essentiel de connaître les conditions sous lesquelles cette approche offre une modélisation correcte du transport électronique. Dans la partie suivante, nous exprimons l'élément de matrice de H_{3D} entre deux états transverses appartenant à différents plans atomiques et soulignons les approximations faites pour simplifier l'hamiltonien.

⁴Rappel : Dans l'approximation de la masse effective, les niveaux d'énergie d'un puits infini 1D s'écrivent : $\varepsilon_i = (\hbar^2 k_i^2)/(2m^*)$. Une faible masse effective augmente donc l'espace entre niveaux d'énergie.

Définissons $|\psi_{n,i}\rangle$ comme le $i^{\text{ème}}$ état transverse du $n^{\text{ème}}$ plan atomique. L'élément de matrice de l'hamiltonien avec $|\psi_{n+1,i'}\rangle$ peut s'écrire :

$$\langle \psi_{n+1,i'} | H_{3D} | \psi_{n,i} \rangle = \beta_x \langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle, \quad (5.24)$$

où T_+ représente la translation d'ensemble de l'état $|\psi_{n,i}\rangle$ du plan n au plan $n+1$. Si le potentiel électrostatique le long de l'axe \mathbf{x} n'est pas constant, c'est-à-dire si (V_n est le potentiel du plan n) :

$$T_+ V_n \neq V_{n+1}, \quad (5.25)$$

l'application de la théorie des perturbations au premier ordre à l'état $|T_+ \psi_{n,i'}\rangle$ donne⁵ :

$$|\psi_{n+1,i'}\rangle = |T_+ \psi_{n,i'}\rangle + \sum_{k \neq i'} \frac{\langle T_+ \psi_{n,k} | V_{n+1} - T_+ V_n | T_+ \psi_{n,i'} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,k}} |T_+ \psi_{n,k}\rangle. \quad (5.28)$$

Le produit scalaire $\langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle$ s'écrit donc :

$$\begin{aligned} \langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle &= \langle T_+ \psi_{n,i'} | T_+ \psi_{n,i} \rangle + \sum_{k \neq i'} \frac{\langle T_+ \psi_{n,i'} | V_{n+1} - T_+ V_n | T_+ \psi_{n,k} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,k}} \langle T_+ \psi_{n,k} | T_+ \psi_{n,i} \rangle, \\ &= \delta_{i'i} + \frac{\langle T_+ \psi_{n,i'} | V_{n+1} - T_+ V_n | T_+ \psi_{n,i} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,i}} (1 - \delta_{i'i}). \end{aligned} \quad (5.29)$$

Considérant deux états de sous-bandes différentes ($i \neq i'$), le terme de couplage de l'hamiltonien (5.24) devient :

$$\begin{aligned} \langle \psi_{n+1,i'} | H_{3D} | \psi_{n,i} \rangle &= \beta_x \langle \psi_{n+1,i'} | T_+ \psi_{n,i} \rangle, \\ &= \beta_x \times \frac{\langle T_+ \psi_{n,i'} | V_{n+1} - T_+ V_n | T_+ \psi_{n,i} \rangle}{\varepsilon_{n,i'} - \varepsilon_{n,i}}, \end{aligned} \quad (5.30)$$

qui est d'autant plus faible que la variation du potentiel électrostatique suivant l'axe source-drain diminue. Si le potentiel est constant dans cette direction, l'équation (5.30) indique que les sous-bandes engendrées par le confinement latéral sont réellement indépendantes et que la matrice de l'hamiltonien du système peut être représentée en sous-blocs diagonaux. Les sous-bandes sont parfaitement découplées les unes des autres et le problème initial à trois dimensions se décompose en plusieurs problèmes 1D indépendants. Cette hypothèse reste valable tant que les variations du profil du potentiel électrostatique d'une section transverse à la suivante sont négligeables. Par exemple, si V_n conserve la même forme, mais qu'il prend différentes valeurs le

⁵Rappel : la correction au premier ordre du $n^{\text{ème}}$ état non perturbé $|\psi_n^{(0)}\rangle$ sous l'effet d'un potentiel V s'écrit :

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + |\psi_n^{(1)}\rangle, \quad (5.26)$$

avec :

$$|\psi_n^{(1)}\rangle = \sum_{k \neq n} \frac{\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle}{\varepsilon_n^{(0)} - \varepsilon_k^{(0)}} |\psi_k^{(0)}\rangle, \quad (5.27)$$

où les $\varepsilon_k^{(0)}$ représentent les énergies propres associées aux états d'ordre 0 : $|\psi_k^{(0)}\rangle$.

long de l'axe \mathbf{x} , les fonctions propres $|\psi_{n,i}\rangle$ restent inchangées bien que les valeurs propres $\varepsilon_{n,i}$ varient.

Pour les transistors SOI, dans lesquels la partie en silicium est uniformément "sandwichée" entre des grilles isolantes, la variation transverse du potentiel est très faible et l'approximation des sous-bandes indépendantes est justifiée. Des vérifications numériques ont été effectuées avec succès dans le cas d'un transistor double-grille 2D [104]. En revanche, les transistors *bulk* ou les SOI dont les dimensions transverses du canal de conduction ne sont pas constantes ne peuvent être modélisés par ce type d'approche.

5.1.4 Résolution de l'équation de Poisson 3D et auto-cohérence

L'équation de POISSON 3D est résolue en utilisant la méthode des différences finies décrite au § (4.2.1) dans le cas à deux dimensions. Le maillage uniforme est ici étendu à la direction \mathbf{z} avec un pas de discrétisation identique à celui de la structure cubique simple adoptée ici pour simuler le silicium ($a = 0.2 \text{ nm}$). Utilisant un développement de TAYLOR au second ordre, les dérivées secondes du potentiel électrostatique $V(x, y, z)$ s'expriment dans le formalisme des différences finies sous la forme :

$$\begin{aligned} \frac{\partial^2 V(x, y, z)}{\partial x^2} &\approx \frac{V(x + \Delta x, y, z) + V(x - \Delta x, y, z) - 2V(x, y, z)}{\Delta x^2}, \\ \frac{\partial^2 V(x, y, z)}{\partial y^2} &\approx \frac{V(x, y + \Delta y, z) + V(x, y - \Delta y, z) - 2V(x, y, z)}{\Delta y^2}, \\ \frac{\partial^2 V(x, y, z)}{\partial z^2} &\approx \frac{V(x, y, z + \Delta z) + V(x, y, z - \Delta z) - 2V(x, y, z)}{\Delta z^2}. \end{aligned} \quad (5.31)$$

Considérant un maillage uniforme de pas égal à a , l'équation de POISSON pour un atome interne au matériau s'écrit :

$$\begin{aligned} &\frac{V_{l+1,m,n} + V_{l-1,m,n} - 2V_{l,m,n}}{a^2} + \frac{V_{l,m+1,n} + V_{l,m-1,n} - 2V_{l,m,n}}{a^2} + \frac{V_{l,m,n+1} + V_{l,m,n-1} - 2V_{l,m,n}}{a^2} \\ &= \frac{\rho_{l,m,n}}{\epsilon}, \end{aligned} \quad (5.32)$$

où $V_{l,m,n}$ est le potentiel électrostatique discrétisé, ρ est densité de charges et l , m et n représentent les $l^{\text{ièmes}}$, $m^{\text{ièmes}}$ et $n^{\text{ièmes}}$ atomes suivant les directions \mathbf{z} , \mathbf{y} et \mathbf{x} respectivement.

