



HAL
open science

Intégration à trois dimensions séquentielle: Etude, fabrication et caractérisation

Perrine Batude

► **To cite this version:**

Perrine Batude. Intégration à trois dimensions séquentielle: Etude, fabrication et caractérisation. Matière Condensée [cond-mat]. Institut National Polytechnique de Grenoble - INPG, 2009. Français. NNT: . tel-00455428

HAL Id: tel-00455428

<https://theses.hal.science/tel-00455428>

Submitted on 10 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SOMMAIRE

INDEX.....	5
INTRODUCTION A L'INTEGRATION 3D SEQUENTIELLE.....	8
A- POURQUOI PASSER EN 3D.....	9
B- LES DIFFERENTES INTEGRATIONS 3D.....	9
C- L'INTEGRATION 3D SEQUENTIELLE (DE 1979 A 2009)	11
D- DESCRIPTION DU MANUSCRIT.....	14
CHAPITRE I	
LES BRIQUES TECHNOLOGIQUES DE L'INTEGRATION SEQUENTIELLE.....	15
A-VUE D'ENSEMBLE DES DEVELOPPEMENTS TECHNOLOGIQUES NECESSAIRES	18
B- STABILISATION DU MOSFET INFERIEUR	21
B-1-STABILISATION DU SILICIURE	21
B-1-1-Mise au point d'un siliciure adapté à l'intégration 3D séquentielle.....	22
B-1-2-Etat de l'art des siliciures dans l'intégration 3D séquentielle.....	24
B-2-STABILISATION DU MODULE DE JONCTION.....	26
C-REALISATION DE LA ZONE ACTIVE SUPERIEURE	28
C-1-LES TECHNIQUES.....	28
C-1-1-Les techniques de type « seed-window ».....	28
C-1-2-Le report de couche.....	29
C-1-3- Seed window versus report de couche.....	31
C-2-LE COLLAGE DANS L'INTEGRATION SEQUENTIELLE.....	33
C-2-1-Les nouvelles problématiques.....	33
C-2-2-Caractérisation du collage.....	35
C-2-3-Amincissement du diélectrique inter-niveau	36
C-2-4-Gravure des contacts à travers l'interface de collage.....	38
C-3-COMPARAISON A L'ETAT DE L'ART.....	39
D-L'ALIGNEMENT.....	40
D-1-MODIFICATION DES MARQUES D'ALIGNEMENT.....	40
D-2-COMPARAISON AVEC L'INTEGRATION 3D PARALLELE.....	41
D-3-STRATEGIE D'ALIGNEMENT DANS L'INTEGRATION 3D SEQUENTIELLE.....	43
E-REALISATION DU FET SUPERIEUR ; PROCEDES A BASSE TEMPERATURE.....	45
E-1-REPARTITION DES BUDGETS THERMIQUES.....	45
E-2-LE MOSFET GERMANIUM.....	45
E-3-LE MOSFET SILICIUM « BASSE TEMPÉRATURE ».....	46
E-3-1-Activation à basse température.....	46
E-3-2-Surélévation des sources et drains à basse température.....	48

F-INTERCONNECTIONS 3D.....	50
F-1-CONTACT CHEVAUCHANT.....	50
F-2-CONTACT TRAVERSANT.....	51
F-3-CONTACT INTERNE.....	52
 G-CONCLUSION SUR LES DEVELOPPEMENTS DES BRIQUES 3D.....	 53
 <u>CHAPITRE II</u>	
<u>CARACTERISATION ELECTRIQUE DE L'INTEGRATION COMPLETE.....</u>	<u>56</u>
A-INTEGRATION REALISEES	58
B-ETUDE DES CARACTERISATION DU FET INFERIEUR.....	60
B-1-STABILITE DU SILICIURE INFERIEUR.....	60
B-2-CARACTERISATION DE L'EMPILEMENT DE GRILLE.....	63
C-CARACTERISTIQUES DE MOSFETS SUPERIEURS.....	67
C-1-PMOS GERMANIUM.....	67
C-1-1-LES RESULTATS OBTENUS.....	67
C-1-2-LES VOIES D'AMELIORATIONS POSSIBLES.....	68
C-1-3-CONCLUSION PRELIMINAIRE SUR L'OPTION GE.....	68
C-2-LE MOSFET SILICIUM BASSE TEMPERATURE.....	70
C-2-1-DESCRIPTION DES VARIANTES DU LOT.....	70
C-2-2-ANALYSE DES TENDANCES DES TESTS PARAMETRIQUES.....	72
C-2-3-PRESENTATION DES CARACTERISTIQUES UNITAIRES.....	75
C-2-4-CARACTERISATION DES RESISTANCES D'ACCES.....	76
C-2-5-CARACTERISATION DE L'EMPILEMENT DE GRILLE.....	77
D-DEMONSTRATION DE CELLULES EN 3D.....	80
D-1-L'INVERSEUR 3D.....	80
D-2-LA SRAM 6T 3D.....	82
 <u>CHAPITRE III</u>	
<u>LES PERSPECTIVES DE GAIN DE L'INTEGRATION 3D SEQUENTIELLE.....</u>	<u>87</u>
A-LES PERSPECTIVES DE GAIN EN DENSITE.....	89
A-1-ETAT DE L'ART.....	89
A-2-METHODOLOGIE DE CONCEPTION.....	90
A-3-MISE AU POINT D'UN MANUEL DE REGLE DE DESSIN ADAPTE A L'INTEGRATION.....	91
A-4-REALISATION DE LA BIBLIOTHEQUE DE CELLULES.....	95
A-5-SYNTHESE LOGIQUE.....	98
B- LES PERSPECTIVES DE GAIN EN PERFORMANCE.....	100
B-1-GAIN EN PERFORMANCE SUR DES TRANSISTORS UNITAIRES.....	100
B-1-1-LE CAS SOI/SOI	101
B-1-2-LE CAS GEOI/SOI	103
B-2-GAIN EN DELAI DANS LES INTERCONNECTIONS.....	103

C-ETUDE DE COUT.....	109
C-1-PRESENTATION DE L'ETUDE.....	109
C-2-RESULTATS.....	111
D-CONCLUSION GENERALE SUR L'ETUDE REALISEE.....	115
<u>CHAPITRE IV</u>	
<u>LES PERSPECTIVES DE GAIN DE L'INTEGARTION 3D SEQUENTIELLE.....</u>	<u>117</u>
A-LA STRUCTURE3D UTILD.....	119
A-1-PREAMBULE SUR L'INTERET DE LA MODIFICATION DYNAMIQUE DE TENSION DE SEUIL.....	119
A-2-DIMENSIONNEMENT POUR OBTENIR UN COUPLAGE CONSEQUENT.....	120
A-3-EVOLUTION EN FONCTIONDU DESALIGNEMENT ET CAS AVEC GRILLES EMPILEES DE LONGUEURS DIFFERENTES.....	124
A-4-LA REALITE TECHNOLOGIQUE	126
B-LES MEMOIRES SRAM.....	129
B-1-LES MEMOIRES SRAM3D.....	129
B-2-LES PROBLEMATIQUES DE CONCEPTION DE CELLULES SRAM.....	130
B-3-LA SRAM 4T 3D UTILD.....	131
B-3-1-Présentation de la SRAM 4T et analyse de sa stabilité.....	131
B-3-2-Positionnement des transistors en 3D.....	132
B-3-3-Analyse des résultats.....	134
B-4-LA SRAM 6T 3D UTILD.....	137
B-4-1-Présentation du modèle compact.....	137
B-4-2-présentation de la SRAM 6T et analyse de sa stabilité.....	138
B-4-3-Positionnement des transistors et analyse des résultats.....	139
C- LES MEMOIRES FLASH.....	142
D- APPLICATION AUX IMAGEURS FORTEMENT MINIATURISES.....	144
<u>CONCLUSION GENERALE.....</u>	<u>148</u>
<u>PUBLICATIONS DE L'AUTEUR.....</u>	<u>156</u>
<u>REFERENCES BIBLIOGRAPHIQUES.....</u>	<u>159</u>

AFM : Microscope à force atomique (Atomic Force Microscopy)

ALD : Procédé de dépôt de couches atomiques (Atomic Layer deposition)

ASICs : Circuit Intégré Spécialisé (Application Specific Integrated Circuit)

Back -End : Ensemble des étapes de réalisation de réalisation d'un circuit après la réalisation des contacts (BE)

BOX : Oxyde enterré (Burried OXide)

BT : Budget Thermique

BTBT : Phénomène de conduction tunnel entre les bandes de conduction et de valence (Band To Band Tunneling)

Buffer : Montage spécifique destiné à amplifier le courant de sortie d'un circuit.

CMP : Polissage mécano-chimique (Chemical Mechanical Polishing)

CNET : Centre National d'Etude des Télécommunication (France)

DGMOSFET : Transistor MOSFET Double Grille (Double gate MOSFET)

DIBL : Effet de canal court induisant un abaissement de la barrière du potentiel du canal par polarisation du drain (Drain Induced Barrier Lowering)

DRAM : Mémoire dynamique à accès aléatoire (Dynamic Random Access Memory)

DRM : Manuel regroupant les règles de dessin d'une technologie (Design Rule manual)

DTMOS : Transistors à modification dynamique de tension de seuil par polarisation de la prise substrat (Dynamic Threshold Voltage MOSFET)

ϵ : permittivité statique

EOT : Epaisseur de diélectrique équivalente en SiO₂ (Equivalent Oxyde Thickness)

FDSOI : Transistor totalement déserté sur SOI (Fully Depleted Silicon On Insulator)

FLIP FLOP: type de bascule munie d'une entrée sensible aux fronts. La sortie reste constante entre deux fronts.

FPGA: Circuit logique programmable (field-programmable gate array)

Front End : Ensemble des étapes de réalisation des transistors avant la réalisation des contacts (FE)

Full custom : Méthodologie de conception où les circuits sont dessinés « à la main », aucune cellule standard n'est alors utilisée.

High-k : Diélectrique à forte permittivité.

HK-UST: Hong Kong University of science and technology

HP : Hautes performances (High Performance)

HSQ : résine à base d'Hydrogène SilsesQuioxane HSQ

ILD : Diélectrique inter-niveau (Inter Layer Dielectric)

IMS : Institute of Microelectronics of Stuttgart

I_{OFF} : Courant du transistor à l'état bloqué

I_{ON} : Courant de conduction du transistor

ITRS : International Technology Roadmap for Semiconductor www.itrs.net/

Latch : Type de bascule avec un fonctionnement proche d'une bascule D. En français, verrou transparent.

L_G : Longueur de grille (dessinée)

LPCVD: Procédé de dépôt en phase vapeur à basse pression (Low pressure chemical vapour deposition)

LSTP: Très basse consommation (Low Standby Power)

μ_{eff} : Mobilité effective des porteurs

MEB : Microscope Electronique à Balayage

Mesa : Isolation latérale des transistors sur isolant par gravure complète entre les zones actives

MIT : Massachusset Institute of Technology

MOSFET : Transistor à effet de champ de type métal/ oxyde/semiconducteur (Metal Oxide Semiconductor Fiel Effect Transistor)

M_x : x ième niveau de métal

N_A : concentration de dopants

NBL : Nombre de bits maximal par ligne

NEC: Nippon Electronic Corporation

Φ : Travail de sortie du matériau

PMD : Diélectrique pré-métal (Pre Metal Dielectric)

Prise ESD: prise visant à protéger un circuit électrique de décharge électrique extérieures (ElectroStatic Discharge)

PVD : Procédé de dépôt en phase vapeur (Physical Vapor Deposition)

RMG : Technique utilisée pour cristalliser une matériau poly-cristallin ou amorphe où le matériau passe à l'état liquide. Rapid Melt Growth (voir [Feng06])

Routage : Etape de réalisation des interconnexions des différentes cellules ou transistors d'un circuit

RTP : Procédé à temps de recuit thermique réduit (Rapid Thermal Processing)

SCE : Effets de canaux court (Short Channel Effect)

SEG: Croissance par épitaxie selective (Selective epitaxy Growth)

SIMS : Spectrométrie de masse à ionisation secondaire (Secondary Ion Mass Spectrometry)

SPE : Recristallisation en phase solide (Solid Phase Epitaxy)

Split-CV: Méthode d'extraction de mobilité utilisant des mesures en courant et des mesures de capacité [Koomen73, Sodini82].

SRAM: Static Random Acces Memory

SW: Fenêtre d'ouverture utilisé pour le transfert d'information cristalline (Seed window)

TED : Diffusion accélérée transitoire (Transient Enhanced Diffusion)

TEM : Microscope électronique en transmission (Transmission Electron Microscopy)

TI: Texas Instrument

T_{OX} : Epaisseur d'oxyde

TU Delft: University of Technology Delft

V_{DD} : Tension d'alimentation

V_{TH} : Tension de seuil du transistor (THreshold Voltage)

**INTRODUCTION A
L'INTEGRATION 3D SEQUENTIELLE**

A - Pourquoi passer en 3D ?

Depuis la commercialisation du premier microprocesseur par INTEL en 1971, la miniaturisation des transistors MOSFETs a été le principal moteur de la croissance du marché des semiconducteurs. Elle a suivi la loi empirique de Moore, à savoir le doublement de la densité de transistors tous les deux ans. Cette réduction a permis de réduire le coût des circuits intégrés tout en améliorant leur vitesse de fonctionnement.

Avec la réduction extrême des dimensions, des phénomènes parasites apparaissent : effets canaux courts, augmentation des fuites de grille, variabilité, augmentation des consommations statiques et dynamiques. Pour répondre à ces problématiques, de nouveaux matériaux ont été introduits (pour la grille, le canal et les accès) et des changements d'architecture ont été proposés (transistors sur SOI, double grille, transistors multi-canaux, FINFETs..).

D'autre part, le nombre de transistors sur les puces ne cesse d'augmenter, ainsi que la longueur des interconnexions servant à les relier. Les délais apparaissant dans ces interconnexions limitent alors les performances des circuits intégrés, réduisant l'apport des améliorations des performances du transistor unitaire. Pour diminuer ce délai, l'aluminium a été remplacé par le cuivre, moins résistif et les diélectriques isolants ont été rendus poreux. Malgré ces efforts technologiques, le délai dans les interconnexions semble encore limiter les performances des circuits intégrés [Meindl01, Meindl03].

Toutes ces difficultés font renaître un intérêt pour l'intégration 3D. On peut pressentir dans cette architecture de nombreux avantages par rapport à une intégration planaire :

- Possibilité d'augmenter la densité de transistors par unité de surface sans réduire leurs dimensions (permettant de limiter les effets parasites dus à la miniaturisation).
- Possibilité d'améliorer les performances des circuits intégrés via un gain en délai dans les interconnexions grâce au remplacement d'interconnexions horizontales longues par des interconnexions verticales plus courtes et via l'optimisation indépendante des nMOS et des pMOS réalisés sur des niveaux distincts.
- Co-intégration de technologies hétérogènes simplifiées.

B - Les différents 3D

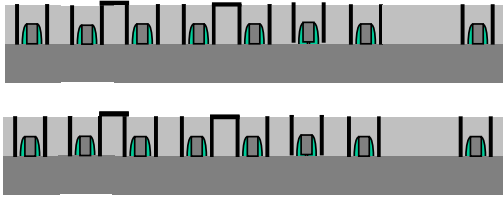
Il existe deux grandes voies technologiques pour réaliser une intégration 3D : l'intégration de type parallèle et l'intégration de type séquentielle (cf. figure 1)

Dans l'intégration de type parallèle, deux plaques sont fabriquées en parallèle puis sont associées pour réaliser l'empilement tridimensionnel.

Dans l'intégration séquentielle, les niveaux de transistors sont fabriqués séquentiellement.

Intégration parallèle

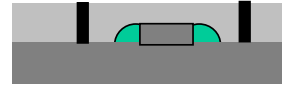
① Réalisation de 2 plaques avec niveau transistor



② Assemblage des deux plaques

***Intégration séquentielle***

① Réalisation du niveau transistor inférieur



② Réalisation du niveau transistor supérieur

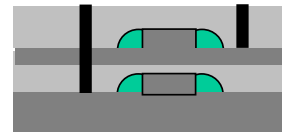


Figure 1 : Description des deux voies technologiques pour réaliser une intégration 3D.

Les briques technologiques nécessaires pour réaliser ces deux types d'intégrations sont profondément différentes.

Dans le cas de l'intégration parallèle, les problématiques sont surtout le collage des deux plaques (collage moléculaire, collage polymère, collage métal-métal), leur alignement, l'amincissement de la plaque supérieure et la réalisation d'interconnexions reliant électriquement les deux niveaux. Pour plus de détails sur cette technologie, le lecteur pourra se référer aux ouvrages de référence [Tan08][Garrou08].

Dans le cas de l'intégration séquentielle, la difficulté réside principalement dans la réalisation d'une zone active supérieure et d'un transistor supérieur à faible température afin de ne pas dégrader les performances du transistor inférieur.

La différence majeure entre ces deux technologies est la densité d'interconnexions réalisable entre les deux niveaux. Cette différence est principalement liée à la précision d'alignement maximale atteignable actuellement pour ces deux technologies : l'écart type minimal démontré est de $\sim 0.1 \mu\text{m}$ dans l'intégration parallèle [Topol05, Steen07] alors qu'elle est de $\sim 10 \text{ nm}$ dans le cas de l'intégration séquentielle [Batude08b]. Quelque soient les dimensions des interconnexions 3D entre les deux couches, nous verrons que cette performance d'alignement limite la densité d'interconnexions 3D à $\sim 10^7$ - 10^8 vias/cm² dans le cas parallèle. Dans le cas de l'intégration séquentielle, la précision d'alignement permet d'atteindre une densité de $\sim 10^{10}$ interconnexions/cm². Elle permet donc de réaliser des interconnexions 3D à l'échelle du transistor.

Malgré ses plus faibles performances en termes d'utilisation de la troisième dimension, (parfois dénommée intégration 2.5 pour cette raison), l'intégration 3D parallèle est la technique la plus étudiée et conséquemment la plus aboutie.

Elle est surtout utilisée pour réaliser des co-intégrations hétérogènes sur une même puce. L'intégration 3D parallèle peut alors être vue comme une extension du SoC, System On Chip (exemple en figure 2) où les différentes fonctions avec des technologies proches sont intégrées verticalement et non plus horizontalement. Elle peut aussi être perçue comme une extension du SiP (System In Package, exemple en figure 3), où les différentes puces de technologies fortement hétérogènes auparavant associées et reliées électriquement par des fils externes au sein du même

boîtier sont reliées par des interconnexions traversant les différents niveaux de l'empilement. Par association de technologies fortement hétérogènes, on entend par exemple une association de circuits intégrés, MEMS, batteries, composants RF, modules de traitement biologique.

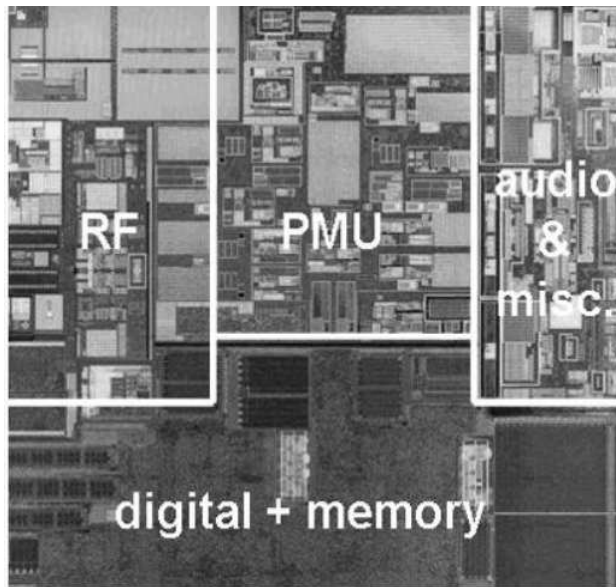


Figure 2 : Exemple de SOC [Hammes07]

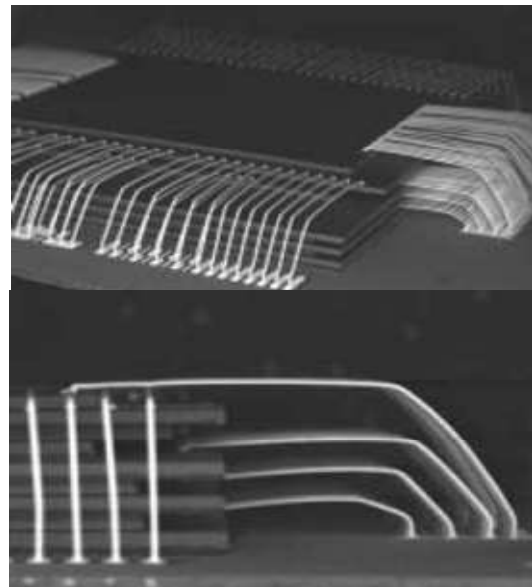


Figure 3 : Exemple de SiP [Toshiba04]

L'option parallèle apparaît actuellement plus mature que l'intégration séquentielle. Cet aspect s'explique par le besoin relativement immédiat de cette technologie pour des applications où la densité d'interconnexions n'est pas particulièrement critique. Nous verrons par la suite (Chapitres III et IV) que pour certaines applications (par exemple la logique pure, les mémoires SRAM et les imageurs fortement miniaturisé, etc.), la réalisation d'un réseau d'interconnexion plus dense peut se révéler un atout fondamental par rapport à l'intégration 3D parallèle.

C - L'intégration 3D séquentielle (de 1979 à 2009)

Les premières études sur l'intégration 3D séquentielle ont débuté dans les années 80 afin de préparer le passage à l'ère des transistors submicroniques. A cette époque, la limitation apparente majeure pour continuer la réduction des dimensions est la définition des motifs par lithographie. Comme l'augmentation de la densité en 2D dans le plan semble être limitée, l'intégration 3D est alors vue comme une solution pour continuer à augmenter la densité des circuits intégrés.

De nombreux articles démontrant la réalisation de transistor empilés ont été publiés entre 1980 et 1993 environ. Les références principales sont présentées dans la table 1.

La difficulté principale à laquelle sont confrontées les équipes de recherche est la réalisation de la zone active supérieure. Les premiers démonstrateurs 3D sont réalisés avec une zone active supérieure en silicium amorphe ou en silicium poly-cristallin. Par la suite, les équipes de recherche développent des techniques pour augmenter la taille des grains afin d'obtenir des caractéristiques de transistors plus proches de celles obtenues sur substrat massif (en particulier en termes de mobilité).

La technique la plus étudiée est la recristallisation laser du dépôt de polysilicium. Certaines études présentent aussi des transistors fonctionnels sur zone active obtenus par recristallisation en phase solide d'un dépôt de silicium amorphe. Dans tous les cas, la zone active supérieure reste polycristalline et les grains sont orientés de façon aléatoire. Les groupes de recherche observent alors que les transistors présentent de fortes variations de paramètres (tension de seuil, épaisseur d'oxyde de grille, etc...) liées à la différence d'orientation des grains et à la présence de joints de grains. Pour permettre d'obtenir des zones actives monocristallines et de même orientation, certains groupes s'orientent sur des techniques permettant de transférer l'information cristalline du substrat inférieur jusqu'à la zone active supérieure via la formation d'un germe cristallin dans une fenêtre d'ouverture. Les techniques utilisant le transfert de l'information cristalline de la zone active inférieure à la zone active supérieure via une fenêtre d'ouverture dans le diélectrique inter niveau seront dénommées techniques « Seed-Window » (SW). Pour plus d'information sur les techniques de recristallisation pour réaliser un substrat SOI, le lecteur pourra se référer à la synthèse réalisée dans [Colinge04].

Groupe	année publication	Technique utilisée pour réaliser l'active supérieure	Référence	Application
CNET	1981-82	poly-Si recristallisé (laser)	[Colinge81a, 81b, 82]	inverseur
Fujitsu	1983-92	poly-Si recristallisé (laser)	[Kawamura83,84,87] [Takao91,92]	inverseur, oscillateur en anneau, SRAM
IMS	1989-93	Si monocristallin (Epitaxie latérale) (SW)	[Zingg89, 90] [Roos91,92, 93]	inverseur, Nand, multiplieur 8X8, ASICs
Matsuhita	1983-90	Poly-amorphe recristallisé (Cristallisation en phase solide) poly-Si recristallisé (laser)	[Uemoto90] [Yamazaki86,90][Akiyama83,84]	SRAM, capteur d'image
MIT	1981	poly-Si recristallisé (laser)	[Goeloe81]	inverseur
Mitsubishi	1983-90	poly-Si recristallisé (laser) Si monocristallin (laser (SW))	[Inoue86, 89] [Aksaka 86] [Nishimura 87]	SRAM capteur d'image
NEC	1988-89	poly-Si poly-Si recristallisé (laser)	[Ando 88] [Kunio89]	SRAM transistor
Sharp	1986	poly-Si recristallisé (laser)	[Ohtake86]	capteur d'image
Stanford	1980-82	poly-Si recristallisé (laser)	[Gibbons80,82]	inverseur
TI	1983-84	poly-Si recristallisé (laser)	[Chen83][Shah 84]	SRAM
Toshiba	1990	poly-Si	[Hayakawa90]	SRAM, miroir de courant

Tableau 1 : Références principales des réalisations et études de l'intégration séquentielle, antérieures à 1993.

Les démonstrations présentées dans ces publications ressemblent peu aux architectures actuelles. Les dimensions (des longueurs de grille, de l'oxyde de grille..) étant très grandes, les paramètres étudiés pour caractériser la faisabilité de l'intégration ont peu de rapport avec ceux classiquement étudiés aujourd'hui (évolution des effets canaux courts, variabilité, réalisation d'oxyde de grille fins, caractérisation des résistances d'accès...). Cependant il est important d'insister sur le nombre d'études que l'intégration 3D séquentielle a suscité à l'époque.

Par la suite, les travaux sur l'intégration 3D séquentielle se sont essouffés, vraisemblablement à cause des difficultés rencontrées pour obtenir une zone active supérieure de qualité cristalline suffisante. D'autre part, les progrès technologiques réalisés, ont permis de continuer à augmenter la densité des circuits intégrés par une simple réduction des dimensions.

C'est dans les années 2000 qu'apparaît un regain d'intérêt pour l'intégration à trois dimensions et notamment pour l'intégration de type séquentielle. Cette fois, l'industrie microélectronique est non seulement confrontée à des limites technologiques, mais aussi à des limites physiques intrinsèques (effets canaux courts, variabilité, délai dans les interconnexions). Cette architecture apparaît alors comme une solution à l'ensemble de ces problématiques en permettant de continuer à gagner en densité tout en relâchant les contraintes de dimensions des transistors (limitation des SCE et variabilité) et en diminuant la longueur des interconnexions. Ce regain d'intérêt se traduit par de nombreux travaux sur le sujet dont les principales références sont récapitulées dans la table 2.

Groupe	Année publication	Technique utilisée pour réaliser l'active supérieure	Application visée	Référence
Samsung	2004-05	Poly-Si ((re)crystallisation)	SRAM	[Kang04][Jung 04,05] [Jang 04][Lim05]
	2006	Poly-Si ((re)crystallisation)	FLASH	[Jung06]
	2007	Si monocristallin (Epitaxie latérale) (SW)	SRAM	[Son 07]
	2007	Si monocristallin (Recrystallisation laser) (SW)	SRAM	[Jung 07]
	2006-2008	Poly-Si ((re)crystallisation)	SRAM innovante	[Sohn06,08]
Hong Kong UST	1999-2001	Poly-Si ((re)crystallisation)	transistors inverseur oscillateur	[Jagar99][Chan00, 01a,b,c,d] [Wang 00a,00b]
	2001-02	Si monocristallin ((re)crystallisation latérale (fusion laser) (SW))	BiCMOS	[Kumar01][Liu 02]
	2004-06	Si monocristallin Utilisation d'un substrat SOI pMOS sur film mince, nMOS sur le substrat massif	transistors inverseur	[Zhang 04] [Wu05][Lin06]
Université de Taiwan	2004-05	Ge monocristallin Report par collage moléculaire	transistors	[Yu 04,05]
Stanford	1998-99	Poly Si ((re)crystallisation)	transistor	[Subramanian 98,99]
	2006-08	Ge monocristallin ((re)crystallisation latérale (fusion thermique) (SW))	transistors, inverseur, étude FPGA	[Feng 06][Wong07] [Lin07][Lin08]
Université Cornell	2001-2006	Si monocristallin (Collage moléculaire SOI)	transistors Etude SRAM, circuits dynamiques	[Xue 01, 03][Tiwari 02] [Liu06]
TU Delft	2008-2009	Poly-Si (Crystallisation laser)	transistors, inverseur	[Mofrad 08,09]
Macronix	2006	Poly-Si ((re)-crystallisation)	FLASH	[Lai 06a,b]
IBM	2007		transistors inférieurs uniquement	[Rajendran07]

Table 2 : Références principales présentant les démonstrations de l'intégration 3D séquentielle dans les années 2000.

Nous noterons que la plupart des études récentes se tournent vers des options qui permettent d'obtenir des îlots de zones actives monocristallines et de mêmes orientations entre elles. Nous reviendrons tout au long du manuscrit sur ces références en sélectionnant les réalisations les plus abouties. Parmi elles, les deux démonstrations majeures de l'intégration 3D séquentielle sont la SRAM réalisée par Samsung avec trois niveaux de zone active en silicium cristallin dont une observation en coupe est présentée en figure 4 et l'inverseur démontré par Stanford avec une zone active supérieure en germanium monocristallin et une zone active inférieure en silicium dont une observation en vue de dessus est présentée en figure 5.

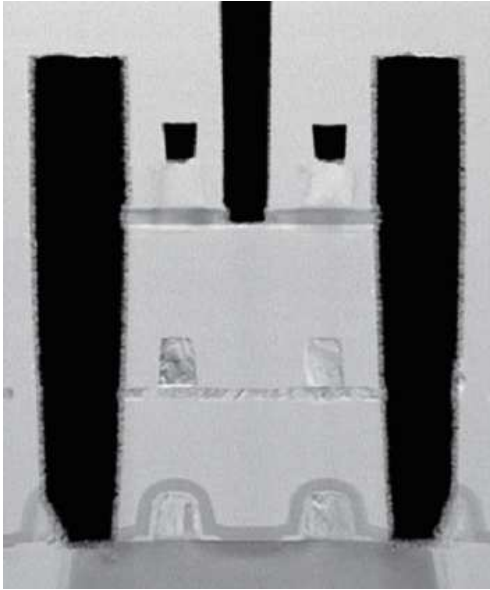


Figure 4 : Observation TEM en coupe de la structure SRAM réalisée par Samsung [Jung07]

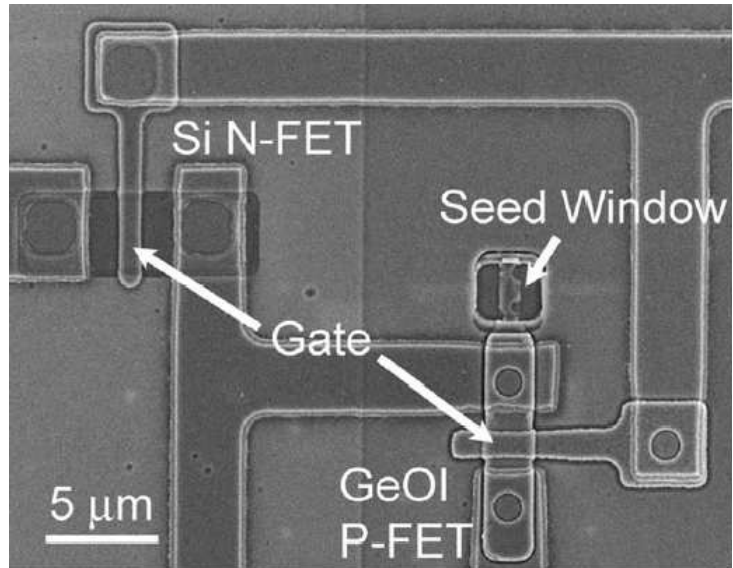


Figure 5: Observation de l'inverseur réalisé par Stanford avec un transistor supérieur en GeOI et un transistor inférieur en SOI [Feng 06]

D - Description du manuscrit

L'objectif de cette thèse a été de concevoir, réaliser et d'étudier un démonstrateur de l'architecture 3D séquentielle.

L'identification des enjeux technologiques et les développements associés seront décrits dans le chapitre 1. Les choix de procédés seront positionnés par rapport à l'état de l'art. Nous chercherons aussi à justifier leur compatibilité avec des nœuds technologiques sub-45 nm.

Le chapitre II sera consacré à la caractérisation électrique des structures réalisées (transistors inférieurs, transistors supérieurs, inverseurs et SRAMs).

Dans un deuxième temps, nous nous sommes attachés à étudier les perspectives de gains (densité, performances, coût) possibles avec l'architecture réalisée pour des applications logiques pures. Cette discussion sera l'objet du chapitre III.

Dans le chapitre IV, nous élargirons le champ d'application de l'intégration 3D séquentielle au domaine des mémoires et aux co-intégrations diverses. Nous tenterons de justifier l'apport de l'intégration 3D séquentielle pour ces technologies.

CHAPITRE I

LES BRIQUES TECHNOLOGIQUES DE L'INTEGRATION 3D SEQUENTIELLE

L'intégration séquentielle présente un avantage majeur par rapport à l'intégration 3D parallèle : les précisions d'alignements atteintes permettent de réaliser des interconnexions inter-niveaux à l'échelle du transistor. La troisième dimension peut alors être efficacement utilisée, rendant l'intégration des cellules CMOS à cheval sur deux niveaux possible.

Cependant les défis technologiques qu'il faut résoudre sont nombreux et placent souvent dans la mentalité générale l'intégration séquentielle comme une intégration à long terme en comparaison à l'intégration parallèle.

Dans ce chapitre, nous présenterons les défis posés par l'intégration 3D séquentielle, les choix technologiques retenus pour les dépasser ainsi que les résultats obtenus lors du développement de chaque brique technologique.

SOMMAIRE DU CHAPITRE I :

LES BRIQUES TECHNOLOGIQUES DE L'INTEGRATION 3D SEQUENTIELLE

A -	VUE D'ENSEMBLE DES DEVELOPPEMENTS NECESSAIRES	18
B -	STABILISATION DU FET INFERIEUR.....	21
B. 1 -	Stabilisation du siliciure.....	21
B.1.1 -	<i>Mise au point d'un siliciure adapté à l'intégration 3D</i>	<i>22</i>
B.1.2 -	<i>Etat de l'art des siliciures dans l'intégration 3D séquentielle.....</i>	<i>24</i>
B. 2 -	Stabilisation du module de jonction.....	26
C -	REALISATION DE LA ZONE ACTIVE SUPERIEURE.....	28
C. 1 -	Les techniques.....	28
C.1.1 -	<i>Les techniques de type « Seed-window »</i>	<i>28</i>
C.1.2 -	<i>Le report de couche.....</i>	<i>29</i>
C.1.3 -	<i>« Seed window » versus report de couche.....</i>	<i>31</i>
C. 2 -	Le collage dans l'intégration 3D séquentielle.....	33
C.2.1 -	<i>Les nouvelles problématiques</i>	<i>33</i>
C.2.2 -	<i>Caractérisation du collage.....</i>	<i>35</i>
C.2.3 -	<i>Amincissement du diélectrique inter-niveau</i>	<i>36</i>
C.2.4 -	<i>Gravure des contacts à travers l'interface de collage.....</i>	<i>38</i>
C. 3 -	Comparaison avec l'état de l'art.....	39
D -	L'ALIGNEMENT	40
D. 1 -	Modification des marques d'alignement.....	40
D. 2 -	Comparaison avec l'intégration 3D parallèle	41
D. 3 -	Stratégie d'alignement dans l'intégration 3D séquentielle.....	43
E -	REALISATION DU TRANSISTOR SUPERIEUR :	45
	PROCEDES A BASSE TEMPERATURE.	45
E. 1 -	Répartition des budgets thermiques.....	45
E. 2 -	Le MOS Ge	45
E. 3 -	Le MOSFET Silicium à bas budget thermique	46
E.3.1 -	<i>Activation à basse température.....</i>	<i>46</i>
E.3.2 -	<i>Surélévation des accès à basse température.....</i>	<i>48</i>

F -	INTERCONNECTIONS 3D	50
F. 1 -	Contact chevauchant.....	50
F. 2 -	Contact traversant.....	51
F. 3 -	Contact interne.....	52
G -	CONCLUSION DEVELOPPEMENT BRIQUES 3D	53

A - Vue d'ensemble des développements nécessaires

Par définition, dans une intégration séquentielle, les transistors supérieurs et inférieurs sont réalisés successivement sur le même substrat support. Dans notre architecture, les deux transistors empilés seront des transistors totalement déplétés (FDSOI pour Fully Depleted Silicon On Insulator). La figure 1.1 présente les principales étapes de fabrication (ainsi que leurs budgets thermiques) d'un FET FDSOI tel qu'il est réalisé au Leti jusqu'aux contacts. Ces étapes sont dites de « Front End » (FE) et précèdent les étapes dites de « Back End » (BE). Les étapes de Back-End correspondent aux étapes permettant d'interconnecter les transistors entre eux et ont un Budget Thermique (BT) classiquement limité à environ 400°C.

Ainsi, dans une technologie planaire, le transistor une fois fabriqué ne subit que des budgets thermiques faibles (<400°C) ne dégradant pas ses performances. Or, dans le procédé de fabrication du transistor FDSOI (cf. figure 1.1), nous observons que de nombreuses étapes sont réalisées à des budgets thermiques souvent très supérieurs à ceux du Back-End. En conséquence, la réalisation du transistor supérieur va nécessairement imposer au transistor inférieur des budgets thermiques supplémentaires pouvant conduire à une dégradation des caractéristiques électriques du transistor inférieur.

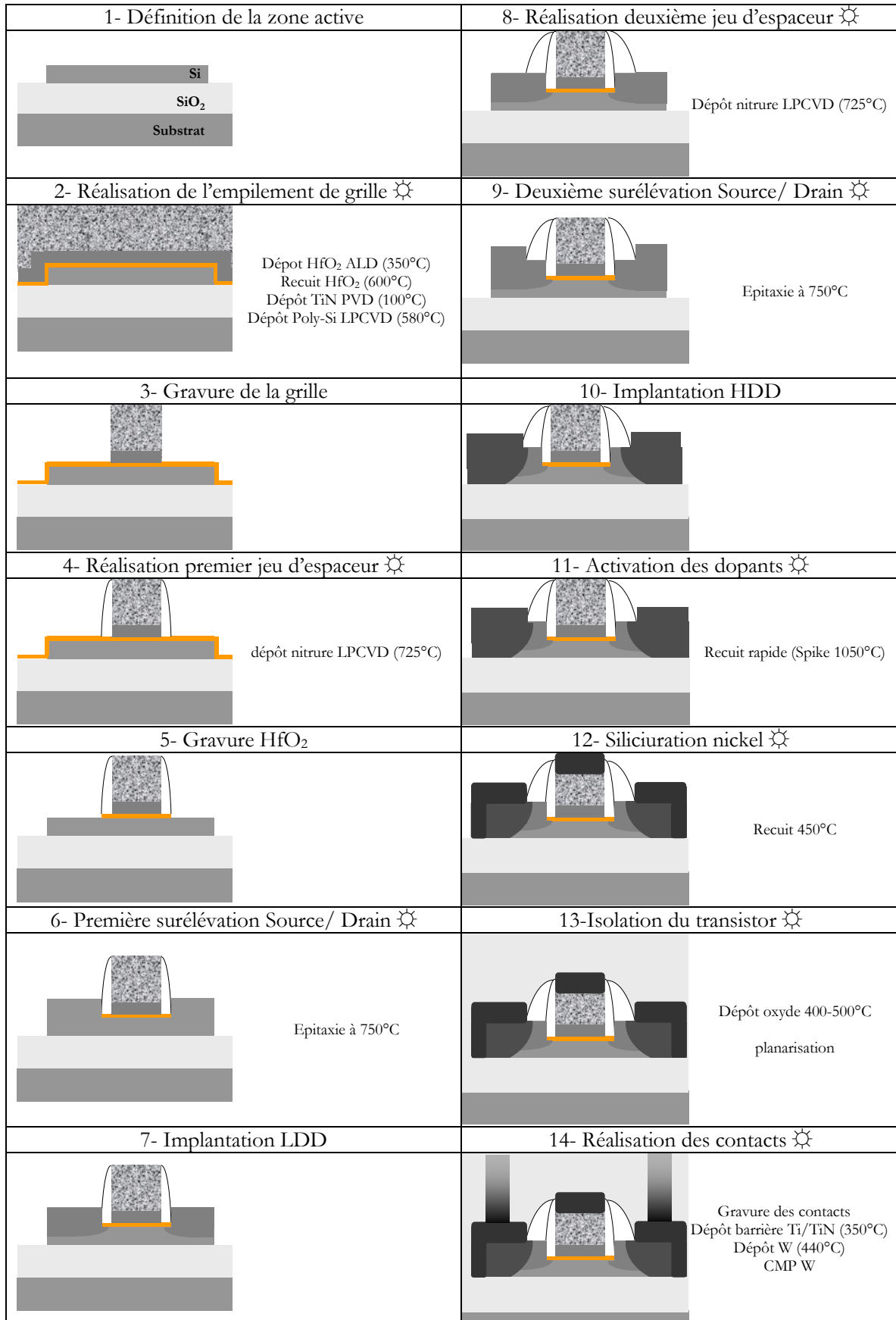


figure 1.1: Schéma de l'enchaînement des principales étapes technologiques de fabrication des transistors FDSOI tel qu'ils sont réalisés au Leti. Le symbole ☼ identifie les étapes technologiques faisant intervenir un budget thermique.

Afin d'obtenir des transistors inférieurs et supérieurs ayant des caractéristiques électriques équivalentes à des transistors planaires, de nombreuses étapes de ce procédé vont être modifiées. Dans ce chapitre nous décrirons l'ensemble des développements technologiques effectués au cours de cette thèse afin de réaliser un démonstrateur de la technologie 3D séquentielle. Ces développements peuvent être regroupés en cinq thématiques récapitulées dans la figure 1.2:

- Stabilisation du FET inférieur, en prévision des budgets thermiques appliqués pour réaliser le FET supérieur (décrits en partie B).
- Réalisation de la zone active supérieure cristalline à faible budget thermique (décrit en partie C).
- Gestion de l'alignement des niveaux lithographiques du transistor supérieur (décrit en partie D).
- Réalisation d'un FET supérieur à faible budget thermique (décrit en partie E).
- Réalisation des interconnexions 3D (décrit en partie F)

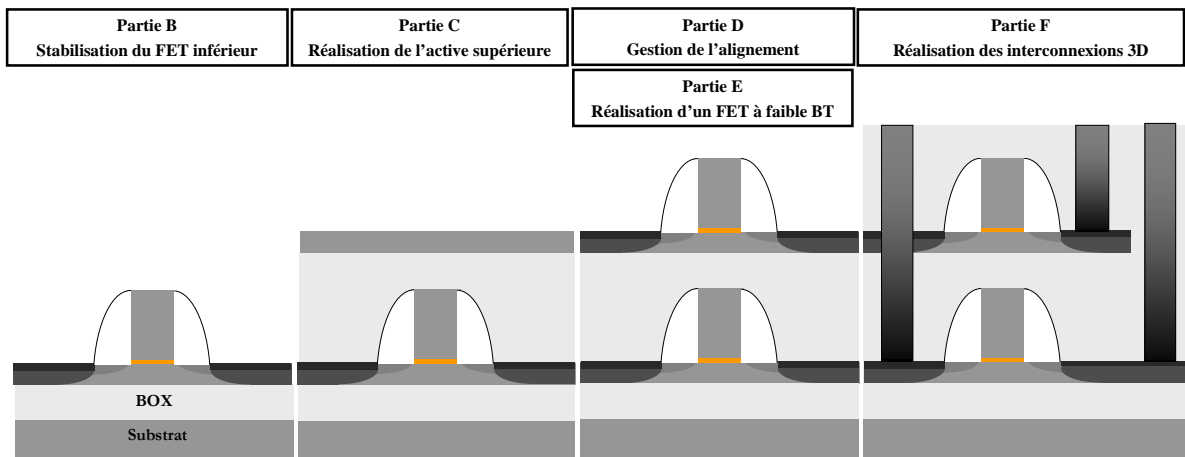


figure 1.2 : Description schématique récapitulant les briques technologiques à modifier par rapport à un procédé d'intégration FDSOI planaire pour réaliser une intégration 3D séquentielle.