Aux interfaces oxyde/silicium, la continuité la composante perpendiculaire de $\epsilon \vec{E}$, impose les relations suivantes :

$$(4V_{l,m,n-1} - 3V_{l,m,n} - V_{l,m,n-2}) \epsilon_{\text{Ox}} = (4V_{l,m,n+1} - 3V_{l,m,n} - V_{l,m,n+2}) \epsilon_{\text{Si}}, \quad (5.33)$$

si l'interface est caractérisée par le $n^{\text{ième}}$ plan atomique perpendiculaire à la direction \mathbf{x} ,

$$(4V_{l,m-1,n} - 3V_{l,m,n} - V_{l,m-2,n}) \epsilon_{\text{Ox}} = (4V_{l,m+1,n} - 3V_{l,m,n} - V_{l,m+2,n}) \epsilon_{\text{Si}}, \quad (5.34)$$

si l'interface est caractérisée par le $m^{\text{ième}}$ plan atomique perpendiculaire à la direction \mathbf{y} ,

$$(4V_{l-1,m,n} - 3V_{l,m,n} - V_{l-2,m,n}) \epsilon_{\text{Si}} = (4V_{l+1,m,n} - 3V_{l,m,n} - V_{l+2,m,n}) \epsilon_{\text{Si}}, \quad (5.35)$$

si l'interface est caractérisée par le $l^{\text{ième}}$ plan atomique perpendiculaire à la direction \mathbf{z} .

Au niveau des contacts des grilles les conditions aux limites de DIRICHLET restent valables :

$$V_{l,m,n} = V_G - V_{FB}, \quad (5.36)$$

où V_{FB} est la tension de bande plate traduisant la différence du travail de sortie entre le métal de grille et le silicium.

Dans les contacts de source et drain, les conditions aux limites de NEUMANN remplacent celles de DIRICHLET adoptées dans le système 2D du chapitre 4 et imposent désormais un champ électrique nul :

$$\frac{\partial V(x, y, z)}{\partial x, y, z} = 0. \quad (5.37)$$

En régime passant, la majorité des électrons issus de la source est transmise à travers le canal, alors que le comportement opposé est valable pour les électrons injectés depuis le drain. La densité électronique de l'extension de source diminue et celle du drain augmente par rapport aux valeurs d'équilibre. Si le potentiel aux extrémités de la zone active est fixe (conditions de DIRICHLET), l'électroneutralité de ces régions (\equiv concentration des dopants donneurs égale celle de la densité électronique) qui résulte de l'auto-cohérence SCHRÖDINGER-POISSON ne peut être conservée. En conséquence, le potentiel électrostatique des extrémités de la zone active doit "flotter" par rapport aux niveaux de FERMI des grands réservoirs à l'équilibre thermodynamique. Ces conditions sont ici applicables grâce au caractère 3D de la zone active. Dans le MOSFET double-grille ultime décrit au chapitre 4, la zone active se réduisait à un fil de silicium idéalement unidimensionnel. Considérer un potentiel flottant aux extensions de source et de drain, aurait conduit à une chute de tension quasi-nulle dans la zone active, phénomène peu probable sachant que les résistances de contact de systèmes réels ne sont jamais totalement négligeables et qu'elles permettent de maintenir une différence de potentiel entre les extrémités de la zone active. Dans un système balistique 1D dépourvu de résistance de contact, seuls les réservoirs limitent la conduction et encaissent la totalité de la tension appliquée (cf § (3.4.2)).

Toutes les frontières qui ne sont pas au contact d'électrodes de grilles remplissent donc les conditions de champ électrique nul, tant au niveau des atomes de silicium des extrémités du nanofil que pour les atomes de l'oxyde des contours du nanofil. La condition de dérivée nulle se

traduit en différences finies par les égalités ci-dessous :

$$\begin{aligned}
 V_{l,m,n} &= \frac{4V_{l,m,n+1} - V_{l,m,n+2}}{3}, \\
 V_{l,m,n} &= \frac{4V_{l,m+1,n} - V_{l,m+2,n}}{3}, \\
 V_{l,m,n} &= \frac{4V_{l+1,m,n} - V_{l+2,m,n}}{3},
 \end{aligned}
 \tag{5.38}$$

suivant que la surface considérée est perpendiculaire à la direction x , y ou z respectivement. Les équations (5.32)-(5.38) permettent de calculer de proche en proche les valeurs discrètes du potentiel électrostatique $V_{l,m,n}$.

L'équation de POISSON est ensuite couplée à celle de SCHRÖDINGER afin de converger vers la solution auto-cohérente. L'approche de l'espace des modes, qui traite séparément le confinement quantique et le transport, se décompose en plusieurs étapes (figure (5.5)) :

- Etape 1 : Résoudre l'équation de POISSON afin d'obtenir un potentiel électrostatique d'essai.
- Etape 2 : Résoudre l'équation de SCHRÖDINGER 2D des plans atomiques transverses et calculer les premières sous-bandes électroniques de chaque vallée le long du nano-fil ainsi que leur fonction d'onde associée.
- Etape 3 : Résoudre l'équation de transport 1D dans le formalisme des fonctions de GREEN hors-équilibre pour en retirer les densités électroniques 1D.
- Etape 4 : Calculer les densités de charges 3D en multipliant les densités électroniques 1D obtenues en 3 par le module au carré des modes transverses trouvés en 2.
- Etape 5 : Revenir à l'étape 1 et calculer le nouveau potentiel électrostatique. Si la convergence est atteinte, le courant est calculé dans le formalisme des fonctions de GREEN ; sinon on reprend les étapes 2 et 3.

5.2 Application au MOSFET Gate-All-Around

Nous allons dans la suite appliquer le formalisme précédemment décrit au MOSFET *Gate-all-around* représenté sur la figure (5.6). Le transistor est caractérisé par un nano-fil de silicium de section carrée ($W_{\text{Si}} = T_{\text{Si}} = 4 \text{ nm}$) entouré d'un oxyde de silicium d'épaisseur constante égale à 1nm.