B - Stabilisation du FET inférieur

Limiter le budget thermique du FET supérieur est inévitable car divers mécanismes activés en température ont des conséquences préjudiciables sur les performances du transistor inférieur. Nous citerons en particulier la diffusion des dopants et le démouillage en phase solide des siliciures. Dans cette partie nous décrirons brièvement ces deux mécanismes et les problématiques qu'ils engendrent.

Nous verrons dans la partie D que le budget thermique du transistor supérieur a été fixé à 650°C et ne peut guère être descendu en dessous de cette limite. En conséquence, il est nécessaire de s'assurer que le transistor inférieur ne subit pas de modifications de ses performances électriques lorsqu'il est soumis à ce budget thermique limite.

B.1 -Stabilisation du siliciure

La brique technologique la plus sensible à une élévation de température est la siliciuration. En effet, dans le cas du siliciure de nickel NiSi, nous observons une augmentation drastique de la résistance du film dès les premières minutes de recuit à 650°C (figure 1.3). Cette augmentation est liée à l'apparition de discontinuité dans le film de siliciure : il y a démouillage en phase solide (figure 1.4).

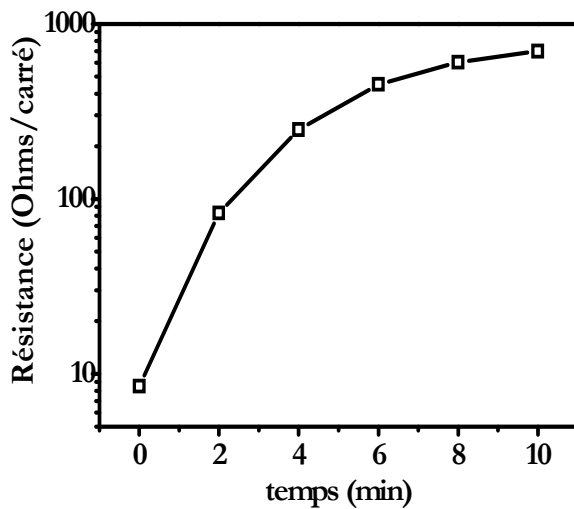


figure 1.3 : Evolution de la résistance de la couche de siliciure de nickel à 650°C.

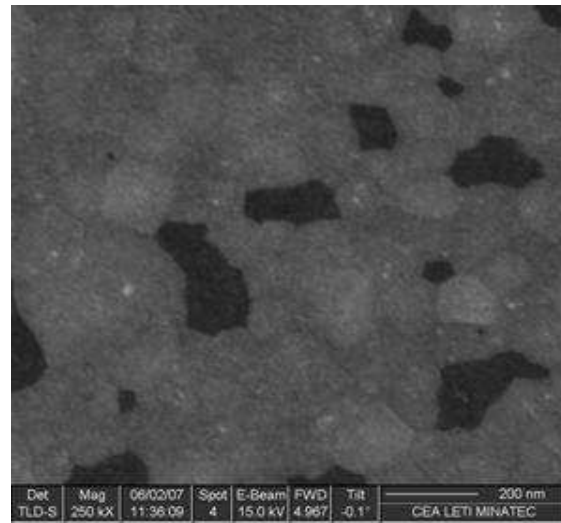


figure 1.4 : Démouillage en phase solide d'un film de NiSi pour un budget thermique de 600°C 10 minutes.

L'agglomération du siliciure va dégrader les performances du transistor via l'augmentation de la résistance d'accès. En effet le courant de conduction I_{ON} des transistors n'est pas seulement limité par la résistance intrinsèque de canal (R_{ch}), mais aussi par les résistances d'accès dont les composantes sont décrites dans la figure 1.5.

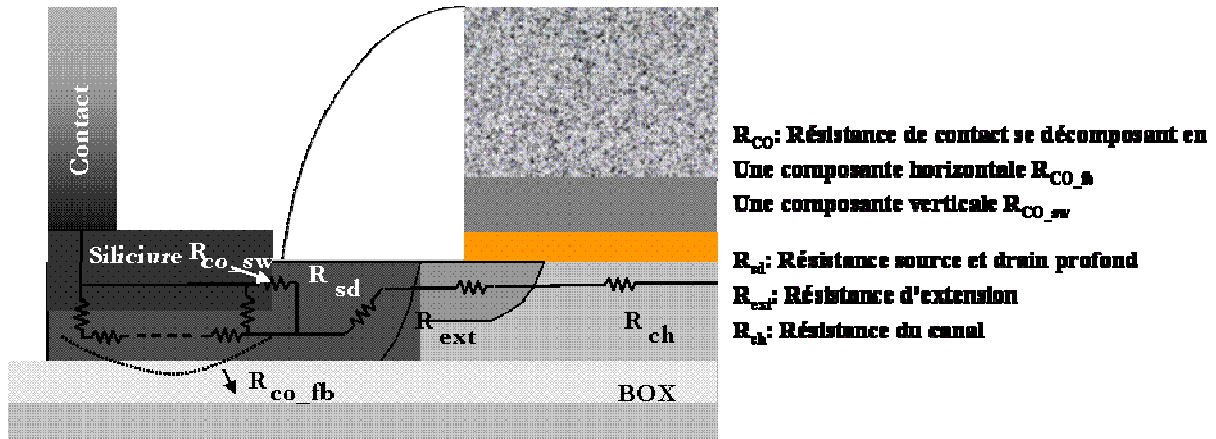


figure 1.5 : Schéma équivalent des résistances séries du module de jonction d'un transistor MOS dans le cas d'une technologie SOI

Avec la réduction des dimensions, ces résistances séries ne sont plus négligeables devant R_{ch} et parmi ces résistances parasites, la résistance de contact entre le siliciure et le silicium joue un rôle prépondérant [Dumont07, Kim04].

Or la minimisation de la contribution de la résistance de contact est fortement liée à la qualité du siliciure. On entend par qualité : le caractère continu de la couche ; une couche discontinue conduit à une augmentation de la résistance de la couche de siliciure mais aussi à une valeur de résistance de contact intermédiaire entre une résistance de contact métal/siliciure et une résistance de contact métal/silicium. La stabilisation du siliciure est donc un développement indispensable afin de ne pas augmenter la contribution de la résistance de contact aux résistances d'accès.

B.1.1 - Mise au point d'un siliciure adapté à l'intégration 3D

Malgré son instabilité morphologique, le NiSi est le siliciure retenu depuis le nœud 65 nm et intégré en production à ce stade. De nombreuses études font état de procédés permettant de stabiliser morphologiquement le monosiliciure de nickel. La méthode la plus utilisée et intégrée dans l'industrie est la stabilisation du NiSi par ajout d'un faible pourcentage de platine (5 à 10%) au nickel [Imbert05] [Mangelinck99] [Cheng00] [Detavernier04]. En marge du platine, l'effet stabilisateur d'autres éléments tels que le tungstène a été démontré [Dedytsche07] [Detavernier06].

Les hypothèses de mécanismes expliquant le gain en stabilité de ces deux procédés sont nombreuses. Dans le cas de l'addition du Pt, il s'agirait d'une modification de la texture de la couche de monosiliciure [Detavernier04] ainsi que d'une diminution de la diffusion du Ni retardant l'agglomération.

Dans le cas de l'addition de W, le retard au démouillage peut s'expliquer par effet mécanique. Le W, élément non soluble dans le NiSi, se redistribue aux joints de grain formant des barrières, isolant les grains entre eux, ainsi la fusion des grains est freinée et l'agglomération retardée [Detavernier06].

Dans un premier temps, nous avons cherché à quantifier l'apport en stabilisation de l'ajout de Pt en association à l'addition de W et F. La description des étapes de la siliciuration à stabilité optimisée est donnée en figure 1.6.

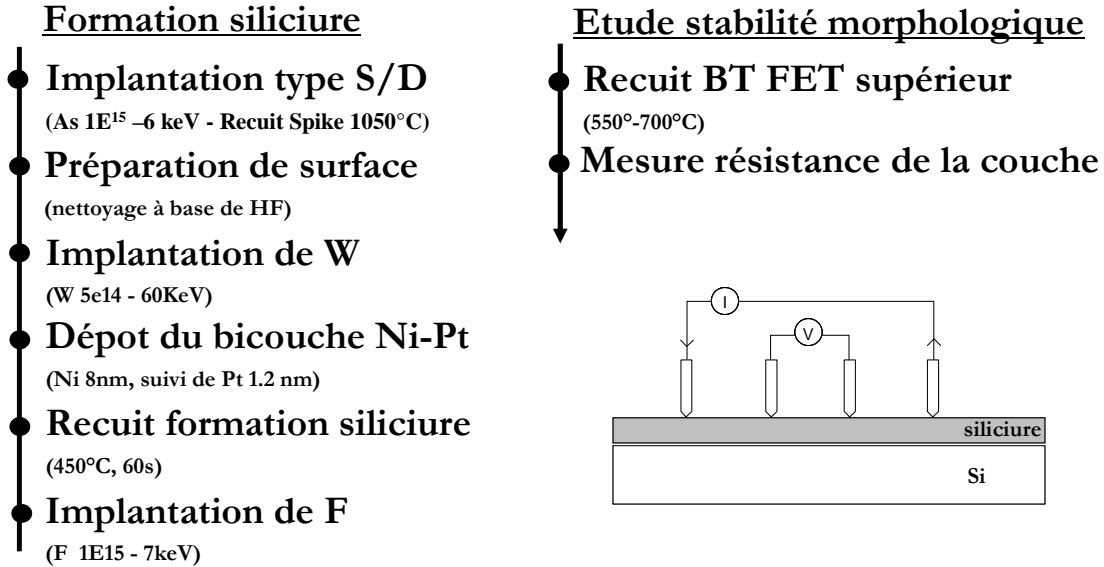


figure 1.6 : Description des étapes de réalisation du siliciure avec incorporation de Pt et addition de W et de F.

Le W est implanté dans la couche de silicium après une préparation de surface (nettoyage à base de HF), puis un empilement de 8 nm de Ni et de 1.2 nm de Pt a été déposé (pourcentage atomique de platine autour de 10%). Le recuit de siliciuration est effectué à $450^{\circ}C$ pendant 60s. Enfin le Fluor est implanté dans le siliciure.

Les échantillons siliciurés ainsi formés sont soumis à des recuits simulant le budget thermique du transistor supérieur, les mesures de résistances des couches par mesure quatre points pour un recuit à $650^{\circ}C$ sont présentées en figure 1.7 et en figure 1.8.

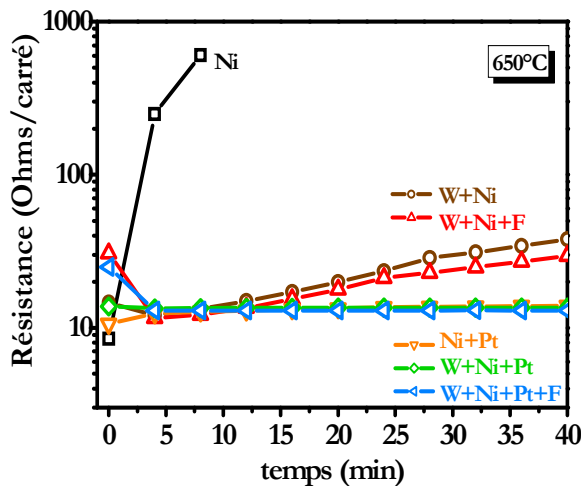


figure 1.7 : Mesure de la résistance du siliciure et suivi de son évolution pour des recuits à $650^{\circ}C$ montrant l'influence de l'addition de W, F et de Pt

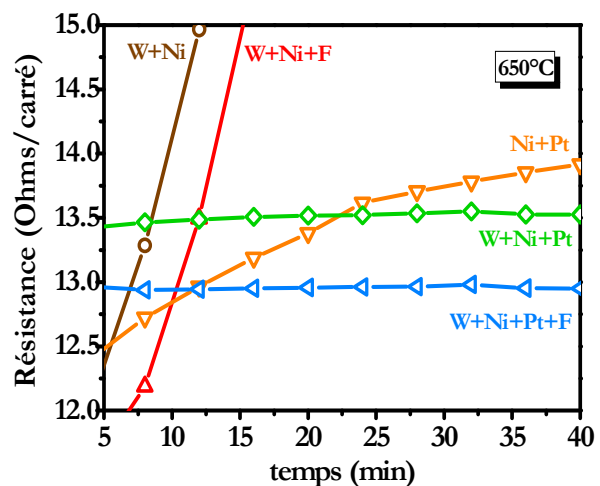


figure 1.8 : Zoom de la figure 1.7, mettant en évidence l'influence de l'addition de W et F au siliciure NiPtSi.

Nous observons que toutes les options proposées ont un effet stabilisateur sur le siliciure. Cependant les variantes incorporant du Pt permettent de maintenir des résistances de couche plus faibles. Le zoom de la figure 1.7 présenté en figure 1.8 met en valeur l'effet cumulatif de la stabilisation apportée par l'addition du W et du F dans le NiPtSi. A cette température, cet effet est mineur, mais il est révélé à plus haute température comme le démontre la figure 1.9. L'effet

stabilisateur de l'addition de platine associé à l'implantation de W et de F est confirmée par les observations MEB qui montrent un film de siliciure continu (cf . figure 1.10)

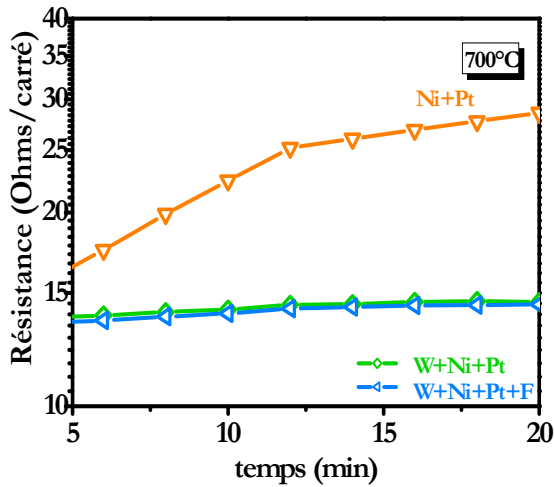


figure 1.9: Influence de l'addition de W et F dans le siliciure NiPtSi à 700°C

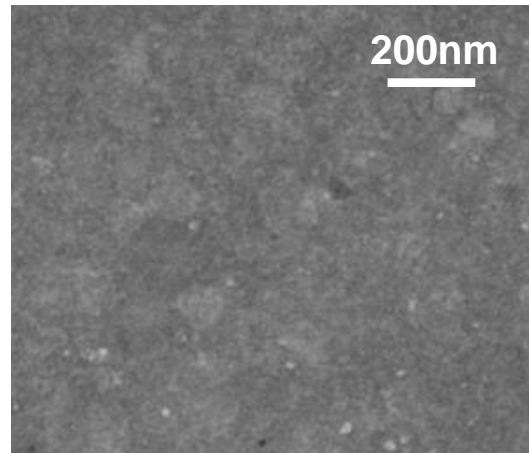


figure 1.10: Observation MEB de la couche de siliciure de nickel optimisé (NiPt+W+F) après recuit à 650°C, pendant 40 minutes

Bien que notre budget thermique maximum soit de 650°C, et qu'à cette température, la stabilité apportée par l'incorporation de Pt semble suffire, nous retiendront l'option avec incorporation de Pt ainsi que W et F apportant une stabilisation morphologique supérieure en prévision des budgets thermiques des étapes du back end (~400°C, 4 heures).

En conclusion, nous observons que la stabilité du morphologique du siliciure de nickel ainsi optimisé est assurée jusqu'à un budget thermique de 700°C sur un substrat de silicium sans motif. Le comportement du siliciure sur pleine plaque et sur motifs comme dans une intégration transistor peut varier fortement. Pour s'assurer de la compatibilité d'intégration, il est nécessaire d'aller jusqu'à l'intégration transistor. Les résultats obtenus seront décrits dans le chapitre II portant sur l'intégration complète.

B.1.2 - Etat de l'art des siliciures dans l'intégration 3D séquentielle

Le siliciure est généralement formé par le procédé dit « auto-aligné » SALICIDE (« Self-Aligned Silicide ») qui permet de silicier simultanément les sources/drains et les grilles sans avoir besoin d'une étape additionnelle de lithographie [Lau82]. Le principe de la siliciuration appliqué aux transistors CMOS est représenté dans la figure 1.11 (exemple dans le cas NiSi).

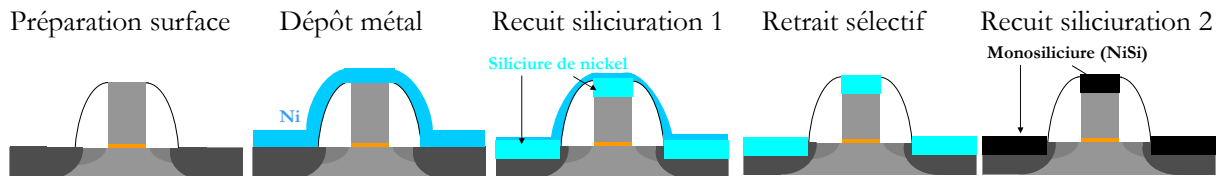


figure 1.11 : Représentation schématique des étapes du procédé SALICIDE

La première étape est constituée d'une préparation de la surface du silicium avant le dépôt de la couche métallique. Le dépôt d'une couche de nickel est suivi par un premier recuit de façon à faire réagir le Ni(Pt) avec le silicium. Un siliciure riche en nickel est alors formé. L'excès de nickel est retiré par retrait chimique sélectif. Enfin la couche de siliciure subit un second recuit afin de se transformer totalement en monosiliciure (NiSi).

A notre connaissance, la siliciuration des accès du transistor inférieur avec le procédé SALICIDE n'a pas été démontrée par d'autres équipes. Seul Samsung a présenté une alternative à ce procédé en ne siliciurant que l'interface de silicium au niveau des contacts comme montré dans la figure 1.12. Cette siliciuration localisée, supprime la problématique de stabilisation morphologique du siliciure puisque cette étape est réalisée à la fin des étapes Front End. Pour réaliser cette siliciuration, le cobalt est alors utilisé à la fois pour former le siliciure (CoSi_x) et remplir les contacts. Cette siliciuration permet de diminuer la résistance de contact par rapport à un contact Ti/TiN directement sur silicium (cf. figure 1.13) et ainsi permettre une augmentation de l'ordre de 20% du courant I_{ON} des transistors (figure 1.14) [Jung05].

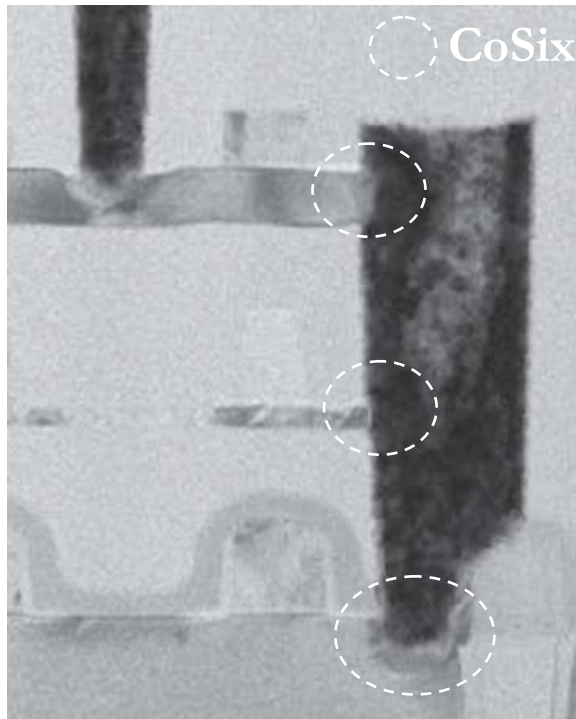


figure 1.12: Vue en coupe d'une cellule 3D séquentielle [Jung05] avec siliciuration des accès au niveau du contact. Le cobalt est utilisé à la fois comme métal de siliciuration et comme métal de remplissage des contacts.

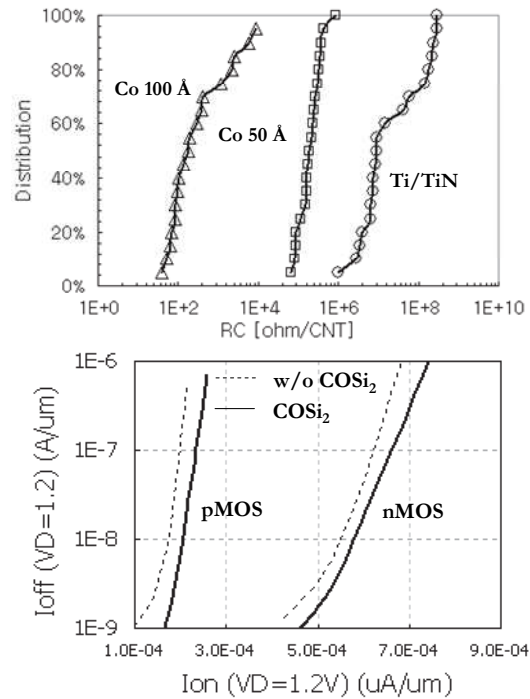


figure 1.13: Distribution cumulée des résistances de contact dans le cas de contacts siliciurés avec du cobalt ou sans siliciuration (dépôt direct de la barrière Ti/TiN) [Jung05]

figure 1.14: Caractéristiques Id-Vg pour un transistor avec ou sans contacts siliciurés [Jung05]

Cette stratégie semble à première vue astucieuse, cependant le contrôle de la phase de siliciure formée est plus compliqué. En effet la présence d'une source de métal infini peut entraîner la formation de phases riches en cobalt non désirées ainsi qu'une siliciuration excessive (risque de court-circuit entre les accès et le canal).

De plus cette solution va entraîner une dégradation de la résistance série par rapport à un procédé SALICIDE. Les composante R_{co_fb} et R_{sd} vont augmenter : R_{co_fb}, du fait de la réduction de la surface de contact et R_{sd} du fait de l'allongement de la zone de source drain à traverser pour atteindre le canal (cf. figure 1.15 et figure 1.16).

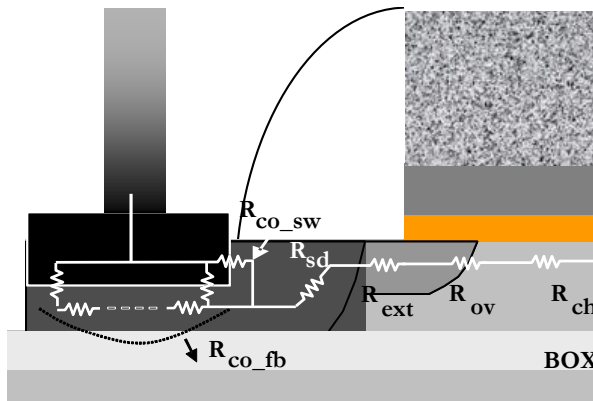


figure 1.15: Représentation schématique des résistances séries du module de jonction dans le cas de l'utilisation du procédé SALICIDE

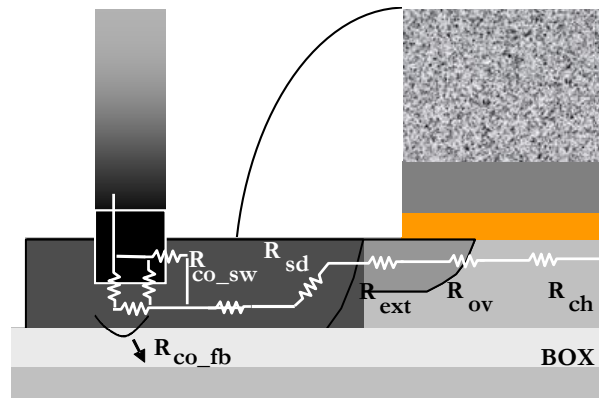


figure 1.16: Représentation schématique des résistances séries du module de jonction dans le cas de l'utilisation de la siliciuration dans les contacts

Les travaux présentés ont été réalisés en collaboration avec le laboratoire « Front End » du LETI

B. 2 -Stabilisation du module de jonction

Le budget thermique du FET supérieur peut dégrader les caractéristiques du module de jonction via deux mécanismes : le mécanisme de diffusion des dopants qui peut altérer le contrôle des effets canaux courts et le mécanisme de désactivation des dopants réduisant le niveau de courant de conduction du transistor.

Concernant le phénomène de diffusion, le budget thermique pour la réalisation du FET supérieur étant limité à 650°C pendant 5 minutes, le phénomène est limité voire inexistant. En effet, si l'on observe le cas du bore, que l'on peut considérer comme un pire cas, car fortement sujet au phénomène de diffusion accélérée transitoire (TED : Transient Enhanced Diffusion), les travaux de Dachs et al [Dachs03] montrent qu'après une recuit à 750 °C pendant 300s (~5min), la profondeur de jonction ne s'est déplacée que de 1nm. Nécessairement à 650°C, la diffusion sera encore plus limitée, la rendant complètement négligeable.

La seconde problématique pouvant se poser lorsque l'on soumet le FET inférieur aux budgets thermiques du FET supérieur est une éventuelle désactivation de ses jonctions. Le phénomène de désactivation peut apparaître lorsqu'une partie des dopants activés le sont en étant situés dans un état métastable. Dans ce cas, l'apport ultérieur d'énergie sous forme d'un budget thermique peut faire repasser l'impureté (le dopant) en site interstitiel afin de minimiser l'énergie du réseau, et ainsi faire chuter le niveau d'activation électrique des dopants. Or la plupart des techniques d'activation alternative (FLASH, Laser, SPER) à l'activation thermique classique des (Rapid thermal Annealing - RTA) conduisent à des jonctions partiellement métastables, donc sujettes à désactivation. En effet, ces techniques actuellement à l'étude afin d'obtenir des jonctions ultrafines (meilleure gestion de la diffusion des dopants) tout en obtenant des niveaux d'activation très élevés, entraînent des niveaux d'activation parfois supérieurs aux valeurs de solubilités limites des dopants dans le semiconducteur à l'équilibre thermodynamique.

Le phénomène de désactivation est un phénomène complexe dont l'amplitude dépend du type de dopant, de la dose implantée, du type de procédé d'activation et du budget thermique post activation subit par les jonctions. Par exemple, dans le cas d'une jonction dopée Bore et activée par procédé SPER, la diminution de la résistance de la couche est de l'ordre de 10% pour un post-recuit à 650°C pendant 90 minutes et de 3% après 800°C pendant 2minutes. [Dachs03].

Nous observons que l'amplitude de la désactivation dépend de la température ainsi que du temps de recuit. Dans le cas d'une activation laser, le phénomène de désactivation du bore apparaît à partir de 800°C alors que pour le phosphore, la désactivation est non négligeable dès 200°C. En conséquence, il nous est difficile de statuer sur le phénomène de désactivation pouvant apparaître dans le cas de recuits alternatifs.

Cependant, l'activation des jonctions du démonstrateur a été réalisée avec un recuit thermique RTA d'une durée de 1s à haute température (1050°C) utilisant des rampes de montée et descente relativement lentes (45°C/s). Dans ce cas l'activation des jonctions est stable, et le budget thermique nécessaire à la réalisation du FET supérieur ne conduit à aucune désactivation des dopants.

Ce qu'il faut retenir

II-A : Stabilité du FET inférieur

- Le budget thermique nécessaire pour réaliser le FET supérieur, même limité à 650°C, va dégrader le siliciure de nickel NiSi du transistor inférieur.
- A cette température, la dégradation de la couche correspond en premier lieu à un démouillage en phase solide du film mince de siliciure. L'apparition de discontinuité dans le film va augmenter la résistance de la couche et ainsi que dégrader la résistance de contact.
- Le procédé développé au Leti associant incorporation de platine (5%) et addition de tungstène ainsi que de fluor permet de stabiliser morphologiquement le siliciure jusqu'à 700°C.
- L'absence de siliciure (ou la dégradation du siliciure) a un impact considérable sur les niveaux de courant des transistors. Il est indispensable de développer un procédé de stabilisation du siliciure pour l'intégration 3D séquentielle afin que le transistor inférieur ait les performances requises pour des nœuds sub-45nm.
- L'option de siliciurer les accès seulement au niveau des contacts semble moins intéressante que l'utilisation du procédé SALICIDE.
- L'intégration d'un siliciure (avec le procédé SALICIDE) au niveau inférieur d'une structure 3D séquentielle est une première.
- Avec un budget thermique pour le FET supérieur limité à 650°C, le phénomène de diffusion dans les jonctions du FET inférieur est limité voire inexistant.
- Si les jonctions du transistor inférieur sont réalisées avec un procédé conduisant à des jonctions métastables, la fabrication du FET supérieur peut entraîner une désactivation partielle des dopants.
- Il est nécessaire d'étudier l'impact de la désactivation avec un budget thermique de 650°C sur des jonctions activées avec des techniques d'activations alternatives à l'activation thermique.

C - Réalisation de la zone active supérieure

C.1 - Les techniques

Les performances des transistors dépendent de la qualité du substrat semiconducteur sur lequel ils sont fabriqués. Idéalement ce substrat doit être monocristallin, sans défaut cristallin (dislocation) et sans particules contaminantes.

Comme nous l'avons vu dans l'introduction, les premiers démonstrateurs de l'intégration 3D séquentielle possédaient des zones actives supérieures en poly-silicium. Par des méthodes de recristallisation, les tailles des grains obtenus ont été suffisantes pour qu'au niveau du transistor, le film d'active soit monocristallin localement. Cependant à l'échelle du circuit, la présence de différentes orientations va induire une variation de performances d'un transistor à l'autre et la présence de joint de grain va diminuer le rendement de transistors fonctionnels. A l'heure où la variabilité des performances des transistors est une problématique majeure dans la fabrication des circuits intégrés, les techniques conduisant à des substrats poly-cristallin ne semblent pas prometteuses.

En conséquence, nous nous concentrerons sur les techniques permettant la réalisation de substrats monocristallins. Elles peuvent être regroupées en deux familles ; Une première famille où l'information cristalline est transférée du niveau inférieur au niveau supérieur via une fenêtre d'ouverture (seed-window) comme décrit dans la figure 1.17. La seconde consiste en un report d'une couche originellement monocristalline au dessus du diélectrique inter-niveau.

C.1.1 - Les techniques de type « Seed-window »

La figure 1.17 récapitule les techniques majeures de la première famille présentées dans la littérature 3D séquentielle (après les années 2000).

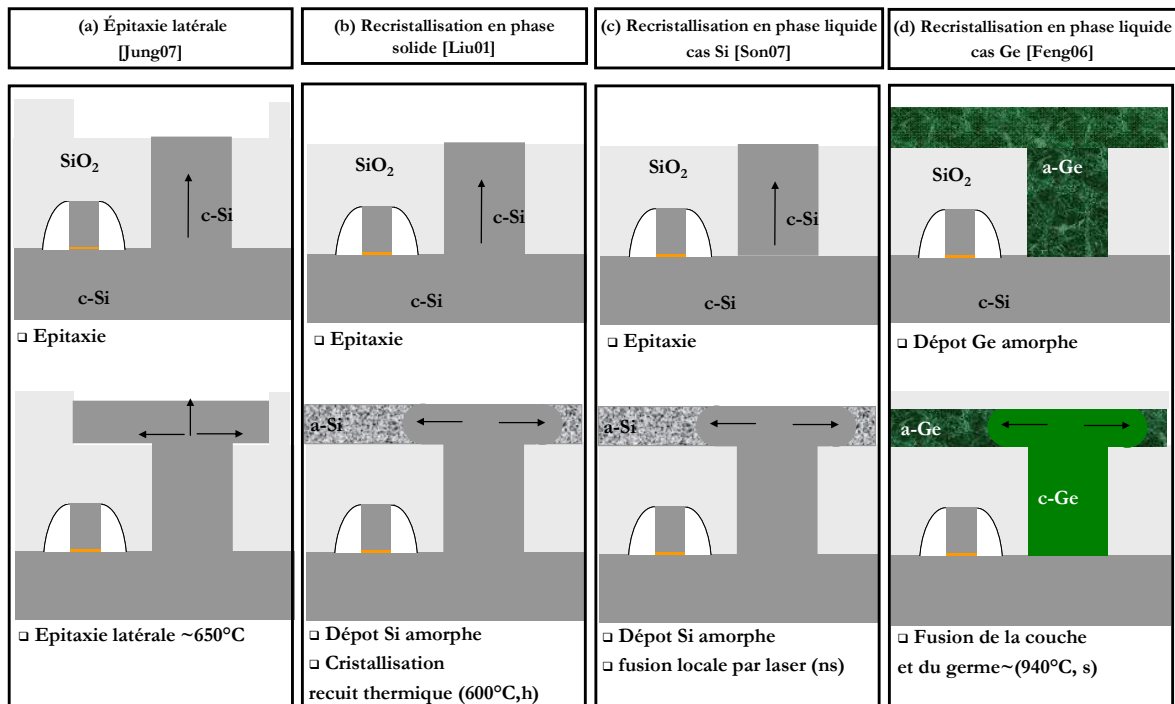


figure 1.17 : Présentation des techniques de réalisations de zone active supérieure présentées dans les références majeures de la littérature 3D séquentielle.

(a) Épitaxie latérale : (ELO : Epitaxial lateral overgrowth ou SEG : Selective epitaxial growth..)

Dans un premier temps, une étape d'épitaxie sélective permet de faire croître verticalement un germe monocristallin dans la fenêtre d'ouverture sur le silicium du niveau inférieur. Quand le germe atteint le haut de la fenêtre d'ouverture, l'épitaxie se poursuit latéralement. Les vitesses de croissance latérales et verticales diffèrent. Afin d'obtenir un film d'épaisseur uniforme, une étape de planarisation est nécessaire. Pour réduire la dispersion induite par cette étape, Jung et al propose de définir des motifs dans le diélectrique délimitant la zone active supérieure comme représentée dans la figure 1.17. Le polissage mécano-chimique (CMP : chemical Mechanical Polishing) est alors réalisé avec arrêt sur l'oxyde. Cette épitaxie est réalisée à 650°C, cette méthode est donc conforme aux exigences de budget thermique limité pour l'intégration 3D séquentielle.

(b) Recristallisation en phase solide (LSPE : Lateral Solid Phase Epitaxy)

Dans [Liu01], la réalisation de la zone active utilise successivement deux techniques. Dans un premier temps une épitaxie sélective est réalisée jusqu'à atteindre le haut de la fenêtre d'ouverture, puis un dépôt de silicium amorphe est réalisé. La recristallisation en phase solide du film est obtenue à l'aide d'un recuit thermique à 600°C. Un recuit de 6 heures est nécessaire pour réaliser une zone active recristallisée de 10 µm de long.

(c) Recristallisation en phase liquide (LEG : Laser induced Epitaxial Growth)

Son et al [Son07] procède de manière similaire à [Liu01] : après une étape d'épitaxie sélective pour atteindre le haut de la fenêtre d'ouverture, ils réalisent un dépôt de silicium amorphe. La recristallisation est cette fois réalisée en phase liquide. La fusion du silicium du film est obtenue à l'aide d'un laser qui chauffe localement le film. La durée du recuit est extrêmement faible (150ns).

(d) Recristallisation en phase liquide (cas du germanium) (Lateral Phase Epitaxy, LPE, Rapid Melt Growth, RMG.)

Dans cette technique, un dépôt de germanium amorphe est réalisé après l'ouverture de la fenêtre sur la zone active inférieure. Après encapsulation, la température du substrat est amenée au dessus de la température de fusion du germanium (937°C). Pendant la phase de refroidissement, l'orientation cristallographique du substrat est propagée de l'interface Ge/ Si à l'ensemble de la couche. Le recuit ne dure que quelques secondes, cependant, à cette température, des phénomènes d'agglomération du silicium et de désactivation et de diffusion des dopants du transistor inférieur sont attendus. La température de fusion du silicium (1415°C), limite l'utilisation de cette technique à la réalisation de zones actives supérieures en germanium.

C.1.2 - Le report de couche

L'autre famille permettant d'obtenir une zone active monocristalline est basée sur la technique de collage moléculaire. Ce phénomène est observé quand deux surfaces parfaitement lisses, propres et non courbées sont mises en contact. Les surfaces adhèrent l'une à l'autre à température ambiante, sans l'utilisation de colles adhésives.

La démonstration du report d'une couche de Si ou de Ge par collage moléculaire au dessus d'un premier niveau de transistor a été démontrée dans [Yu04] et [Xue03] respectivement.

Les deux équipes utilisent le procédé Smart CutTM pour transférer ensuite une fine couche d'active supérieure sur le substrat support sur lequel le premier niveau de transistor a été préalablement réalisé.

La figure 1.18 présente le report d'une couche de Germanium réalisé par Yu et al [Yu04]. Ce report a été réalisé au dessus d'un niveau de transistors et de ses interconnexions. (cf. figure 1.18(a)). Sur la figure 1.18(b), nous observons que le transfert de la zone active supérieure est partiel (localisé au dessus des plots de métal du niveau inférieur). Ce transfert partiel s'explique par la non planarisation de l'oxyde de collage qui a une topographie de l'ordre du μm (cf. figure 1.18(c)).

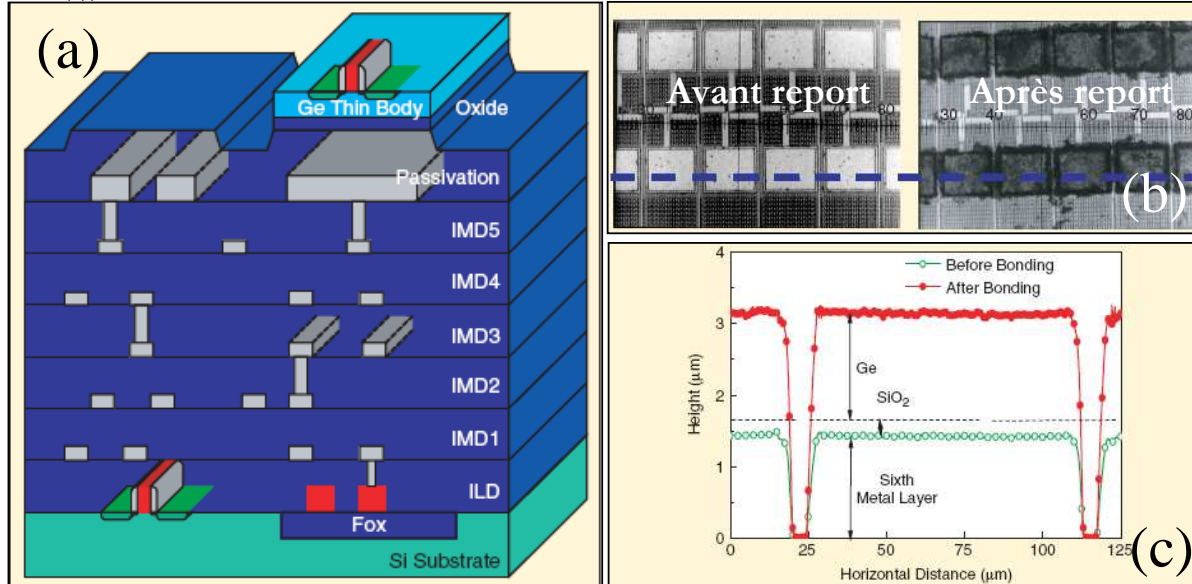
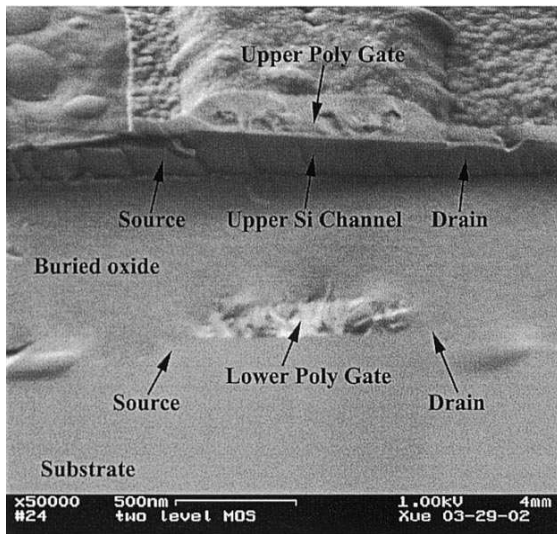


figure 1.18 : Réalisation d'un transistor GeOI au dessus des interconnexions. La zone active Ge est obtenue par collage et procédé SMART CUT. [Yu04]



Xue et al, ont aussi démontré la réalisation d'une active supérieure en utilisant le collage. La figure 1.19 présente une vue en coupe de l'empilement réalisé.

Nous observons que la zone active a un aspect relativement perturbé. Aucune information n'est donnée sur la qualité du collage.

figure 1.19: Observation de la coupe d'un empilement de transistors. La zone active supérieure est en silicium. Elle est obtenue par collage puis clivage avec implantation Hydrogène. [Xue03]

Dans ces deux références, les transistors inférieurs et supérieurs ne sont pas interconnectés. En conséquence aucune cellule 3D n'a pu être démontrée avec le report de couche.

C.1.3 - « Seed window » versus report de couche

Par rapport aux procédés de type « seed-window » précédemment décrits, le report basé sur le collage possède de nombreux avantages.

En premier lieu, il permet d'obtenir des films monocristallins correspondants aux exigences de qualité requises pour des applications industrielles [ITRS]. La qualité cristalline du film reporté dépend uniquement de la qualité du substrat de départ car le procédé de collage ne crée pas de dislocations. La présence de dislocation diminue le rendement de dispositifs fonctionnels (fuite de diode et de courts circuits entre la source et la grille).

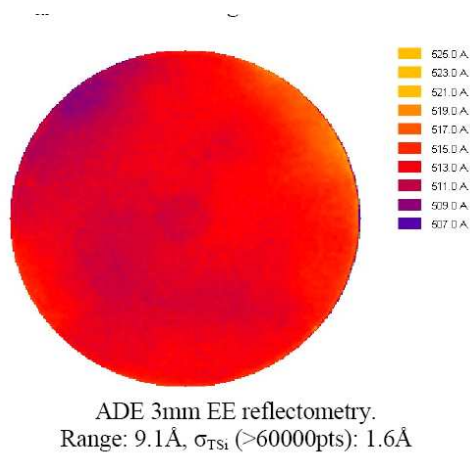


figure 1.20 : Cartographie d'épaisseur de zone active d'un substrat SOI 300mm. [Weber08]

De plus ce procédé permet un contrôle très précis de l'épaisseur reportée à l'angström près (cf. figure 1.20).