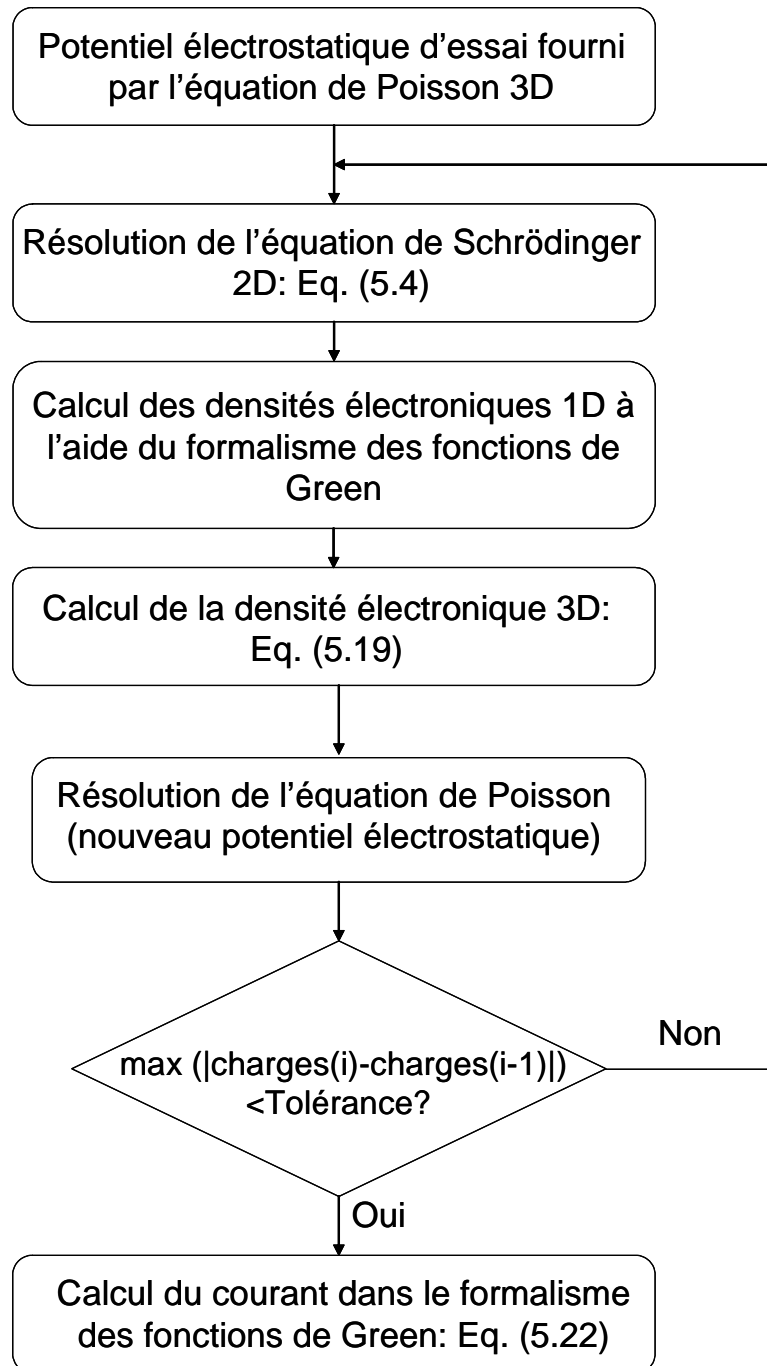


FIG. 5.5: Graphique détaillant les différentes étapes de l'auto-cohérence dans l'approche de l'espace des modes 3D.

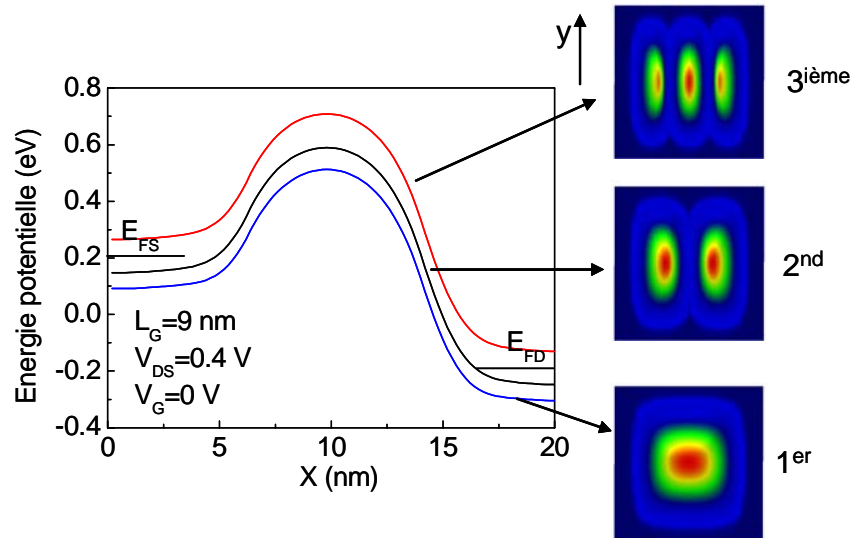


FIG. 5.7: Représentation des trois premières sous-bandes d'énergie associées aux vallées (010) et des modules au carré de leurs états propres transverses. E_{FS} et E_{FD} définissent les niveaux de FERMIE de la source et du drain respectivement.

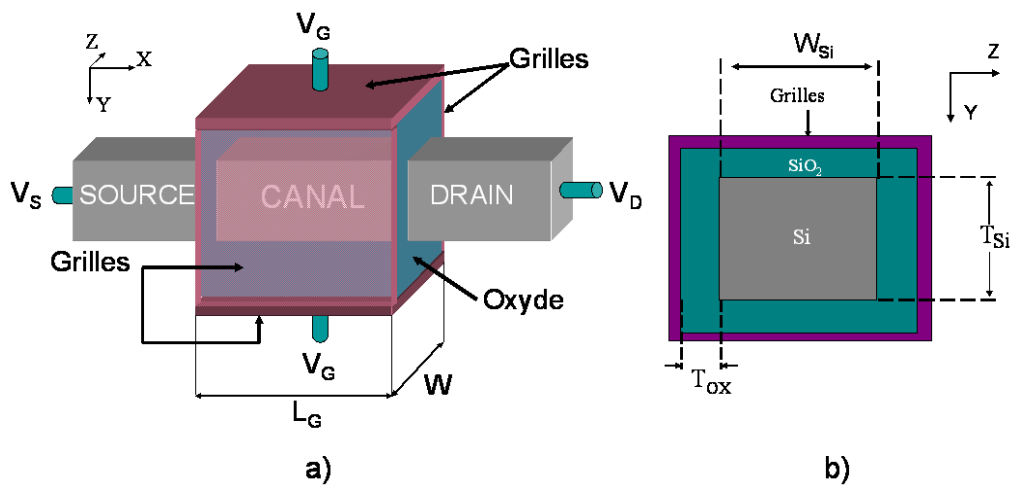


FIG. 5.6: a) Représentation 3D d'un MOSFET Gate-all-around. Par volonté de clarté, l'oxyde de grille n'est pas représenté autour de la source et du drain. b) Définition des principaux paramètres de la section transversale.

La figure (5.7) illustre la séparation de l'équation de SCHRÖDINGER 3D en une équation 2D définissant le confinement suivant les directions y et z et une équation 1D décrivant le transport balistique le long du canal à l'aide du formalisme des fonctions de GREEN. Le confinement transverse induit une discrétisation des énergies qui se répartissent en sous-bandes. L'approximation de l'espace des modes, qui consiste à considérer les sous-bandes indépendantes, permet de ne retenir que celles peuplées d'électrons, c'est-à-dire celles qui se situent sous les niveaux de FERMIE des réservoirs macroscopiques. Afin de légitimer l'application de cette méthode aux transistors confinés, il a été vérifié que les modes restaient rigoureusement identiques d'un plan atomique

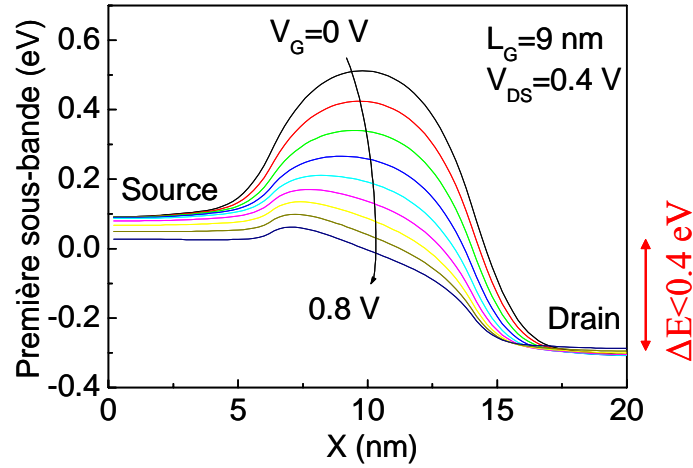


FIG. 5.8: Evolution de la première sous-bande pour V_G variant de 0 à 0.8 V par pas de 0.1 V. La tension appliquée aux extrémités de la zone active est inférieure à V_{DS} (0.4 V), illustrant ainsi la notion de résistance des réservoirs.

à un autre⁶. D'autre part le calcul des sous-bandes et des vecteurs propres de chaque section transverse constitue plus de 70% du temps de calcul cpu total. L'approche "mode-space" peut donc être numériquement optimisée en ne calculant que les fonctions propres de la première section atomique et en les répétant dans les autres plans de la zone active. Les valeurs propres doivent en revanche être déterminées dans tout le système, mais leur obtention est beaucoup plus rapide.