Le contrôle de l'épaisseur du film est essentiel pour limiter la dispersion des performances des transistors. La figure 1.21 illustre l'impact de la variation d'épaisseur sur la tension de seuil des transistors FDSOI. Pour une longueur de grille de 25 nm, une modification de 1 nm d'épaisseur entraîne une modification de 30 mV environ. Pour pouvoir maintenir un écart type de la variation de V_{TH} à 3% de V_{DD} pour les générations avenir, Weber et al ont montré que l'écart type de l'épaisseur de silicium doit être de l'ordre de quelques angstrœms pour une plaque de 300mm (cf. figure 1.22).

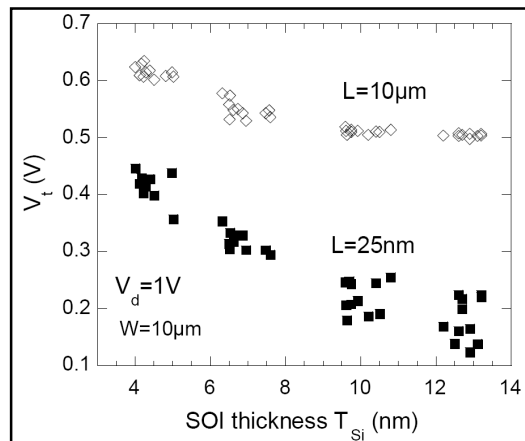


figure 1.21 : Corrélation entre la dispersion de l'épaisseur du film de zone active et la dispersion en tension de seuil des transistors FDSOI [Weber08]

	Year	'08	'10	'12	'14	'16	'18	'20
ITRS roadmap [20]	L (nm)	38	28	22	18	14	12	10
	T_{ox} (nm)	1.6	1.4	1.2	1.1	0.9	0.8	0.8
	V_{dd} (V)	1.1	1	1	0.9	0.8	0.8	0.8
	σ_{Vt} (mV) (3% V_{dd} - Q_{ox} contribution)	31	28	28	25	25	22	22
	σ_{Tox} (4% T_{ox} 3 σ) (Å)	0.22	0.19	0.16	0.15	0.12	0.11	0.11
	This work	DIBL (mV/V)	100	100	100	100	100	100
T_{Si} (nm)		8	6.5	6	5.5	5	4.5	4
σ_L (12% L 3 σ) (nm)		1.48	1.12	0.88	0.72	0.56	0.48	0.4
$\sigma_{TSi,max}$ (Fig.19) (Å)		7.5	4.4	3.6	2.8	2.7	2.4	2.2
σ_L constant (nm)		1	1	1	1	1	1	1
$\sigma_{TSi,max}$ (Fig.19) (Å)		8	4.5	3.5	2.5	2.2	1.8	1.7

figure 1.22 : Roadmap des épaisseurs de Silicium pour dispositifs FDSOI basé sur l'ITRS [Weber08]

L'utilisation de méthodes type « seed-window » nécessitant une planarisation par CMP semble compromettre un contrôle de l'épaisseur de zone active compatible avec des exigences industrielles pour des applications de type FDSOI sur le niveau supérieur.

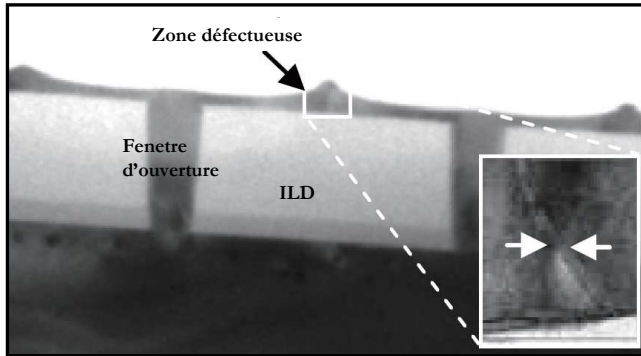


figure 1.23 : Observation d'une zone active supérieure réalisée par Laser induced Epitaxial Growth [Son07]

recristallisées à l'aide d'un germe est limitée à cause de phénomènes parasites (phénomène de nucléation aléatoire, dans cas de la LSPE, perte de l'information cristalline dans le cas de la SEG..etc).

Enfin, il faut noter que le collage moléculaire, contrairement aux techniques SW, offre la possibilité de co-intégrer différentes orientations de surface. Par exemple, le nMOS inférieur peut être réalisé sur un substrat <100> et le pMOS supérieur sur un substrat <110> sans étapes de procédé additionnels. Le choix des orientations de manière indépendante pour le nMOS et le pMOS permet d'optimiser leurs performances.

D'autre part, il faut remarquer que dans les techniques de type seed-window, la zone active inférieure doit rester à nu au niveau des fenêtres d'ouverture. La siliciuration de ces zones doit alors être supprimée ou, si l'étape de siliciuration est maintenu, des étapes supplémentaires doivent être effectuées : protection des zones de germe pendant l'étape de siliciuration, retrait de la protection avant l'étape de réalisation de la fenêtre d'ouverture.

Cette particularité de l'intégration de type seed-window est peut être aussi la raison qui a poussé Samsung à ne pas réaliser une siliciuration classique avec le procédé SALICIDE mais à développer une technique alternative : la siliciuration localisée dans les fonds de contacts.

En conclusion, le collage semble la méthode la plus adaptée à la réalisation de zones actives supérieures pour des nœuds avancés.

Le Leti a acquis une grande expérience dans le collage moléculaire afin de réaliser des substrats SOI [Bruel95, Moriceau01]. Cependant de nouvelles problématiques apparaissent quand ce collage doit être fait au dessus d'un niveau de transistors. La partie suivante décrit les développements réalisés pour l'adapter à la réalisation de zones actives supérieures dans l'architecture 3D séquentielle.

Un autre critère de distinction du collage par rapport aux techniques précédemment citées est la taille de la zone active utilisable pour réaliser des circuits. En effet la surface utilisable pour les transistors est amputée de la surface occupée par les fenêtres d'ouverture. De plus les zones où se rejoignent les fronts de cristallisation peuvent elles aussi être défectueuses (cf. figure 1.23).

En pratique la taille des zones

C. 2 -Le collage dans l'intégration 3D séquentielle

C.2.1 - Les nouvelles problématiques

La figure 1.24 présente le procédé de réalisation de l'active supérieure tel qu'il a été réalisé au Leti.

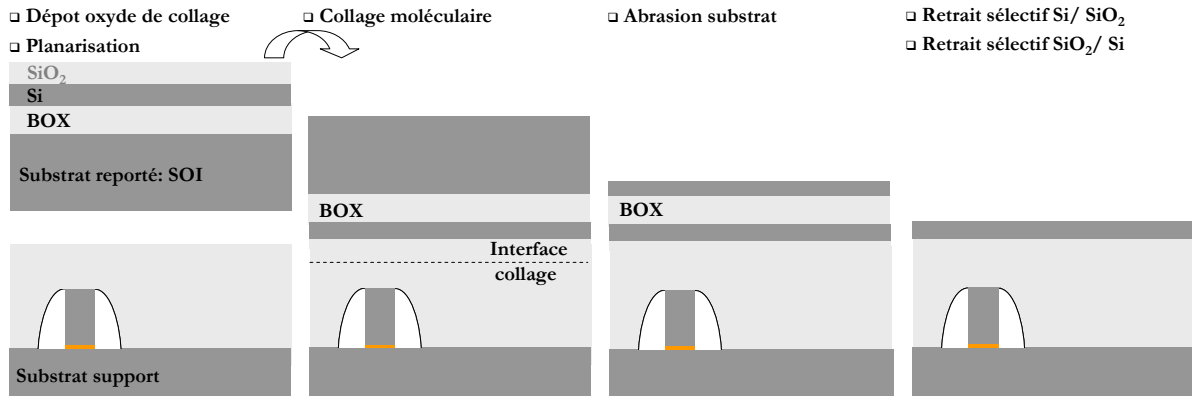


figure 1.24 : Description schématique des étapes essentielles des techniques de réalisation du film d'active supérieur par collage moléculaire.

Le collage effectué est un collage hydrophile oxyde/oxyde. De part la présence du transistor inférieur, l'oxyde de collage du substrat support est nécessairement déposé. Un oxyde de silicium de type High Density Plasma (HDP) est déposé sur la plaque support et sur la plaque reportée. La topographie du diélectrique sur la plaque support doit être supprimée afin de pouvoir mettre en regard deux surfaces planes. Un traitement de surface permet de saturer les deux interfaces avec des groupements hydroxyles. Le collage est réalisé à température ambiante et initialisé manuellement. A cette étape, l'adhérence entre les deux plaques est assurée par des liaisons hydrogènes entre les molécules d'eau adsorbées à la surface. Le recuit de solidification de l'interface (200°C, 1 heure) permet de créer des liaisons covalentes entre les deux oxydes. L'eau de l'interface est dispersée dans l'ensemble du diélectrique qui se comporte alors en quelque sorte comme un réservoir. Le silicium du substrat reporté est éliminé par abrasion mécanique puis par attaque chimique (TMAH) avec arrêt sur l'oxyde enterré. Cet oxyde sera retiré sélectivement par rapport au Si, laissant la fine couche de silicium collée au dessus du transistor inférieur. Le transistor supérieur peut alors être réalisé sur cette nouvelle zone active vierge.

Pour obtenir le collage des deux plaques, les oxydes de collage doivent répondre à certains critères récapitulés dans la figure 1.25.

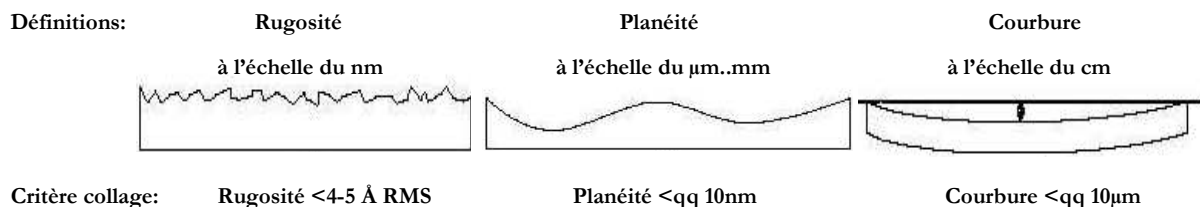


figure 1.25 : Définitions de la courbure, planéité et microrugosité d'une plaque.

Le critère le plus important est la rugosité de surface. C'est à l'échelle micrométrique qu'apparaissent des microrugosités. La mesure de la rugosité de surface est réalisée à l'aide d'un AFM (Atomic Force Microscope). La rugosité limite pour la réalisation d'un collage moléculaire

est de 0.4-0.5nm RMS (RMS pour Root Mean Square, soit la moyenne quadratique de la hauteur pour l'analyse de la topographie de l'image).
L'oxyde de type HDP choisi répond bien à ce critère comme le montre la mesure de rugosité présentée en figure 1.26.

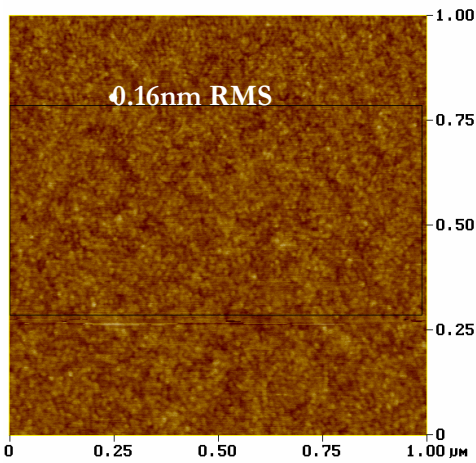


figure 1.26: Mesure de la rugosité de l'oxyde HDP après dépôt.

Pour obtenir la planéité de l'oxyde de collage côté substrat support, une étape de polissage mécano-chimique est réalisée. Les relevés profilométriques (cf. figure 1.27) ont montré que la topographie après planarisation est inférieure à 5 nm (mesure dans le bruit). Avec une surface mal planarisée, il se peut que le collage ne s'initialise pas du tout ou que le transfert du film semiconducteur soit partiel.

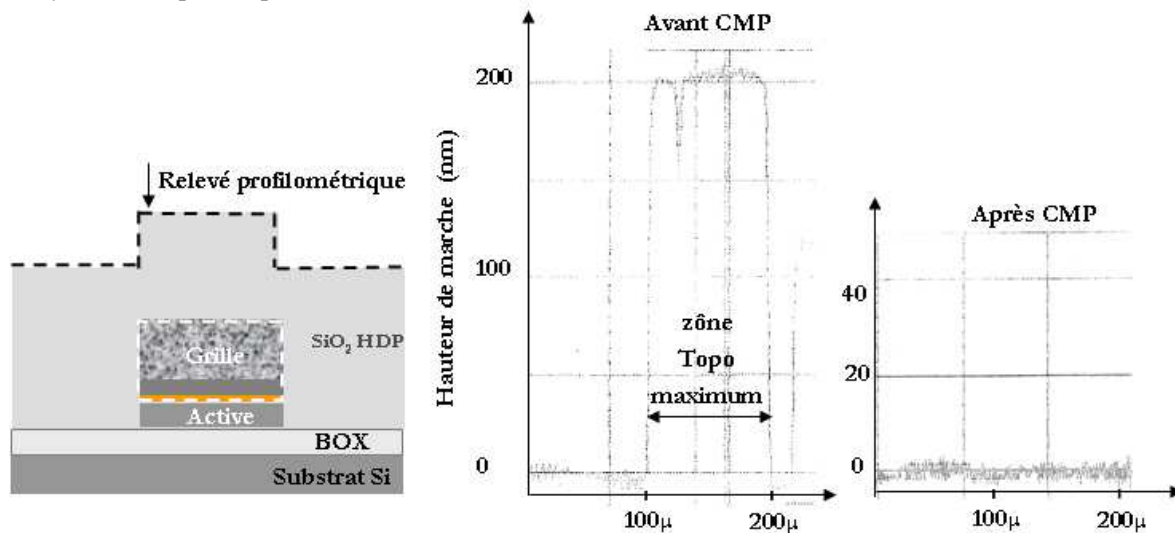


figure 1.27 : Mesure profilométrique de l'oxyde de collage côté substrat support après planarisation. (cas d'un motif du niveau inférieur avec topographie maximum: présence du niveau grille sur le niveau zone active).

L'étape de CMP permet de planariser la surface mais induit une dispersion d'épaisseur à l'échelle de la plaque non négligeable. La figure 1.28 illustre l'impact de l'étape de CMP sur la dispersion à l'échelle de la plaque. Partant d'un dépôt de HDP de 110 nm en moyenne sur une plaque vierge de silicium et une dispersion avec un écart type σ de 1%, le retrait de 70 nm d'oxyde par CMP conduit à une dispersion de 8%.

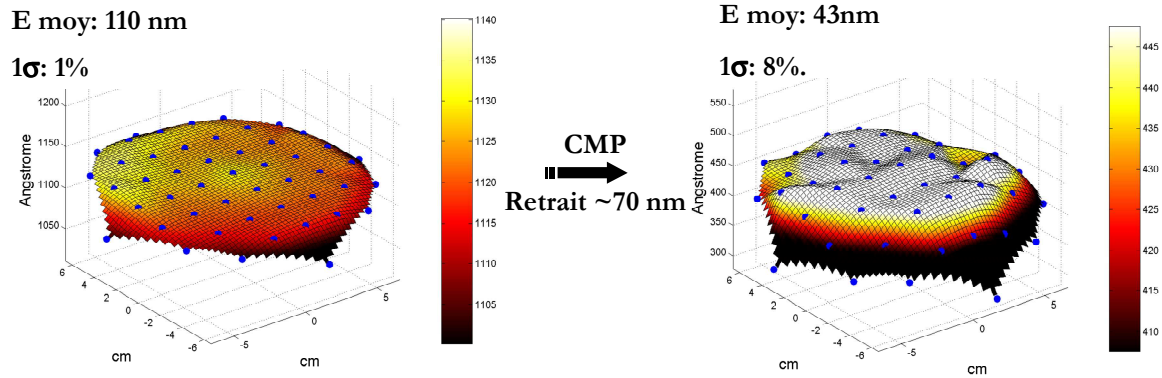


figure 1.28 : Evolution de l'épaisseur de diélectrique avant et après un amincissement du diélectrique par CMP (Mesure par ellipsométrie).

Les variations d'épaisseur de diélectrique induite par la CMP n'ont pas provoqué de difficulté à faire adhérer les deux plaques car la tolérance du collage à la courbure est de l'ordre de la dizaine de μm .

C.2.2 - Caractérisation du collage

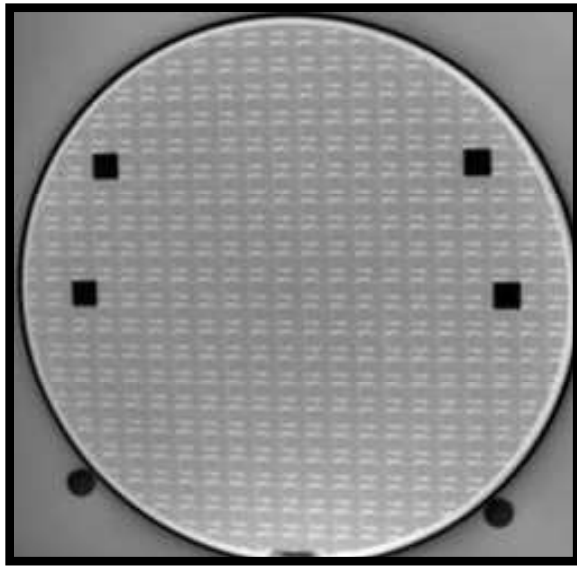


figure 1.29 : Caractérisation infrarouge du collage

Après la mise en adhérence des deux plaques et un recuit de densification à 200°C pendant une heure, la qualité du collage est caractérisée par diverses techniques.

Les défauts de collage peuvent être visualisés à l'aide d'une caméra infrarouge. Le silicium étant transparent aux rayonnements infrarouges, les zones enterrées non collées sont révélées, à condition que la distance entre les deux plaques soit supérieure à $0,5 \mu\text{m}$. A l'aide de cette méthode, aucun défaut de collage n'est détecté (cf. figure 1.29).

La microscopie acoustique à balayage (Scanning Acoustic Microscopy, SAM) est une technique qui permet de visualiser les zones non collées à l'interface de plus petites dimensions. Avec cette technique, la résolution latérale est 1 mm et verticale est 30 nm (fréquence 50 MHz). Le principe de la SAM repose sur la propagation d'une onde sonore au travers de différents milieux. Le changement d'un milieu d'un point de vue acoustique correspond à une variation de l'impédance acoustique du matériau traversé.

Lorsque l'interface de collage est fermée, il y a transmission de l'onde acoustique d'une plaque à l'autre. Dans le cas des défauts de collage qui peuvent contenir de l'eau, de l'air ou du vide, le changement de milieu provoque une rupture d'impédance et donc

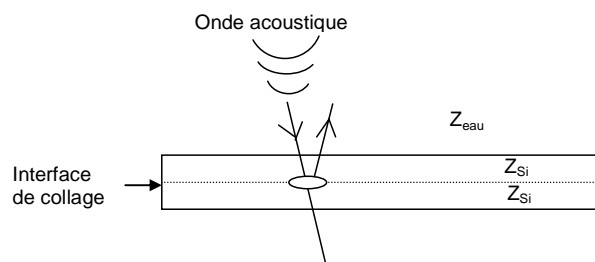


figure 1.30 : Descriptif du principe de fonctionnement de la caractérisation acoustique du collage.

une réflexion de l'onde (écho) (cf. figure 1.30). L'analyse de cet écho qui permet la visualisation des défauts.

L'analyse acoustique du collage au dessus du niveau inférieur de transistor n'a pas révélé de défauts de collage de surface importante comme le montre la figure 1.31.

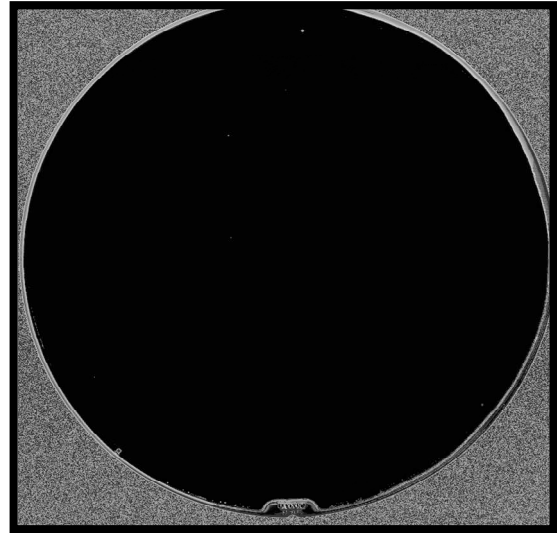


figure 1.31 : Caractérisation acoustique du collage

La qualité du collage peut être quantifiée au travers de la mesure de l'énergie de collage. La méthode de Maszara est la technique de mesure de l'énergie de collage la plus répandue dans la littérature. Elle consiste à insérer une lame au niveau de l'interface de deux plaques collées et de mesurer la longueur décollée via une caméra infrarouge. Cette mesure permet de remonter à l'énergie de collage, pour plus de détails, le lecteur pourra se reporter à [Maszara88]. La valeur de 900 mJ.m^{-2} extraite permet de conclure que le collage est de bonne qualité et permet d'assurer la tenue mécanique des collages pendant l'étape de retrait du substrat.

C.2.3 - Amincissement du diélectrique inter-niveau

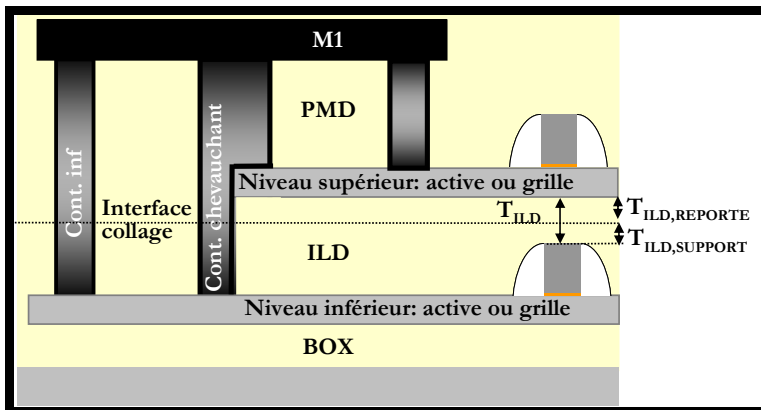


figure 1.32 : Description des différents types de vias dans une intégration 3D séquentielle et un seul niveau contact.

diélectrique pré-Métal (PMD : Pre-Metal Dielectric) et la couche de diélectrique inter-niveau (ILD : Inter Layer Dielectric). Nécessairement le facteur de forme de ce contact, c'est-à-dire le rapport entre la taille de l'ouverture du contact et la profondeur gravée va être augmenté. Pour limiter ce facteur de forme, nous chercherons à diminuer l'épaisseur du diélectrique inter-niveau. Cet aspect est d'autant plus essentiel que la taille des contacts influence énormément la densité des circuits intégrés avec des dimensions avancées (sub-65nm)¹.

Cependant, T_{ILD} doit posséder une borne inférieure, nécessaire pour rendre négligeable le couplage entre les deux transistors empilés. Nous montrerons, dans le chapitre IV partie A, que cette épaisseur se situe aux environs de 100 nm pour des longueurs de grille de 35nm.

¹ Nous reviendrons plus en détail sur la problématique de réalisation des contacts dans la partie F

Afin de réduire l'épaisseur du diélectrique pré-métal de l'intégration planaire à une épaisseur de 100 nm, nous avons procédé en deux étapes. Tout d'abord, l'empilement de diélectriques composant classiquement le diélectrique pré-métal au LETI a été modifié. Cet empilement est décrit dans la figure 1.33.

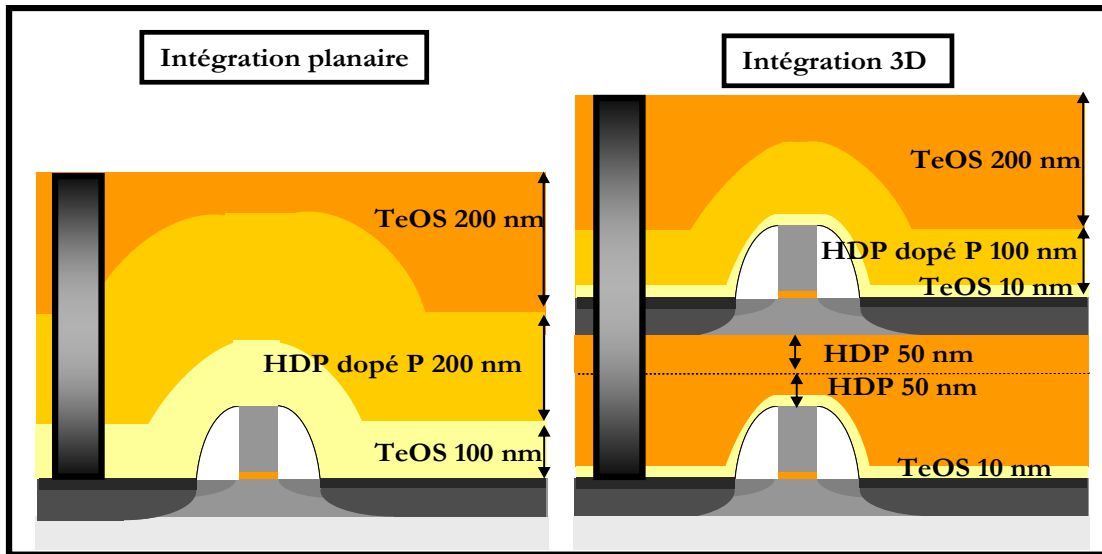


figure 1.33 : Description de la composition du diélectrique pré-métal tel qu'il est réalisé au Leti pour une intégration planaire et description de l'empilement de diélectriques PMD et ILD choisi pour l'intégration 3D séquentielle afin de limiter le facteur de forme des contacts atterrissant sur le niveau inférieur.

Le rôle de la couche d'oxyde de type HDP dopé Phosphore est de bloquer la diffusion des alcalins. Un seul exemplaire au dessus du transistor supérieur est donc suffisant. L'épaisseur d'oxyde de type TeOS (précurseur : tetra ethoxy-silane) est diminuée, son épaisseur est suffisante pour bloquer la diffusion du phosphore.

Dans un deuxième temps, nous avons optimisé les épaisseurs de dépôt sur les substrats support et reporté ainsi que l'épaisseur retirée par CMP. Comme décrit sur la figure 1.34, l'épaisseur de diélectrique inter-niveau correspond à la somme de deux contributions : $T_{ILD, \text{reporté}}$ et $T_{ILD, \text{support}}$.

La contribution $T_{ILD, \text{support}}$ est la plus difficile à contrôler : d'un côté l'épaisseur du diélectrique déposée doit être suffisante pour permettre la planarisation (T_{DEPOSE} = typiquement 3 * la topographie) et l'épaisseur retirée doit être la plus faible possible car la dispersion augmente avec l'épaisseur retirée (~10% de l'épaisseur retirée). Un amincissement trop important peut conduire au découvrage des grilles à la surface du diélectrique planarisé. La figure 1.34 témoigne de la bonne optimisation de cette étape et démontre la réalisation d'un diélectrique fin (~100nm).

Pour mieux contrôler la contribution $T_{ILD, \text{reporté}}$, l'oxyde de type HDP a été remplacé par un oxyde thermique dont l'épaisseur est contrôlable au nm près. Cette étape nous permet de démontrer un diélectrique inter-niveau ultra fin (60nm) (cf. figure 1.35)

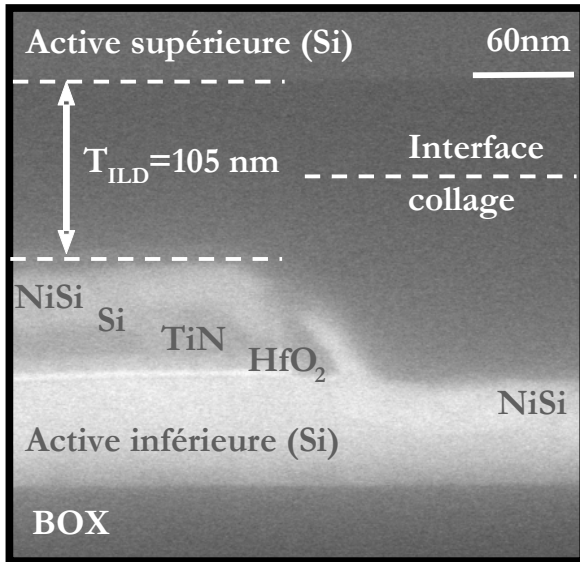


figure 1.34 : Photo MEB d'un report d'une zone active Si au dessus du transistor inférieur. Cas ILD fine~100nm

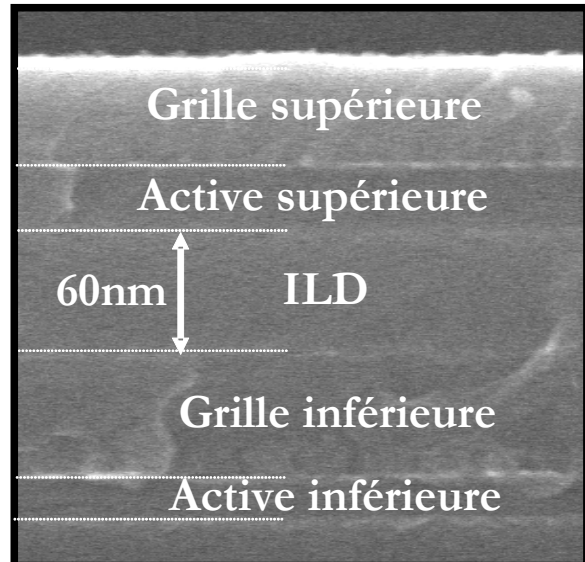


figure 1.35 : Photo MEB de l'empilement transistor supérieur au dessus du transistor inférieur. Cas ILD ultra fine=60nm

C.2.4 - Gravure des contacts à travers l'interface de collage

Pendant la gravure des contacts ou l'étape de stripping, des faiblesses dans l'interface peuvent apparaître conduisant à des courts-circuits entre les contacts voisins. La figure 1.36 montre un contact traversant l'interface de collage après stripping et dépôt de la barrière Ti/TiN. Aucune détérioration de l'interface de collage n'a été observée.

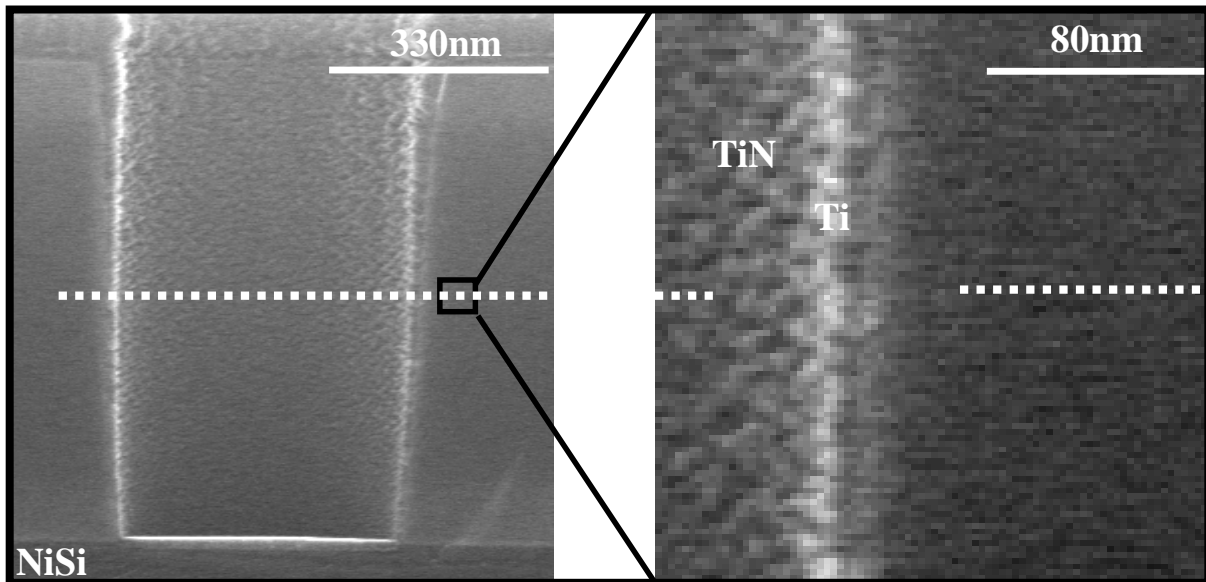


figure 1.36 : Observation MEB d'un contact traversant l'interface de collage après stripping et déposition de la barrière Ti/TiN. L'interface de collage ne présente pas de dégradation.

C. 3 -Comparaison avec l'état de l'art

Les résultats obtenus au Leti en termes de collage peuvent être mis en valeur par rapport à l'état de l'art de la technologie 3D séquentielle présenté dans le paragraphe C.1.2, par de nombreux aspects : Il s'agit du premier report complet sur plaque de 200 mm. Les épaisseurs de diélectrique inter-niveaux, ainsi que celles des films reportés correspondent aux plus petites démontrées actuellement. ($T_{Si} = 10 \text{ nm}$, $T_{Ge} = 50 \text{ nm}$, $T_{ILD} = 60 \text{ nm}$).

Les caractérisations du collage réalisées montrent une excellente qualité de collage.

Les travaux présentés ont été réalisés en collaboration avec le laboratoire « transfert de Film et Circuit » du LETI

Ce qu'il faut retenir

II-C : Réalisation de la zone active supérieure par collage moléculaire

- La réalisation de la zone active par collage moléculaire se démarque par rapport aux techniques de type « seed-window » par :
 - La qualité de la zone active qu'elle permet d'obtenir (qualité de l'actif supérieur dépendant uniquement de la qualité du substrat de départ,
 - Le contrôle de la dispersion d'épaisseur de film atteint,
 - La réalisation d'une zone active pleine plaque (pas de perte de densité due aux fenêtres d'ouvertures, défauts aux fronts de cristallisation..),
 - le choix de l'orientation cristalline des deux zones actives empilées.
- La topographie du diélectrique inter-niveau du côté substrat support est supprimée grâce à une étape de planarisation.
- Nous n'avons pas observé de défaut lié au fait que l'oxyde du côté du substrat support soit un oxyde déposé et non un oxyde thermique. La faible rugosité de cet oxyde déposé (0.16nm) rend l'étape de collage possible.
- La caractérisation du collage démontre que le report est d'excellente qualité (énergie de collage = 900 mJ.m^{-2}), sans défaut majeur sur l'ensemble de la plaque (caractérisation infrarouge et acoustique).
- Des reports de film de Ge et de Si d'épaisseur 50 et 10 nm respectivement ont été réalisés au dessus du niveau de transistor inférieur.
- L'épaisseur du diélectrique inter-niveau a été optimisée afin de minimiser le rapport des formes des contacts. (Gain en densité, utilisation du procédé de gravure classique pour la réalisation des contacts).
- L'étape de gravure des contacts ne fragilise pas l'interface de collage.

D - L'alignement

D. 1 -Modification des marques d'alignement

La méthode d'alignement dans une intégration séquentielle est la même que dans une intégration planaire. L'alignement des niveaux de masques se fait à l'aide de marques d'alignement, puis les motifs de la dernière couche sont définis par lithographie. Seul le nombre de couches à aligner est modifié (un niveau active et un niveau grille supplémentaire).

Deux types de lithographies sont utilisés au LETI, la lithographie optique en ultra violet profond (DUV) pour définir des motifs supérieurs à 300nm, et la lithographie électronique pour définir des motifs en dessous de cette limite.

Les marques d'alignement DUV sont constituées d'un réseau de traits de 8µm espacé de 16 µm et de 120nm de profondeur dans le silicium du substrat comme décrit dans la figure 1.37. La détection des marques se fait par diffraction dans le réseau de franges des marques. La topographie générée par les marques d'alignement, de l'ordre de 300 nm va augmenter la dispersion du diélectrique inter-niveau de 90 nm environ (10% de 3 * 300 nm). Pour limiter cette dispersion, les marques ont été modifiées et réalisées dans le silicium de la zone active d'épaisseur 10 nm (cf. figure 1.37(b)).

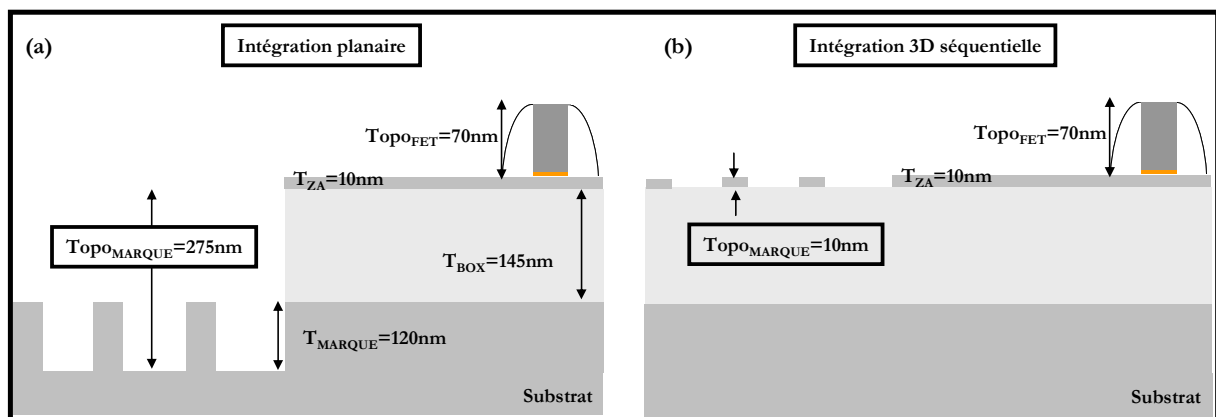


figure 1.37 : Description des marques DUV de l'intégration planaire (a) et modification pour l'intégration 3D séquentielle (b).

Malgré la réduction de topographie, l'équipement repère correctement les marques. La figure 1.38 récapitule les résultats d'alignement obtenus avec l'équipement de lithographie 248 nm. De plus nous observons que la précision d'alignement lorsque l'on aligne la zone active supérieure par rapport à la zone active inférieure n'est pas dégradée par rapport à l'alignement d'un niveau grille sur un niveau active.

	Ecart type (σ) Grille / active	Ecart type (σ) Active supérieure/ active inférieure
X	7	7
Y	10	7

figure 1.38 : Contrôle du désalignement dans le cas de la grille inférieure alignée sur l'active inférieure et dans le cas active supérieure sur active inférieure

Remarque : Dans le cas d'une lithographie e-beam, les marques d'alignement sont encore plus profondes (700nm dans le substrat). Pour supprimer cette topographie, un dépôt de 700 nm

d'HDP suivi d'une lithographie contremasque marque e-beam permet de combler les marques. Après le report de la zone active supérieure, ces marques seront régénérées et alignées sur les marques DUV. Les lots réalisés avec lithographie e-beam étant toujours en cours de fabrication, nous ne pouvons pas donner de résultats sur les précisions d'alignement obtenues.

D. 2 - Comparaison avec l'intégration 3D parallèle

La possibilité d'aligner les niveaux de masques nécessaires à la réalisation du transistor supérieur par rapport au niveau inférieur avec une telle précision ($\sigma \sim 10$ nm) est un atout majeur de la technologie 3D séquentielle par rapport à l'intégration 3D parallèle.

En effet, les meilleures performances d'alignement reportées dans la littérature parallèle varient entre $\sigma \sim 0.5 \mu\text{m}$ et $\sigma \sim 0.1 \mu\text{m}$ [Topol05], [Steen07] [Chen 07].

Les plus faibles performances d'alignement de l'intégration parallèle par rapport à l'intégration séquentielle s'expliquent principalement par l'enchaînement des étapes qui diffèrent entre intégration parallèle et intégration séquentielle. La figure 1.39 illustre l'enchaînement des étapes ayant un impact sur les performances d'alignement.

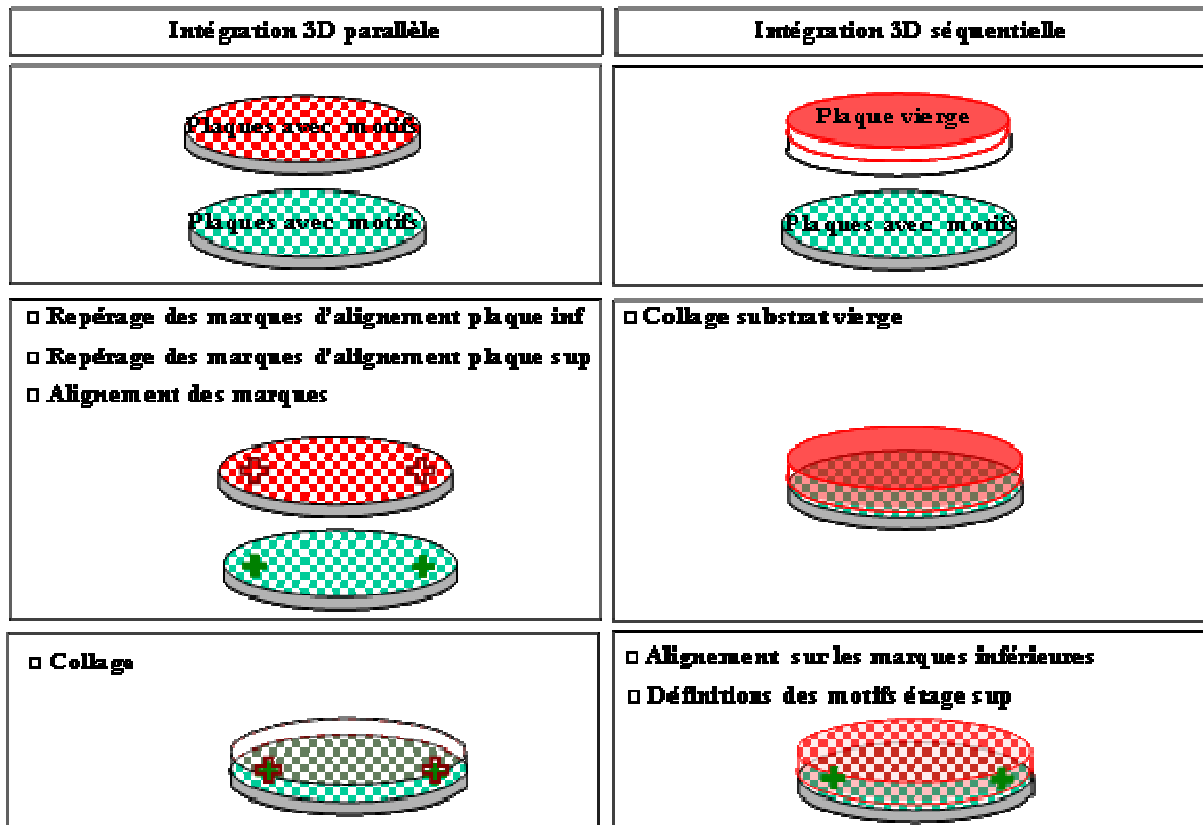


figure 1.39 : Description des étapes ayant un impact au niveau de l'alignement entre les niveaux inférieur et supérieur dans l'intégration parallèle et l'intégration séquentielle.