La figure (5.8) représente la première sous-bande des vallées (010) le long de la zone active pour différentes tensions de grille. A faible V_G , la barrière de potentiel du canal interdit le passage des électrons depuis la source vers le drain : le transistor est "bloqué". La résistance du canal est dans ce cas suffisante pour que la totalité de la tension de drain V_{DS} appliquée aux électrodes se retrouve au niveau de la zone active. Lorsque V_G augmente, la barrière de potentiel du canal diminue et les électrons traversent la zone active sans subir de réflexion. Seule la tension drain-source décide du sens du transport. Il en résulte un état hors-équilibre prononcé caractérisé par un déficit de charges dans l'extension de la source et un excédent dans l'extension du drain (figure (5.9)). Plus précisément pour une tension $V_{DS} > 0$, un électron injecté dans l'extension de la source sera transmis à travers le canal, alors que son homologue du drain sera réfléchi. Le nombre d'électrons par niveau d'énergie passe au final de la valeur d'équilibre 2 à 1 dans l'extension de source et à 3 dans celle drain. Le rôle des réservoirs est de s'opposer à cet état hors-équilibre et de maintenir l'électroneutralité en injectant (resp. absorbant) plus d'électrons dans l'extension de la source (resp. du drain). Cependant la quantité d'électrons fournie ou absorbée par un réservoir à travers un seul mode (\equiv sous-bande) est limitée par le quantum de conductance $G_0 = 2e^2/h$ (cf § (3.3.1)). La distribution de FERMI-DIRAC, qui régit la répartition des électrons dans les réservoirs, est à l'origine de cette limitation en imposant un maximum de deux électrons (spin) par niveau d'énergie. Les réservoirs présentent ainsi une

⁶La variation du produit scalaire de vecteurs propres de la sous-bande i appartenant à différentes sections est inférieure au millièème ($\langle T^+ \psi_{n,i} | \psi_{n',i} \rangle \approx 0.999$).

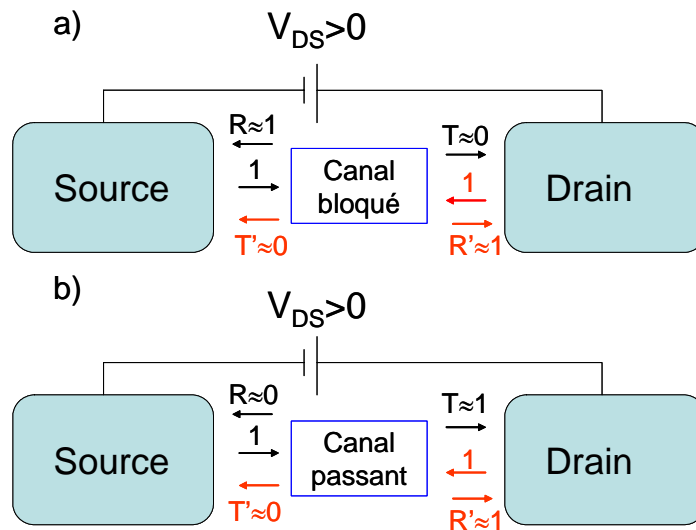


FIG. 5.9: Représentation du transport en termes de probabilités de transmission (T, T') et de réflexion (R, R') d'un état normé ($=1$) dans un : a) transistor passant, b) transistor bloqué.

résistance finie et "encaissent" une partie de la tension V_{DS} appliquée aux électrodes (figure (5.8) pour $V_G = 0.8$ V). L'abaissement des sous-bandes côté source augmente le nombre d'électrons de cette région alors que le décalage positif en énergie des sous-bandes côté drain produit l'effet inverse.

La figure (5.10) montre le coefficient de transmission à travers la barrière source-canal-drain associé aux vallées (010). Malgré un système électrostatique 3D complexe, le coefficient de transmission présente une allure simple en escaliers illustrant la contribution des trois premières sous-bandes d'énergie au processus de transport des porteurs. Concrètement, les énergies pour lesquelles la transmission augmente d'une unité indiquent le début d'une sous-bande et la vitesse de croissance dépend de la forme de la barrière de potentiel entre la source et le drain qui détermine la proportion de la composante tunnel et thermo-ionique. Chaque plateau correspond à un quantum de conductance $G_0 = 2e^2/h$. Nous avons ici un exemple d'évolution des propriétés électriques pour les dimensions mésoscopiques. La conductivité du canal n'est plus continue et change brutalement d'une énergie à une autre. La relation de linéarité entre courant et tension (loi d'OHM) n'est plus vérifiée ou tout du moins ne l'est que sur une plage restreinte de tension où le nombre de sous-bandes participant au transport est constant.

La figure (5.11) représente la densité électronique le long de l'axe source-drain, prise au centre de la section transverse, pour les deux régimes de fonctionnement du transistor. En régime bloqué, la diminution de la quantité de charges du canal suit une loi exponentielle, caractéristique d'une transmission par effet tunnel. Les électrons issus de la source sont clairement dissociés de ceux provenant du drain. A l'état passant, la densité électronique des charges augmente dans le canal : les électrons de la source traversent désormais la zone active. La séparation spatiale des particules de chaque réservoir est moins apparente. Nous observons également des oscillations côté drain, associées aux réflexions au niveau du contact. Les états injectés depuis le drain se

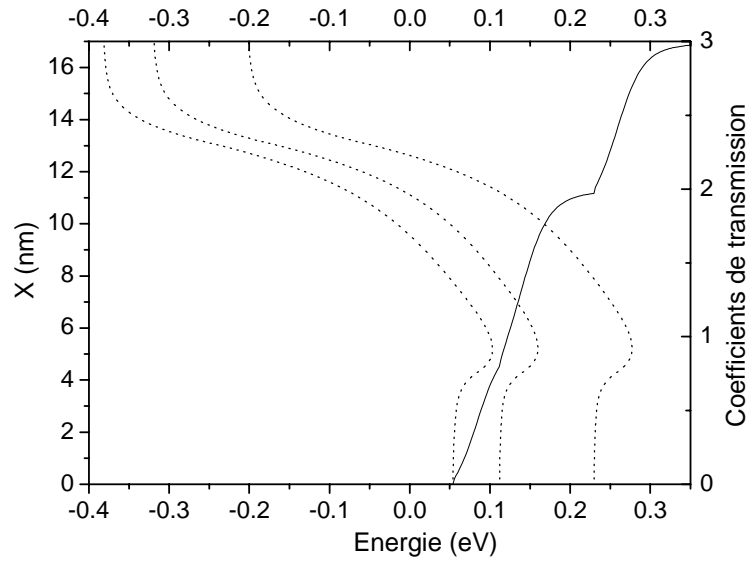


FIG. 5.10: Coefficient de transmission à travers les vallées (010) le long de l'axe source-canal-drain et les sous-bandes associées. $V_{DS} = 0.5$ V, $V_G = 0$ V et $L_G = 9$ nm.

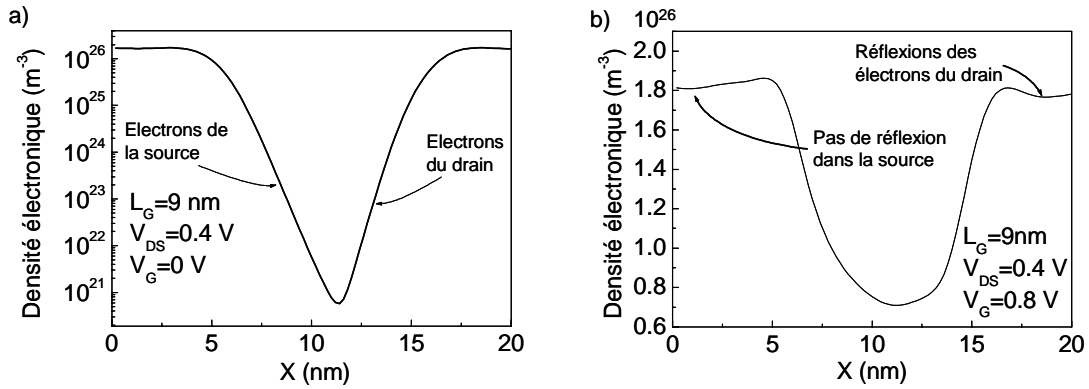


FIG. 5.11: Profil de la densité électronique le long de la zone active du MOSFET GAA simulé : a) en régime bloqué, b) en régime passant. $\theta = 300$ K.

réfléchissent sur la barrière du canal et interfèrent avec les nouveaux états entrants dans la zone active. Côté source, la majorité des électrons traverse le canal et les interférences avec les électrons réfléchis sont plus faibles. Notons également, que la présence d'oscillations repose sur la conservation de la cohérence de phase dans les composants balistiques. En présence de diffusion les interférences sont détruites [202]. Ces oscillations sont plus visibles si l'on étudie la densité d'états locale (équation (5.18) sans les fonctions de distribution) le long de la zone active pour chaque régime (figure (5.12)). Dans l'état bloqué, les extensions de source et de drain présentent des franges d'interférence traduisant les réflexions des états incidents sur la barrière de potentiel du canal et leur avec les nouveaux états entrant dans la zone active. En régime passant, seule l'extension de drain conserve les oscillations. On observe également une densité non nulle sous les sous-bandes ce qui est dû au tunnel source-drain.