Dans le cas de l'intégration séquentielle, la plaque collée est vierge. En conséquence l'alignement des deux plaques au moment du collage n'a pas d'importance. C'est seulement après collage que l'on aligne le masque de la zone active sur les marques du substrat inférieur. La transparence du film semi-conducteur aux rayonnements ultra violets permet de réaliser l'alignement. L'alignement des masques du transistor supérieur est similaire à l'alignement d'un niveau grille ou d'un niveau contact dans une architecture planaire.

Dans l'intégration parallèle, le collage est réalisé après définition des motifs sur les deux plaques. En conséquence les deux plaques doivent être parfaitement alignées au moment du collage.

Actuellement, les équipements d'alignement développés spécifiquement pour l'intégration parallèle ne permettent pas de contrôler le désalignement avec précision au moment du collage des plaques. D'autre part, comme l'alignement se fait avant le collage, la courbure des plaques ainsi que la position des plaques entre le moment où l'alignement est réalisé et après réalisation du collage vont être modifiées. Pour obtenir plus d'informations sur les techniques d'alignement dans l'intégration parallèle, le lecteur pourra se reporter aux ouvrages [Tan08] et [Garrou08].

Cette limite en précision d'alignement implique une limite en termes de densité de contacts liant les plaques inférieures et supérieures. La figure 1.40 représente schématiquement un empilement 3D de type parallèle. Une simple application numérique nous permet d'estimer la densité maximale de connexions 3D en fonction de la précision d'alignement.

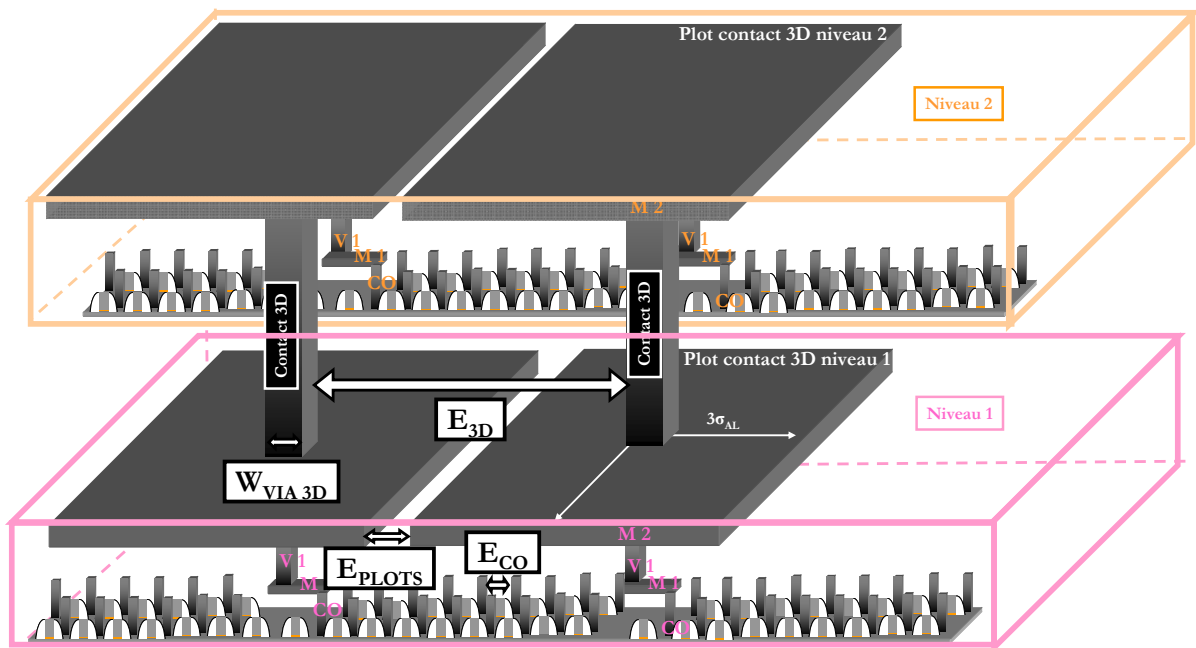


figure 1.40 : Estimation de la densité d'interconnexions 3D (liant plaque supérieure et inférieure) pour une intégration parallèle. Les faibles performances d'alignement entraînent une faible utilisation de la troisième dimension.

Des plots de contacts sont prévus de manière à venir contacter le contact 3D au niveau haut et au niveau bas. Ces plots sont généralement réalisés en Métal 1 ou en Métal 2. Pour tenir compte des désalignements éventuels, les plots ne sont pas dessinés avec les mêmes dimensions que les contacts 3D mais comportent des marges. Pour obtenir un rendement d'interconnexions fonctionnelles suffisant, ces marges sont dessinées au moins à 3σ .

Dans le cas d'une performance d'alignement à $1\sigma=0.5\mu\text{m}$, la taille des marges totales du plot sera alors de $3\mu\text{m}$. En négligeant la taille du via ($W_{\text{VIA}3\text{D}}$) (dépendant de la manière de réaliser l'intégration parallèle : via first, via last...) et la distance entre deux plots (E_{PLOTS}), on obtient une densité maximale de contacts 3D de 10^7 contacts/ cm^2 .

La meilleure performance d'alignement démontrée actuellement mène à $\sigma=0.1\mu\text{m}$ [Chen07], la densité maximale est alors de 2.10^8 vias/ cm^2 .

Dans le cas d'une intégration planaire, la densité de contacts est de 10^{10} contacts/ cm^2 (pour le nœud 45 nm). Cette densité prend bien en compte l'espacement minimal entre deux contacts et la taille du contact.

Ainsi le transfert d'information de la couche inférieure à la couche supérieure est réduit à 1 contact 3D (entre le deux niveaux empilés) tous les 100 contacts à l'intérieur d'un niveau empilé. En conséquence nous observons que la troisième dimension n'est que très peu utilisée.

Il est important de souligner que le rapport d'un contact 3D sur 100 contacts au sein d'une couche empilée est un scénario très optimiste car l'évaluation de la densité des vias 3D ne prend pas en compte la taille du via (W_{VIA3D}), ni de l'espacement entre le plots de connexions (E_{PLOTS}) (considérés nulle), alors que la valeur de la densité de contacts dans un niveau planaire pour le nœud 45 nm donnée, prends bien en compte l'espacement entre les contacts ainsi que la distance entre les contacts.

Dans l'intégration séquentielle, au contraire, les performances d'alignements ($\sigma=0.01\mu m$ avec un stepper 248 nm) permettent de réaliser des interconnexions 3D à l'échelle du transistor. Une cellule CMOS peut donc parfaitement être dessinée à cheval sur les deux niveaux.

D. 3 -Stratégie d'alignement dans l'intégration 3D séquentielle

Pour aligner les différents niveaux de masques, la référence peut être prise sur le tout premier niveau de marque généré avant même le début de procédé de fabrication (marques dans le substrat) ou sur les différents niveaux de marques générés au cours du procédé de fabrication.

Le choix de la référence d'alignement dépend alors de la criticité de l'alignement au niveau densité et de l'état des marques d'alignement. Par exemple dans le cas de l'alignement d'un niveau contact dans une intégration planaire, la référence est souvent prise préférentiellement sur le niveau grille, car l'espacement contact grille est plus critique que l'espacement contact sur active au niveau densité.

D'autre part, les marques d'alignement peuvent être dégradées au cours du procédé, les performances d'alignement seront alors meilleures si on aligne sur les marques les moins dégradées.

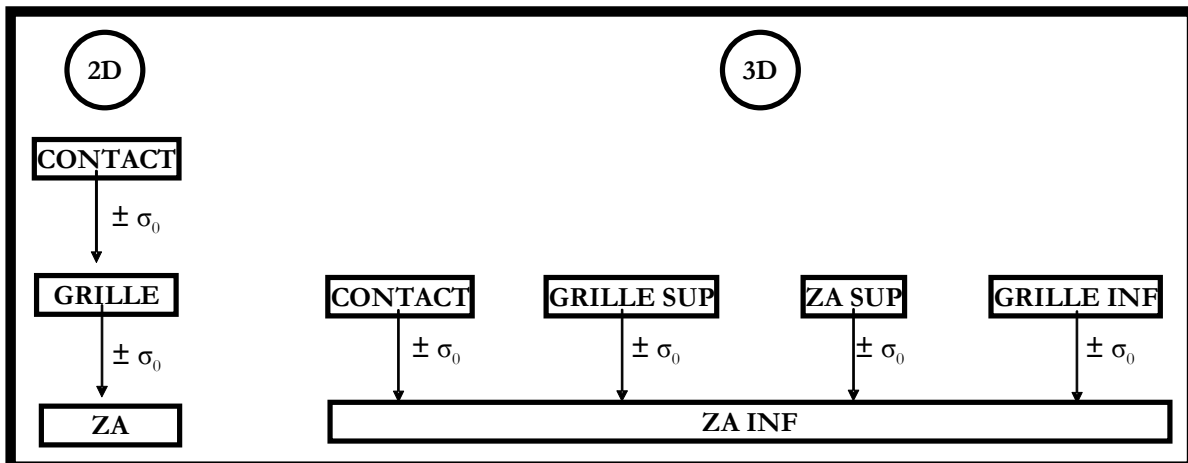


figure 1.41 : Description des stratégies d'alignement dans le cas d'une intégration planaire et de l'intégration 3D séquentielle réalisée au Leti.

Dans le cas d'une intégration planaire, une telle discussion sur le choix de la référence d'alignement, semble pertinente. Dans le cas de l'intégration 3D, le nombre de niveaux lithographiques augmentant, il semble plus intéressant de garder toujours la même référence au

cours du procédé (à la condition, que les marques d'alignement ne subissent pas de dégradation). C'est donc le choix de stratégie d'alignement que nous prendrons pour réaliser notre structure (cf. figure 1.41).

Ce qu'il faut retenir

II-D : L'alignement dans les structures 3D monolithique

- Pour limiter la topographie au moment du collage, la profondeur des marques d'alignement doit être modifiée par rapport au standard Leti.
- La réalisation des marques DUV dans le silicium de la zone active (10 nm) permet d'aligner les niveaux lithographiques nécessaires à la réalisation des transistors inférieurs et supérieurs avec un écart type σ de l'ordre de 10 nm pour un stepper 248nm (même performance que l'alignement d'un niveau grille sur un niveau active en planaire).
- Une telle performance d'alignement permet de réaliser des interconnexions 3D (liant niveau inférieur et supérieur) à l'échelle du transistor. En conséquence, les cellules CMOS peuvent être dessinées à cheval sur les deux niveaux.
- Dans le cas d'une intégration parallèle où la meilleure performance d'alignement reportée à $1\sigma = 0.1\mu\text{m}$. La densité d'interconnexion 3D maximale sera limitée à un maximum de 10^8 vias 3D par cm^2 . La densité de contact d'une technologie 45 nm étant de $10^{10}/\text{cm}^2$, les contacts 3D relieront donc des blocs logiques comportant environ 100 transistors. La troisième dimension n'est en conséquence que très peu utilisée par rapport à une intégration 3D séquentielle.
- Dans l'intégration réalisée au Leti, la référence d'alignement de tous les niveaux de masques de la partie Front end, contact inclus a été prise sur les marques d'alignement préliminaires réalisées avant le démarrage de procédé de fabrication. Cette référence unique permet de limiter le désalignement global entre tous les niveaux lithographiques.

E - Réalisation du transistor supérieur :

Procédés à basse température.

E. 1 -Répartition des budgets thermiques

Nous avons vu que l'intégration séquentielle impose une limitation du budget thermique du FET supérieur. La table suivante récapitule les budgets thermiques nécessaires à la réalisation d'un FET planaire et les modifications apportées afin de les réduire au maximum.

Transistor FDSOI planaire		Transistor FDSOI supérieur 3D séquentiel	
Recuit HfO ₂	600°C, 15 min	Recuit HfO ₂	515°C, 5 min
Dépôt Poly	580, 45 min	Dépôt Poly	515°, 40min
Espaceurs	725°C	Espaceurs	480°C
Epitaxie Si	750°C	Epitaxie Si _{0,7} Ge _{0,3}	650°C
Activation RTP	1050°C	Activation SPE Activation dans Ge	600°C

Table 2.1 : Récapitulation des budgets thermiques dans la cas d'un transistor FDSOI à budget thermique classique et dans le cas d'un transistor à bas budget thermique destiné à être réalisé sur le niveau supérieur de l'empilement 3D séquentiel.

Le recuit d'activation des dopants est clairement l'étape la plus coûteuse en budget thermique du procédé FDSOI standard. Pour le réduire, deux voies ont été suivies, la première est la réalisation du transistor supérieur en germanium (la température d'activation y est plus faible ~600°C), la deuxième est de remplacer le recuit thermique haute température par la recristallisation en phase solide (SPE : Solid Phase Epitaxy Regrowth) permettant d'activer les dopants dans le silicium à 600°C seulement.

Notons que le remplacement de l'oxyde thermique de silicium par des high-k déposés permet de prime abord de supprimer une contribution au budget thermique importante (~1000°C pour une oxydation thermique classique).

E. 2 -Le MOS Ge

Le FET GeOI est développé au Leti depuis 2004. Nous avons pu utiliser l'ensemble des développements préalablement réalisés.

Par rapport au transistor sur silicium, le transistor sur germanium est fabriqué à plus faible température, typiquement de l'ordre de 600°C, il est donc adapté à l'intégration 3D séquentielle. Le choix du diélectrique de grille pour les transistors germanium reste encore à l'heure actuelle une question ouverte. Actuellement, le Leti opte pour un empilement de grille dite « *Si-like* » : une fine épitaxie de silicium (1nm) est réalisée au dessus de la zone active suivie par le dépôt de l'empilement classique réalisé sur silicium (HfO₂/TiN/Poly-Si). Avant dépôt du high-k, le capping de silicium est partiellement oxydé, conduisant à un film de silicium de 0.5nm sur le substrat GeOI, l'interface oxyde-semiconducteur revient alors à l'interface bien maîtrisée Si/SiO₂ (cf. figure 1.42).

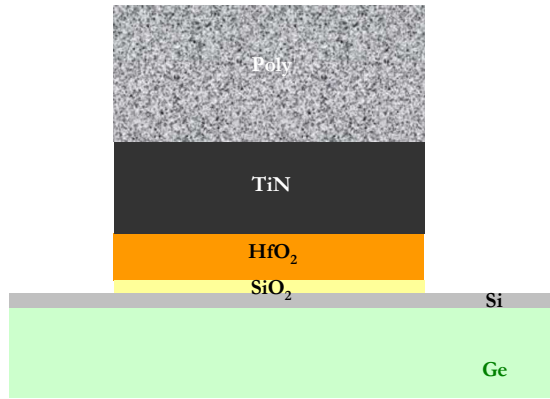


figure 1.42 : Description de l'empilement de grille de type « Si like » réalisée sur germanium.

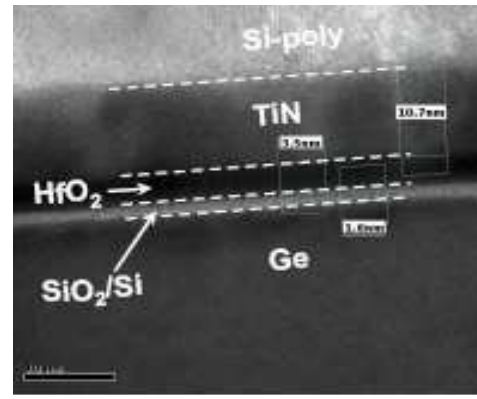


figure 1.43 : Photo MEB empilement de grille sur germanium.

A l'heure actuelle, les extractions de mobilité sur les dispositifs réalisés au Leti montrent un gain en mobilité pour les transistors pMOSFETs. Dans le cas du nMOS, de faibles mobilités (inférieures à celles obtenues sur silicium) sont obtenues. De plus la difficulté éprouvée à activer les dopants de type n conduit à des courants de conduction résultants encore plus dégradés. En conséquence nous nous concentrerons sur le transistor pMOS que nous cherchons à co-intégrer avec des transistors nMOS silicium afin de réaliser les cellules CMOS.

E. 3 -Le MOSFET Silicium à bas budget thermique

E.3.1 - Activation à basse température

E.3.1.a - Présentation de la technique

L'activation des dopants par recristallisation en phase solide a été originellement étudiée pour réaliser des jonctions ultrafines avec des niveaux d'activation élevés afin de limiter les effets canaux courts et obtenir des niveaux de courants importants dans des dispositifs de dimensions avancées.

Cette technique utilise la recristallisation à basse température ($\sim 600^{\circ}\text{C}$) du silicium amorphe permettant aussi l'activation des dopants à l'intérieur de la zone amorphisée. La figure 1.44 présente schématiquement les trois étapes de l'activation des dopants par recristallisation en phase solide. Dans un premier temps, l'arrangement cristallin est détruit par une implantation d'atomes lourds, c'est l'étape dite de « pré-amorphisation ». Ensuite les atomes dopants sont implantés. Puis le recuit à basse température va permettre de recristalliser le silicium tout en plaçant les dopants en site substitutionnel. Si l'espèce dopante est assez lourde, l'étape d'implantation peut suffire à auto-amorphiser le silicium.

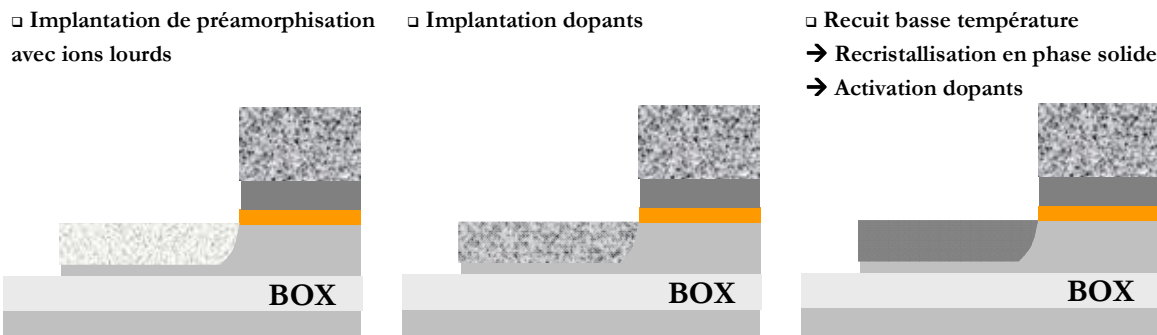


figure 1.44 : Procédé de recuit par recristallisation en phase solide.

E.3.1.b - Résultats SPER pleine plaque

Des tests pleines plaques sont tout d'abord réalisés afin de valider les conditions d'implantations calibrées à l'aide du logiciel de simulation CTRIM de SILVACO. Le film de silicium du substrat SOI mesure 30 nm. Les paramètres doivent être choisis pour ne pas amorphiser l'ensemble du film de manière à pouvoir le recristalliser à partir d'un germe monocristallin (15 nm et 20 nm amorphisé dans le cas LDD et HDD respectivement). Les conditions d'implantations et les types de dopants choisis sont récapitulés dans la table 2.2.

	N		P	
	Type LDD	Type HDD	Type LDD	Type HDD
Agent de pré-amorphisation			Ge	Ge
Dose (cm ⁻²)			5.10 ¹⁴	5.10 ¹⁴
Energie implantation			9	13
Dopant	As	As	B	B
Dose (cm ⁻²)	8.10 ¹⁴	2.10 ¹⁵	8.10 ¹⁴	2.10 ¹⁵
Energie implantation (keV)	8	10	3	3

Table 2.2 : Récapitulation des conditions d'implantation SPE sur SOI pleine plaque.

L'arsenic est auto-amorphisant, il n'est pas nécessaire de réaliser une étape de pré-amorphisation. Dans le cas du bore, le dopage des accès nécessite une pré-amorphisation.

Après implantation, les plaques sont soumises au recuit de recristallisation (600°C, 2min). La figure 1.45 présente la mesure de la résistance de la couche en fonction de la profondeur de la jonction. Ces résultats sont comparés aux résultats de la littérature avec recuit thermique rapide (RTP : Rapid Thermal Processing) et avec des techniques de recuit avancées (SPE, Laser, Flash). Pour les mesures sur substrat massif, la profondeur de la jonction est prise pour une concentration de dopants de 5.10¹⁸ .cm⁻³, et dans le cas du SOI, égale à l'épaisseur de silicium de la zone active. Les mesures de résistances sont réalisées par mesures 4 pointes.