Les caractéristiques $I_D - V_G$ calculées pour deux longueurs de grille différentes (figure (5.13))

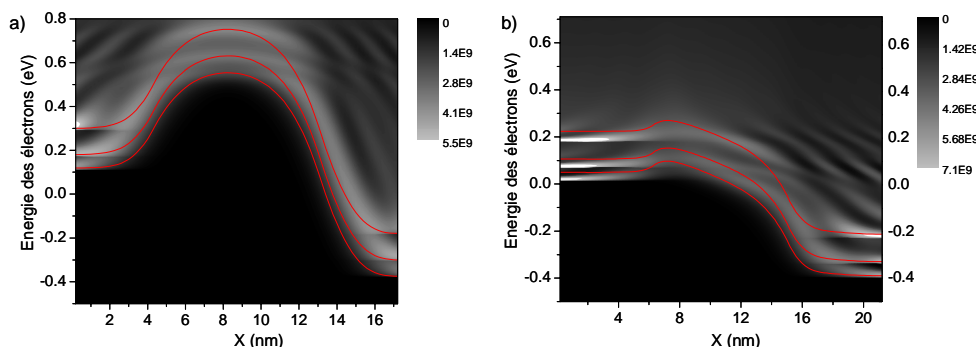


FIG. 5.12: Densités locales d'états et sous-bandes (lignes continues) de la vallées (010) d'un MOSFET GAA a) en régime bloqué ($V_G = 0$ V), b) en régime passant ($V_G = 0.7$ V). $L_G = 9$ nm, $V_{DS} = 0.5$ V.

montrent que l'effet tunnel entre la source et le drain s'amplifie avec la diminution de L_G sous les 10 nm. A ces dimensions, l'épaisseur de la barrière du canal décroît : la composante tunnel à faible V_G dégrade la pente sous le seuil ainsi que le courant I_{off} , alors que le courant thermo-ionique présente des performances sous le seuil bien meilleures. Le transistor devient "tunnel" [204], mais sa tension de grille contrôle toujours le courant du canal.

5.3 Comparaison des différentes architectures émergentes

Les MOSFETs multi-grille, tels que le *Tri-gate*, le *Pi-gate*, le *Omega-gate* et le *Gate-all-around* (figure (5.14a)) sont pressentis comme de sérieux candidats pour continuer la diminution des composants jusqu'à l'échelle nanométrique [32]. Une limitation importante à ce niveau d'intégration est la prédominance des effets quantiques, principalement le confinement du canal (qui induit une séparation des sous-bandes) et la conduction par effet tunnel à travers la barrière de potentiel du canal. Il a été théoriquement prouvé [201] que les transistors à nano-fil de silicium avec deux/trois grilles et un bord flottant (pas de grille arrière) peuvent présenter un meilleur contrôle électrostatique que les MOSFETs double-grille. L'objectif de cette section est précisément de simuler à l'aide du formalisme des fonctions de GREEN des transistors à nano-fils présentant différentes architectures de grille. Nous comparons les performances électriques en fonction des configurations de grille (figure (5.14b)) et explorons la dépendance des performances face aux variations d'épaisseur, de largeur et de longueur du fil de silicium afin d'évaluer les capacités et les limites de chaque architecture. La figure (5.15) illustre l'ensemble des paramètres utilisés pour définir les formes des multi-grilles. Nous considérons toujours un oxyde de silicium d'épaisseur $T_{ox} = 1$ nm. La section transverse de la partie en silicium reste également inchangée ($T_{Si} = W_{Si} = 4$ nm). Seule l'influence de la structure des grilles est dans un premier temps étudiée.

La figure (5.16) compare les caractéristiques $I_D - V_G$ de chaque architecture. Nous constatons que les paramètres sous le seuil ainsi que le courant I_{on} s'améliorent nettement avec l'augmen-

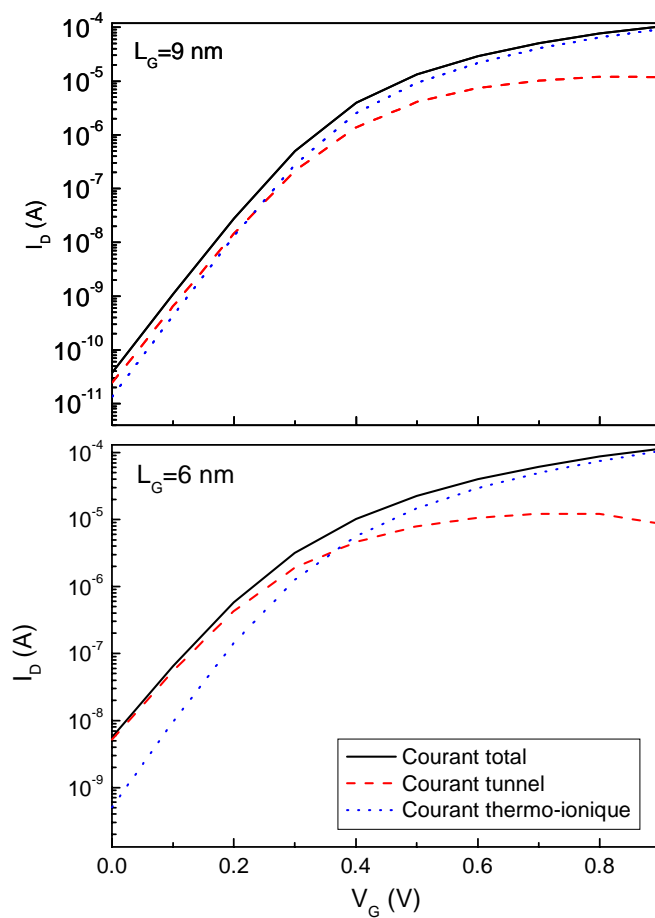


FIG. 5.13: Caractéristiques $I_D - V_G$ calculées pour un MOSFET GAA avec $L_G = 9$ nm et $L_G = 6$ nm. $T_{Si} = W_{Si} = 4$ nm et $V_{DS} = 0.5$ V. La composante tunnel dégrade substantiellement la pente sous le seuil du transistor court.

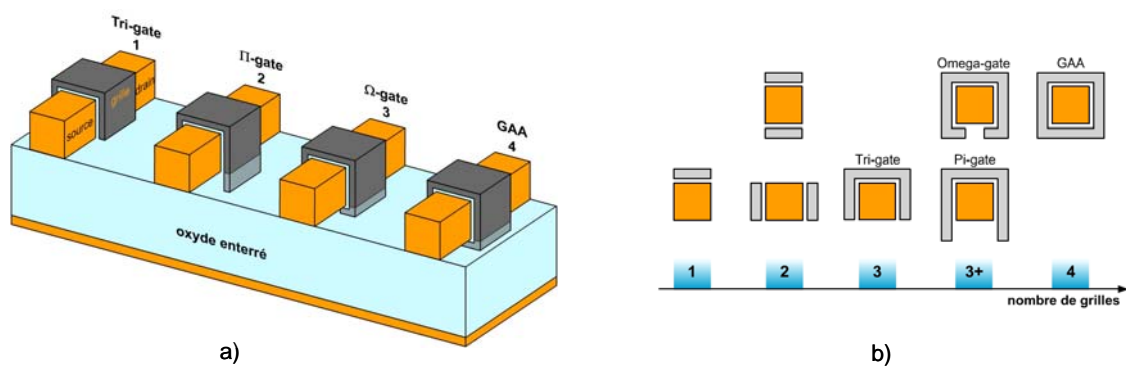


FIG. 5.14: a) Différentes configurations de MOSFETs multi-grilles. b) Sections transverses schématiques classées en fonction du nombre équivalent de grilles.

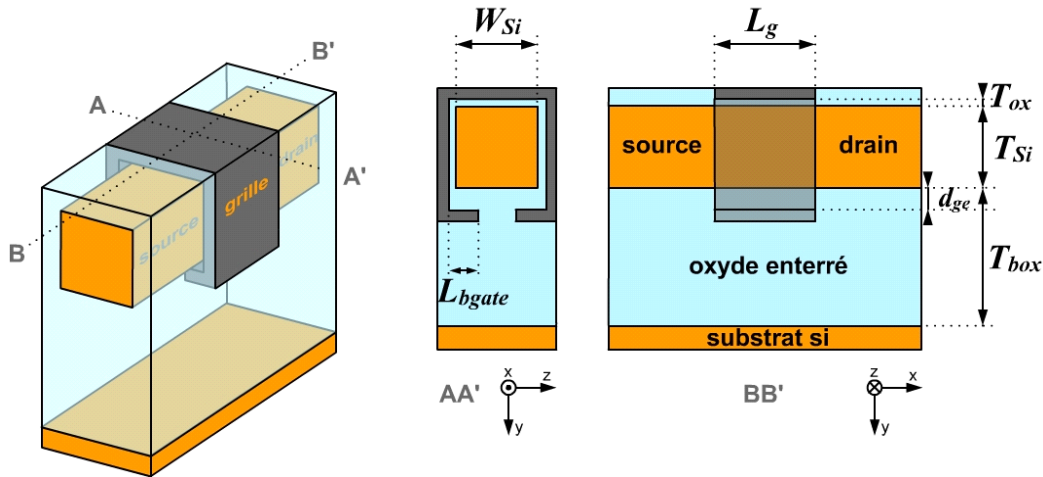


FIG. 5.15: Définition des principaux paramètres associés aux architectures présentées figure (5.14). Pour toutes les structures simulées, la concentration de dopants S/D est 10^{20} cm^{-3} et le canal est intrinsèque. Il n'y a pas de recouvrement de grille avec les régions S/D et les jonctions sont supposées abruptes.

tation du nombre de grilles équivalentes, c'est-à-dire avec le renforcement du contrôle électrostatique des grilles sur la conduction du canal. Cet effet est illustré par la figure (5.17) où la densité électronique dans le plan $y-z$ du milieu du canal est représentée. Sans tenir compte de l'amplitude des densités elle-même, on observe une inversion volumique plus prononcée au centre de la section transverse dans le cas du transistor *Gate-all-around*. L'influence de chaque grille sur les charges apparaît nettement depuis la structure *Tri-gate* jusqu'au MOSFET *Gate-all-around* en passant par le Π -*gate* et le Ω -*gate* et les densités électroniques gagnent en homogénéité avec l'augmentation du nombre équivalent de grilles. Les caractéristiques $I_D - V_G$ pour des transistors plus courts ($L_G = 6 \text{ nm}$) sont représentées figure (5.18). Nous observons que la dégradation en régime sous le seuil est plus importante que dans le cas précédent (*i.e.* $L_G = 9 \text{ nm}$) lorsque le nombre équivalent de grilles est réduit. La figure (5.18) montre également que le GAA contrôle le mieux la barrière du canal. Les coefficients de transmission à travers la barrière source-drain associés aux vallées (010) sont représentés figure (5.19). En régime passant, les courbes en escalier traduisent la contribution des trois premières sous-bandes d'énergie. A énergie donnée l'augmentation du coefficient de transmission avec le nombre équivalent de grilles constitue clairement un indice de qualité du contrôle électrostatique. Le comportement est inversé en régime bloqué, *i.e.* que le GAA fournit la plus petite transparence est le *Tri-gate* la plus élevée. L'influence des extensions de grille pour des structures Π -*gate* et Ω -*gate* est reportée sur la figure (5.20). Cette série de résultats montre que les performances globales de l'architecture *Tri-gate* sont faiblement améliorées par la pénétration verticale de la grille dans l'oxyde enterré. Il en est autrement pour le Ω -*gate* dans lequel un mince recouvrement de la grille arrière se traduit par une forte modification des caractéristiques électriques; ces dernières se rapprochant de celles du MOSFET GAA. L'impact de l'extension de grille est plus visible en observant la pente sous le seuil (S) et le courant I_{on} de chaque architecture séparément (figure (5.21)). Alors qu'une pénétration verticale de 3 nm de l'oxyde d'un Π -*gate* n'induit qu'une amélioration de 7.5% de S et de 15% du

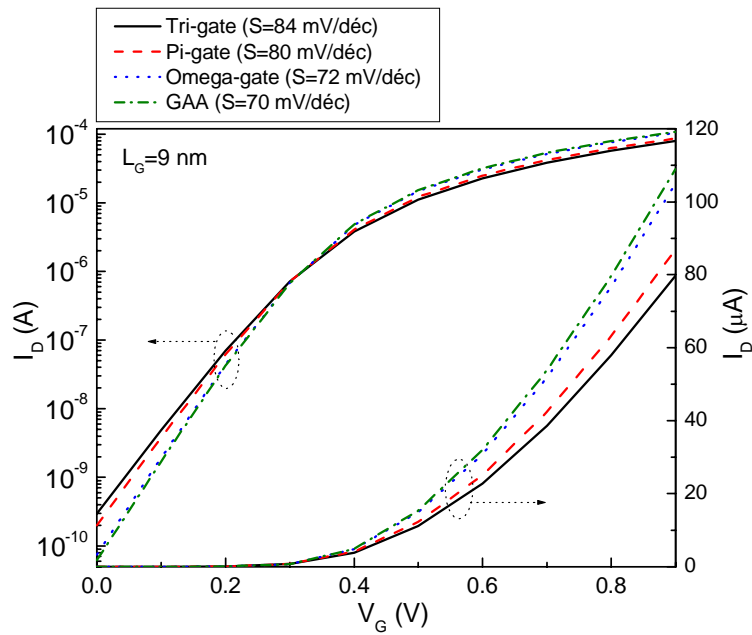


FIG. 5.16: Caractéristiques $I_D - V_G$ des différentes architectures. Augmenter le nombre de grilles améliore de manière générale les performances du transistor. Les transistors Π -gate et Ω -gate présentent respectivement une extension de grille égale à $d_{ge} = 0.8 \text{ nm}$ et $L_{bg} = 0.4 \text{ nm}$. $T_{Si} = W_{Si} = 4 \text{ nm}$.

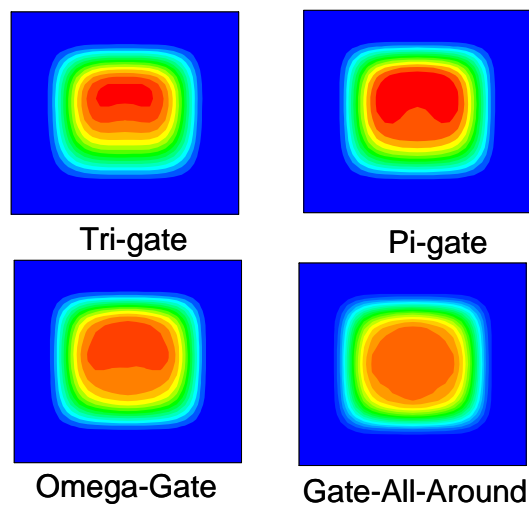


FIG. 5.17: Evolution des densités électroniques dans la section transverse du milieu du canal pour différentes structures. $V_{DS} = 0.5 \text{ V}$, $V_G = 0.7 \text{ V}$. Les autres paramètres sont identiques à ceux de la figure (5.16).

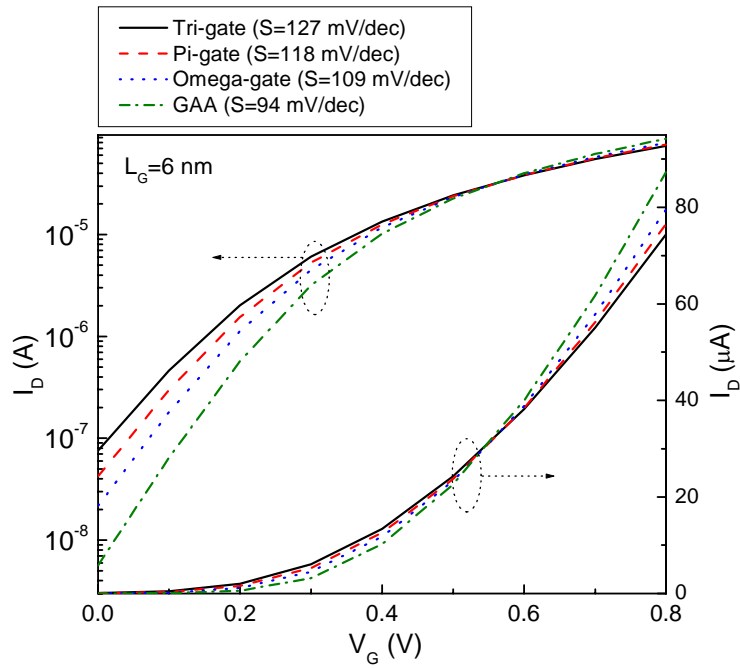


FIG. 5.18: Caractéristiques $I_D - V_G$ des différentes architectures pour une longueur de grille $L_G = 6$ nm. Les autres paramètres sont ceux de la figure (5.16).

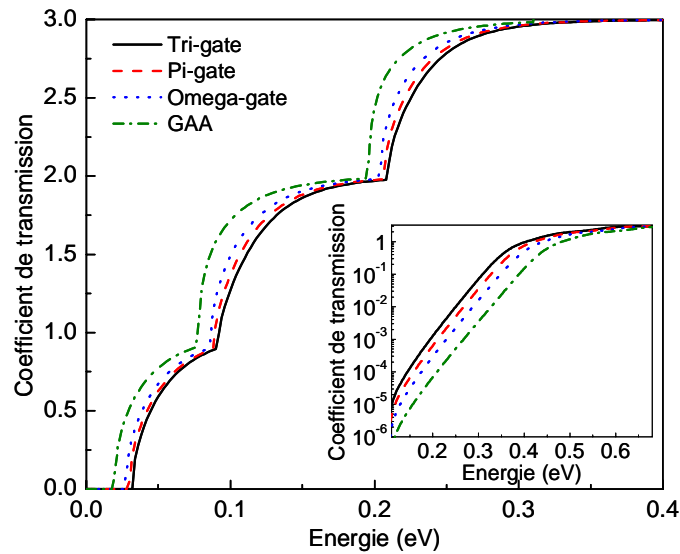


FIG. 5.19: Coefficients de transmission des vallées (010) pour les différentes architectures avec $L_G = 6$ nm en régime passant ($V_G = 0.7$ V). Figure insérée : coefficient de transmission pour les mêmes architectures en régime bloqué ($V_G = 0$ V). Les autres paramètres sont ceux de la figure (5.16)

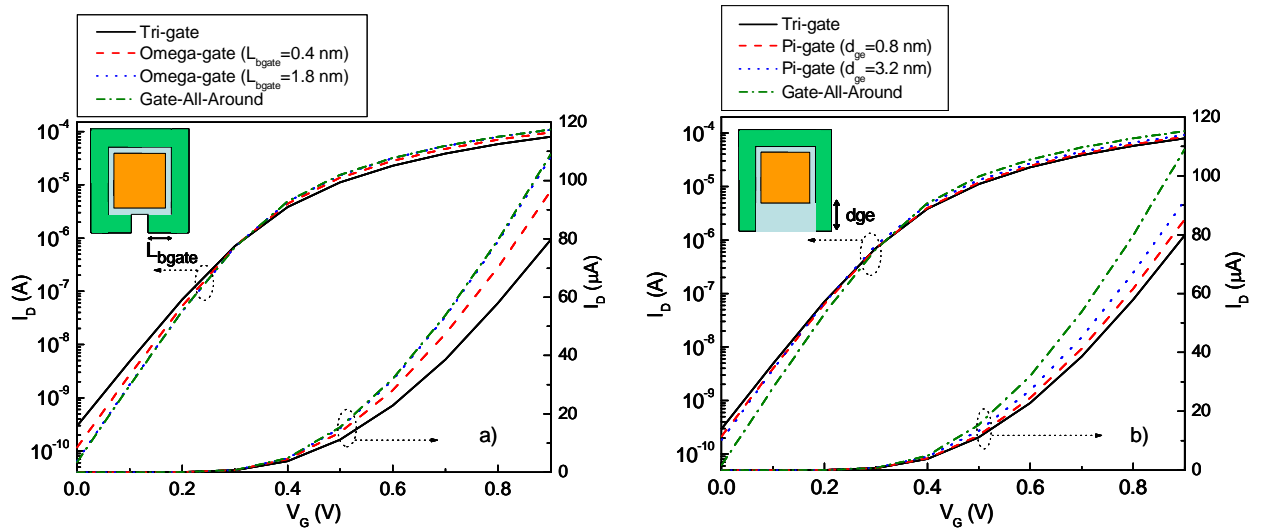


FIG. 5.20: Evolution des caractéristiques de courant $I_D - V_G$ en fonction a) de la pénétration de l'oxyde vertical des Π -gate b) du recouvrement de l'oxyde arrière des Ω -gate. $L_G = 9$ nm, $T_{Si} = W_{Si} = 4$ nm, $V_{DS} = 0.5$ V.

I_{on} , un recouvrement de la grille arrière d'un Ω -gate de 0.4 nm améliore les mêmes paramètres de 10% et 21% respectivement.

Finalement, la figure (5.22) montre l'évolution de la pente sous le seuil en fonction de la section transverse d'un fil de silicium carré pour le GAA et le *Tri-gate*. L'augmentation de la section transverse induit une divergence rapide de S vers des valeurs inacceptables (>110 mV/décade) pour le *Tri-gate* alors que la pente du GAA est moins modifiée. D'autre part, nous observons pour de faibles sections transverses, que l'influence de la configuration de grille s'atténue et que les pentes S des deux structures sont sensiblement les mêmes. Ce comportement est expliqué par le fait qu'en-dessous 2 nm, la combinaison du contrôle électrostatique des grilles et des effets du confinement quantique conduit à des distributions de densité de porteurs similaires dans le film pour toutes les structures de grille. Nous illustrons cet effet figure (5.23) dans le cas le plus défavorable du *Tri-gate*. La structure présente une distribution électronique 2D centrée identique à celle du GAA lorsque la section transverse décroît sous 2 nm. À ce niveau ultime d'intégration, on peut supposer que toutes les architectures multi-grilles deviennent sensiblement équivalentes.

5.4 Conclusion

Dans ce chapitre, nous avons présenté une simulation quantique 3D de transistors à nano-fil de silicium basée sur le formalisme des fonctions de GREEN hors-équilibre. L'approche de l'espace des modes est adoptée pour décomposer l'hamiltonien 3D et ainsi obtenir un gain significatif du temps de calcul. Ceci nous a permis de nous intéresser à des architectures MOSFETs émergentes. Nous avons notamment comparé les performances électriques de différents compo-

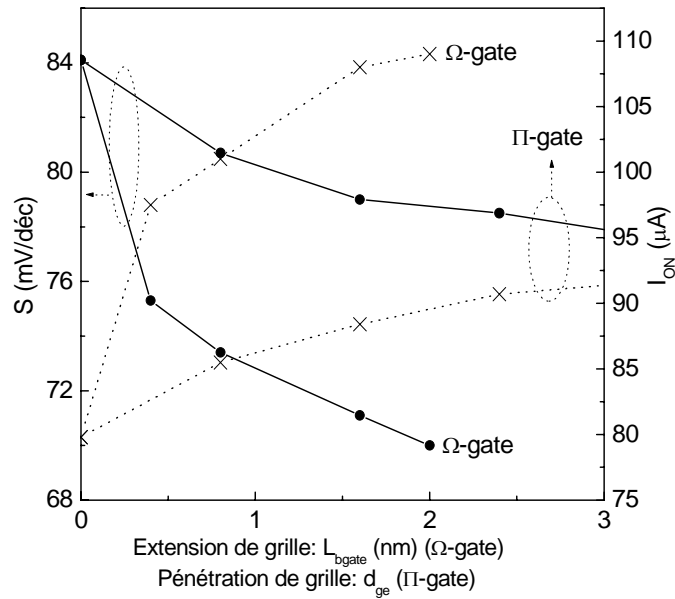


FIG. 5.21: Influence des extensions de grille dans les structures Π -gate et Ω -gate sur la pente sous le seuil S (continu) et le courant I_{on} (pointillés). $L_G = 9$ nm, $T_{Si} = W_{Si} = 4$ nm, $V_{DS} = 0.5$ V.

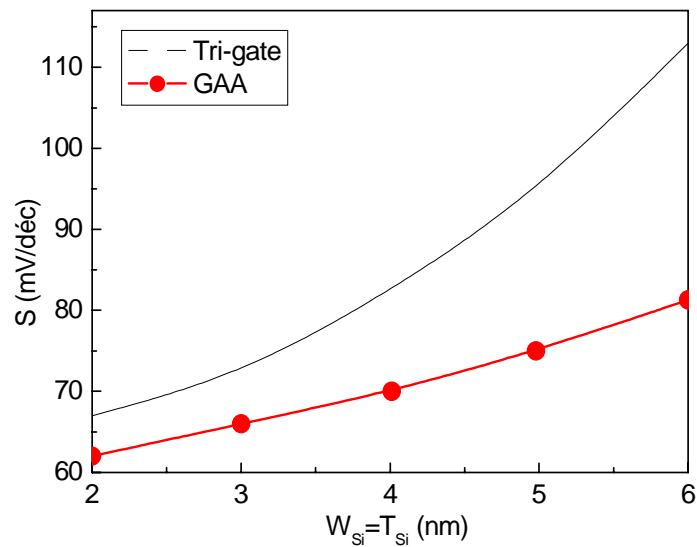


FIG. 5.22: Pente sous le seuil en fonction de la section transversale carrée pour les architectures Tri-gate et GAA. Les autres paramètres sont : $L_G = 9$ nm, $V_{DS} = 0.5$ V.

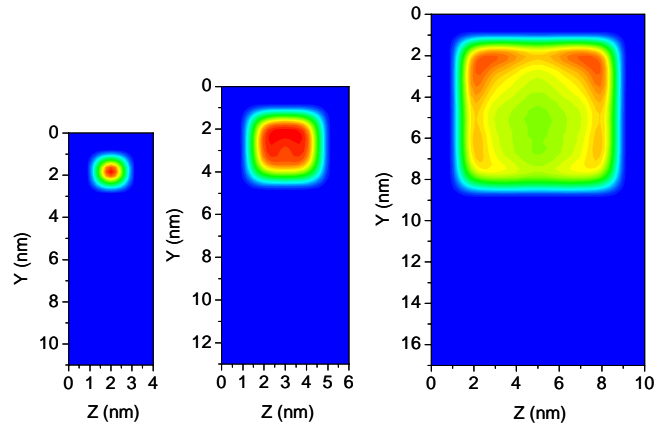


FIG. 5.23: *Distribution de la densité électronique 2D dans le plan y - z d'une structure Tri-gate avec : 1) $T_{Si} = W_{Si} = 2$ nm, 2) $T_{Si} = W_{Si} = 4$ nm, 3) $T_{Si} = W_{Si} = 8$ nm.*

sants multi-grilles en fonction de leur architecture. Le transistor gate-all-around semble présenter les meilleures performances grâce à son excellent contrôle électrostatique des charges du canal et reste le favori pour parvenir à des longueurs de grille sous les 10 nm. La structure Ω -gate qui propose des performances très proches de ce dernier sans nécessiter de grille arrière (plus difficilement réalisable) pourrait bien constituer une autre solution. Enfin, nous avons montré que les différences électriques entre architectures s'estompaient avec la diminution de la section transverse.

Les informations obtenues jusqu'à présent concernent des composants parfaits, c'est-à-dire sans défaut cristallin. Ce cas de figure est en réalité peu probable et la prévision de l'influence de défauts ponctuels dans de tels systèmes soulève des interrogations majeures. Le chapitre suivant tentera d'y répondre en modélisant des défauts ponctuels dans les transistors 3D, toujours à l'aide du formalisme des fonctions de GREEN.

Chapitre 6

Traitement et influence de défauts ponctuels dans les nano-transistors

La diminution de la taille des MOSFETs vers les limites physiques et technologiques pose le choix des structures des prochaines générations CMOS en des termes nouveaux. Parmi les difficultés envisagées, les fluctuations aléatoires du nombre et de la position des impuretés dopantes des composants peuvent détériorer la fiabilité des transistors (variation de la tension de seuil et du courant I_{off}) et les rendre impropres aux opérations CMOS [205, 206]. En effet, lorsque les MOSFETs atteignent l'échelle nanométrique, la variation des paramètres du transistor due au faible nombre de dopants et à leur position dans le canal devient préoccupante. Du fait de la réduction de la tension d'alimentation des transistors ultimes, les nouveaux circuits intégrés sont de surcroît plus sensibles aux fluctuations des caractéristiques des MOSFETs et les positions aléatoires des dopants se répercutent désormais sur les performances et la fonctionnalité des circuits analogiques et numériques.

Dans ce chapitre, nous utilisons une description quantique basée sur le formalisme des fonctions de GREEN hors-équilibre pour étudier l'influence de défauts ponctuels sur les paramètres électriques de structures MOSFETs émergentes. La première section présente la solution analytique d'un défaut ponctuel représenté par un potentiel intra-atomique, introduit dans une chaîne linéaire à l'équilibre. La section (6.2) applique cette approche au MOSFET double-grille dont le canal, ultimement confiné, est réduit à une chaîne atomique (cf chapitre 4). Cependant, l'influence du potentiel coulombien associé à l'écrantage par les charges du canal de l'impureté ionisée est négligé. La section (6.3) étend donc le modèle à une description quantique plus réaliste, capable de simuler un défaut dans des dispositifs émergents 3D décrits au chapitre 5 à l'aide de l'approximation "mode-space". La section (6.4) étudie l'évolution des paramètres physiques qui accompagne la présence d'un défaut et discute l'influence de la position et du type de ce dernier sur les caractéristiques électriques. La dernière section établit les conclusions et perspectives.