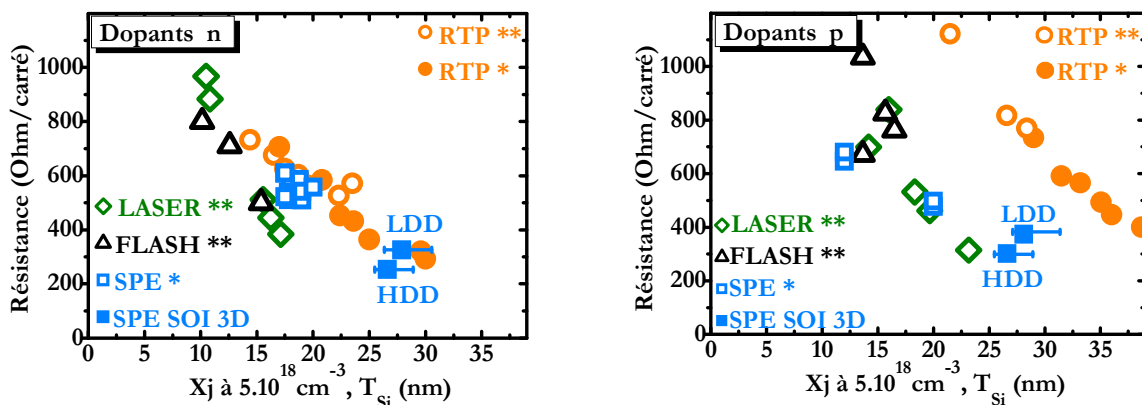


figure 1.45 : Mesure des résistances des couches dopées par SPE et comparaison avec les résultats obtenus par recuit RTP et recuits alternatifs sur substrat massif. *[Pouydebasque05], **=[Sugoro04]

Ces figures correspondent aux figures de mérite classique pour des jonctions. L'objectif étant de se placer dans la région ayant la profondeur de jonction la plus fine et la résistance de couche la plus faible. Nous observons que dans le cas des dopants n, l'activation SPE sur film mince donne des résultats similaires aux autres techniques de recuit. Dans le cas des dopants p, nos résultats sont équivalents à ceux obtenus avec les techniques de recuits avancés (FLASH, Laser) et bien meilleurs que ceux obtenus avec un recuit haute température classique.

Dans le cadre d'une intégration transistor, la première série d'espaceurs doit être supprimée. En effet, à la température du recuit de cristallisation, la diffusion des dopants est quasi nulle, avec un espaceur, il n'existe pas de zone de recouvrement grille/accès. Un recouvrement trop faible limiterait le courant I_{ON} .

E.3.2 - Surélévation des accès à basse température

Les surélévations des sources et drains ont pour rôle de permettre la fabrication de transistors sur films fins, il s'agit donc d'une brique technologique indispensable à la réalisation de transistors FDSOI avec des dimensions avancées (sub-45nm).

L'épaississement des accès permet (i) de conserver une zone non amorphisée après l'implantation des dopants afin de permettre la recristallisation des sources et drain et (ii) de disposer d'une épaisseur de film suffisante pour la siliciuration.

L'épitaxie de Si classiquement réalisée à 750°C en technologie planaire a été remplacée par une épitaxie de $Si_{0.7}Ge_{0.3}$ à 650°C. Nous observons sur la figure 1.46 que passer de l'option silicium pur à l'épitaxie de SiGe avec un pourcentage de Ge de l'ordre de 30% permet d'augmenter significativement la vitesse de croissance (à température donnée). On dispose ainsi de temps de croissance raisonnable à 650°C pour la surélévation des S/Ds [Hartmann08a,b]. De plus, la réduction du budget thermique permet de limiter les phénomènes d'agglomération apparaissant sur les films minces sur isolant.

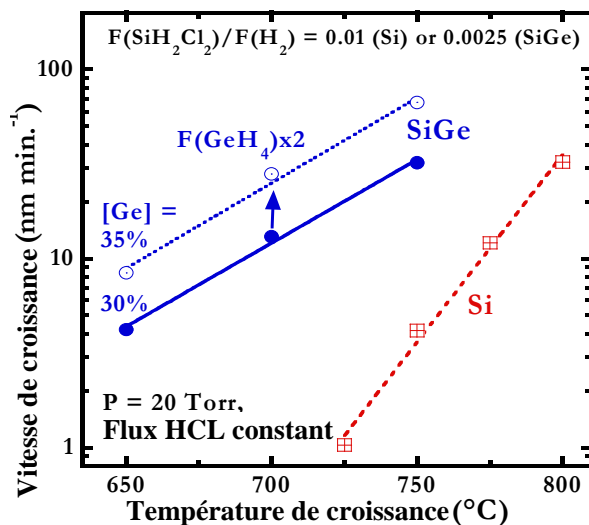


figure 1.46: Comparaison des vitesses de croissance en fonction de la température.

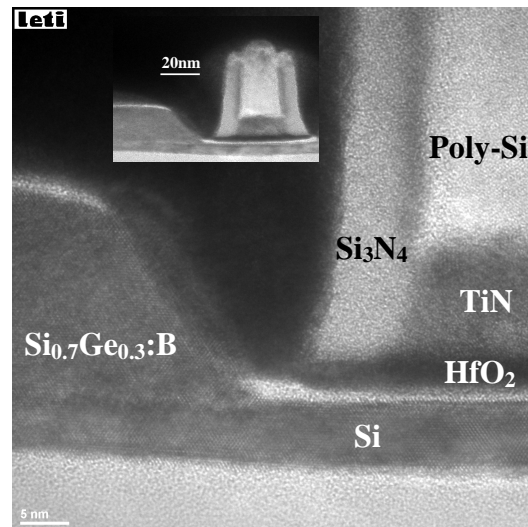


figure 1.47 : Observation MEB d'une surélévation d'accès, épitaxie sélective de SiGe (Ge:30%) à 650°C sur film de silicium de 8nm.

Cette étape est indispensable pour réaliser des transistors FDSOI. Elle impose que le FET inférieur soit stable jusqu'à 650°C, alors qu'à présent, le budget thermique nécessaire pour réaliser le FET supérieur était de 600°C.

Ce qu'il faut retenir

II-D : Réalisation du FET supérieur à basse température

- L'intégration 3D séquentielle requiert le développement d'un FET réalisé à faible budget thermique afin de ne pas dégrader le FET inférieur. Le transistor inférieur avec sa siliciuration développé peut théoriquement avoir des performances inchangées si le BT supérieur reste en dessous de 650°C. Cependant nous chercherons à supprimer toute contribution au BT non indispensable à la fabrication du FET supérieur même si elle est inférieure à cette valeur limite.
- Il est indispensable de trouver une alternative au recuit d'activation des dopants à 1050°C.
- Pour répondre à cette problématique, la réalisation d'un FET en germanium au premier étage est envisagée. La température d'activation des dopants dans le germanium est aux alentours de 600°C.
- Pour réaliser l'activation des dopants dans le silicium, nous proposons d'utiliser la technique de recristallisation en phase solide (SPER). Après amorphisation d'une partie du film semiconducteur et l'implantation des dopants, un recuit à faible température permet la recristallisation du film tout en activant les dopants.
- La figure de mérite (résistance de la couche en fonction de la profondeur de jonction) des activations basse température réalisées confirme la pertinence de cette méthode pour l'activation des dopants du transistor supérieur. Les résultats obtenus sont dans la lignée des résultats obtenus avec les techniques de dopage les plus avancées (FLASH, LASER) dans le cas des dopants de type n et p.
- La réalisation des surélévations des sources et drains, indispensable pour le dopage des films minces peut être rendue compatible avec l'intégration 3D séquentielle ($T_B < 650^\circ\text{C}$) en réalisant une épitaxie de $\text{Si}_{0.7}\text{Ge}_{0.3}$ au lieu d'une épitaxie Si.

F - Interconnexions 3D

F. 1 - Contact chevauchant

Pour la réalisation des contacts, nous disposons d'un seul masque de contact pour graver les

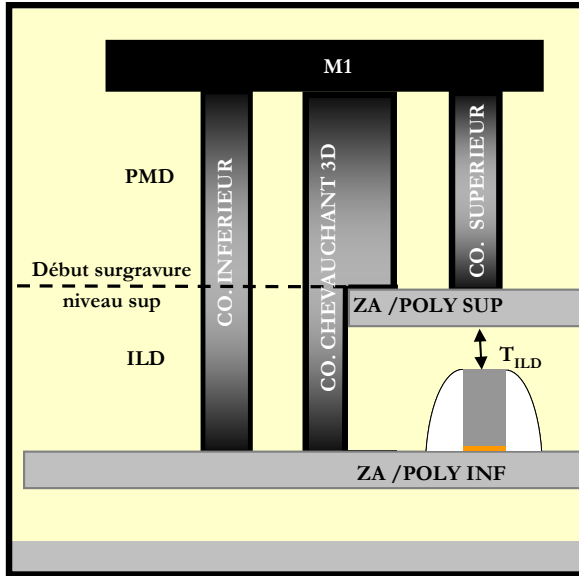


figure 1.48 : Description des contacts réalisés en un seul niveau lithographique.

contacts débouchant sur le niveau supérieur, ceux débouchant sur le niveau inférieur et ceux qui relient le niveau supérieur avec le niveau inférieur (cf. figure 1.38). Pour réaliser ces contacts simultanément, il est nécessaire de disposer d'une gravure hautement sélective par rapport au niveau supérieur. Afin de pouvoir réaliser des interconnexions dans les deux options choisies (option Ge ou Si pour le transistor supérieur), la chimie de gravure devra être sélective à la fois sur le siliciure de nickel et le germaniure de nickel (la grille est toujours en poly silicium).

Pour vérifier la bonne tenue du siliciure (ou du germaniure) lors de la surgravure permettant d'atteindre le niveau inférieur, nous avons imposé des temps de surgravure permettant de graver 300 nm

d'isolation inter-niveau. Les résultats obtenus sont présentés en figure 1.49.

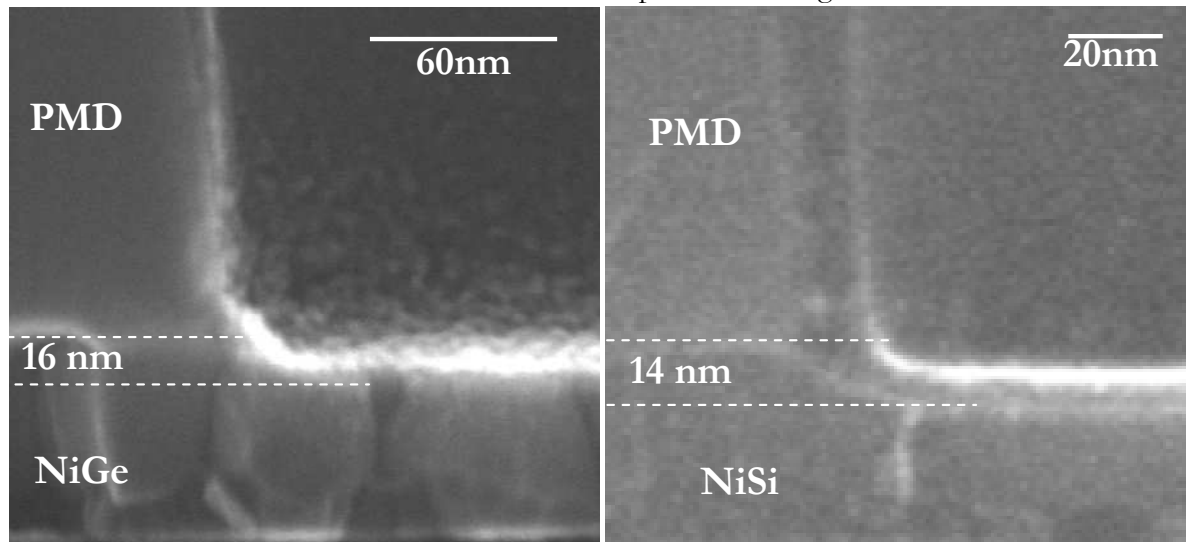


figure 1.49 : Observation MEB des contacts atterrissant sur les niveaux supérieurs avec une surgravure de plus de 300nm.

Rappelons que cette épaisseur correspond à une borne supérieure puisque nous avons réduit l'épaisseur de diélectrique inter-niveau à 100nm ramenant l'épaisseur supplémentaire à graver autour de 200nm (en prenant en compte l'épaisseur d'active supérieure et la hauteur du transistor inférieur). Nous observons que la consommation de siliciure ou de germaniure permettant à la gravure de déboucher sur le niveau inférieur situé 300 nm au dessous est de l'ordre de 15nm. Il

est donc possible de contacter le niveau inférieur sans traverser le niveau supérieur. L'arrêt dans le siliciure (germaniure) permet de garantir de faible résistance de contact.

Les figures suivantes témoignent de la validité de cette approche pour réaliser les contacts avec une seule lithographie.

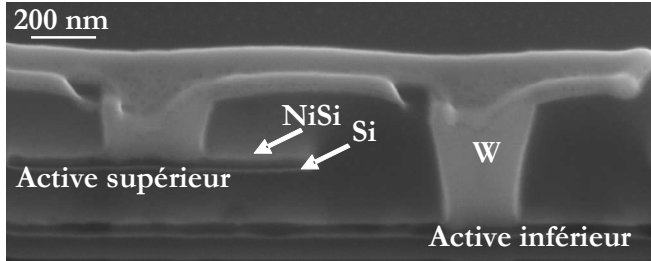


figure 1.50 : Observation de contacts atterrissant sur l'étage supérieur et inférieur avec un seul niveau lithographique (vue en coupe).

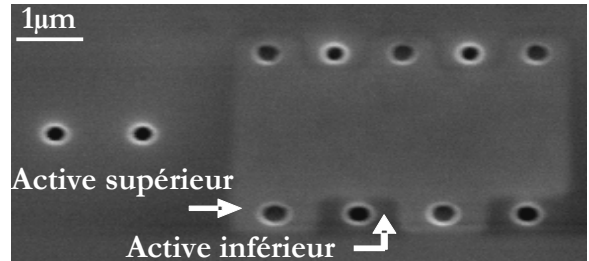


figure 1.51 : Observation de contacts atterrissant sur l'étage supérieur et inférieur avec un seul niveau lithographique (vue du dessus).

F. 2 -Contact traversant

On appelle contact « traversant », un contact qui lie le niveau supérieur au niveau inférieur en traversant la couche supérieure (cf. figure 1.52).

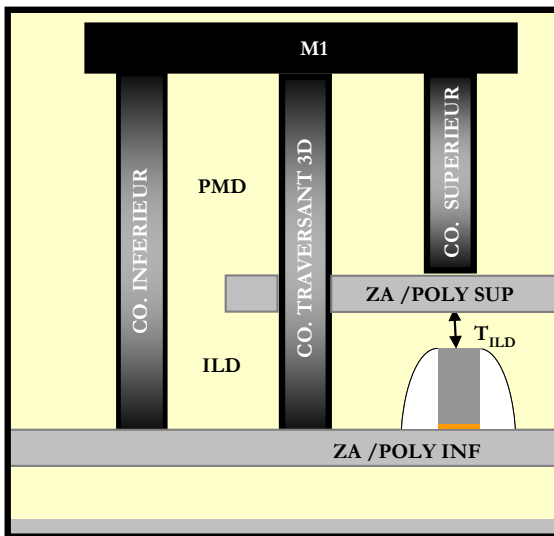


figure 1.52: Description du contact traversant

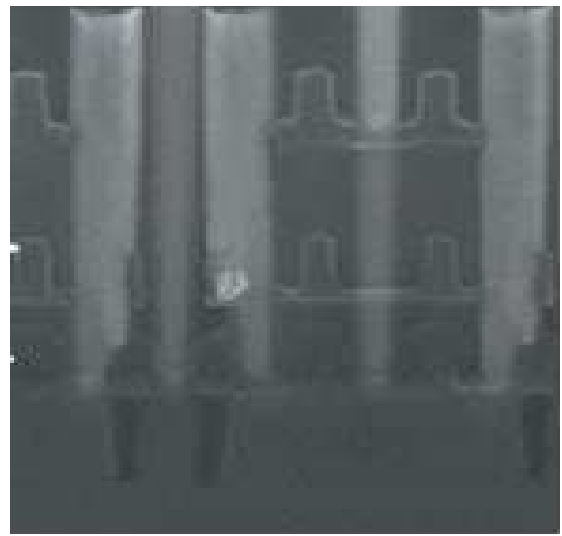


figure 1.53 : Vue en coupe des contacts traversants de l'architecture démontrée dans [Jung05]

Dans ce cas, au moins deux étapes de lithographie seront alors nécessaires pour réaliser tous les contacts : une première lithographie définissant les contacts atterrissant uniquement sur le niveau supérieur ou le niveau inférieur (suivie d'une gravure sélective) et une deuxième lithographie définissant les contacts connectant les deux niveaux en traversant la couche supérieure (gravure non sélective).

Ce type de contact n'a pas pu être démontré au cours de cette thèse. Pour la suite des études, un autre masque a été dessiné au cours de la thèse, avec des structures comportant des contacts traversants.

Dans la littérature, les deux types de contacts ont été présentés. L'exemple de contact traversant le plus abouti a été présenté par Samsung où la gravure traverse deux zones actives avant d'atterrir sur la zone active inférieure (cf. figure 1.53).

En termes de densité d'intégration, le contact traversant est nettement plus intéressant que le contact chevauchant. Dans le contact traversant, la surface de contact est latérale et permet de diminuer de moitié la taille du contact 3D (cf chapitre III partie A). L'impact sur la résistance série sur ce type de contact est une problématique complexe et n'a pas été étudiée au cours de cette thèse.

F. 3 -Contact interne

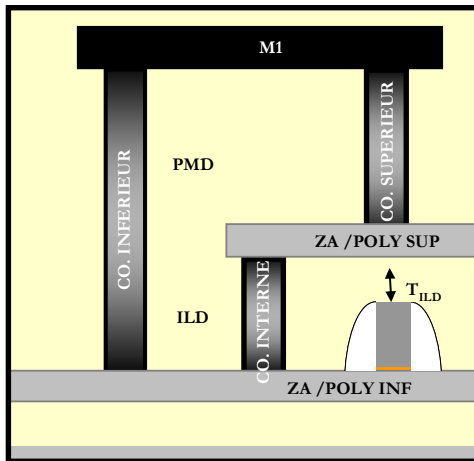


figure 1.54 : Description du contact interne

Nous appelons contact « interne », un contact qui ne remonte pas jusqu'au niveau métal 1, il est confiné entre le niveau supérieur et inférieur (cf. figure 1.54). L'espace situé au dessus peut alors être librement utilisé par du M1. Cette interconnexion locale entre deux niveaux facilite la connexion entre le niveau n et le niveau n+1 dans une intégration à plus de deux niveaux.

L'aspect réalisation technologique de ce contact n'a pas été abordé dans cette thèse, cependant il nous apparaît comme une étude importante à réaliser par la suite. Nous verrons dans le chapitre III consacré aux gains apportés par la 3D que ce type de contact permet un gain en densité supplémentaire.

Ce qu'il faut retenir

II-E : Interconnexions 3D

- Nous avons démontré la réalisation de l'ensemble des contacts permettant de fabriquer des cellules 3D en une seule lithographie.
- Cependant en termes de densité, l'option contact chevauchant n'est pas la plus intéressante. En conséquence, il sera nécessaire de développer une technologie de contact de type traversant.
- Une interconnexion de type local, reliant uniquement le niveau inférieur et supérieur et ne remontant pas jusqu'au niveau métal 1 semble prometteuse, surtout dans des intégrations avec plus de deux niveaux de transistors empilés.

G - Conclusion développement briques 3D

L'ensemble des défis technologiques majeurs pour réaliser une intégration 3D séquentielle a été relevé et nous a permis de réaliser une intégration complète dont les résultats électriques sont présentés dans la chapitre II.

La table suivante permet de récapituler les options technologiques démontrées et de les comparer aux démonstrations récentes de l'état de l'art 3D séquentiel. Nous ne sélectionnons que les publications où des FET supérieurs et inférieurs sont réalisés de manière séquentielle.

	LETI [Batude09]	SAMSUNG [Jung05]	SAMSUNG [Jung07]	STANFORD [Feng06]	CORNELL [Xue03]	TAIWAN [Yu05]
FET Inférieur						
Siliciuration niveau inférieur	Oui SALICIDE Ni	Partielle fonds de contacts Co	Partielle fonds de contacts Co	non	non	non
Active supérieure						
épaisseur diélectrique inter-niveau	60-100nm	~60nm	~100nm	~500nm	~500nm	~200nm
Technique de réalisation	Collage moléculaire Retrait substrat	recristallisation à 650°C	Selective epitaxy growth SEG (SW)	Rapid Melt growth RMG Ge (SW)	Collage exfoliation	Collage exfoliation
Epaisseur active	~10nm (Si) ~50nm (Ge)	~20nm (Si)	~20nm (Si)	250nm (Ge)	150nm (Si)	1.5μ (Ge)
FET supérieur						
oxyde de grille	HfO ₂ ALD	SiO ₂ oxydation plasma	SiO ₂ oxydation plasma	Nitruration à 600°C	?	IrO ₂ /LaAlO ₃
surélévation accès	oui	non	non	non	non	non
Type d'activation	SPER (Si) Thermique (Ge)	Spike (Si)	Spike (Si)	Thermique (Ge)	Thermique (Si)	Thermique Ge
BT max	≤650°C	≤750°C	≤650°C	≤945°C	≤800°C	≤500°C
Lg (niveau supérieur)	0.5μm ² 50nm	~65nm	~80nm	~1μm	~1μm	10μ
Contacts	Chevauchants	Traversant	Traversant	Chevauchants	Aucun contact 3D	Aucun contact 3D

Les choix technologiques ainsi que les développements réalisés placent l'intégration proposée dans les schémas d'intégration les plus aboutis. Ces choix technologiques sont en adéquation avec des applications haute densité (sub-45nm) :

- L'architecture FDSOI permet un meilleur contrôle des effets canaux courts. Cette architecture sur le niveau supérieur est rendue possible par le développement du report de couche par collage moléculaire au dessus du transistor inférieur permettant de contrôler avec précision l'épaisseur de la zone active supérieure, ainsi que par le développement de surélévation de S/D à 650°C.

² Une longueur de grille minimale de 0.5μm a été démontrée pour les transistors supérieurs. Les résultats pour des longueurs de grilles plus avancées (jusqu'à 50nm) sont obtenus avec un procédé basse température mais ne sont pas empilés au dessus d'un niveau de transistor inférieur)

- La siliciuration du niveau inférieur permet d'obtenir les niveaux de courants de conduction requis pour des nœuds avancés.
- L'utilisation du collage moléculaire plutôt que des techniques SW permet de ne pas perdre de place avec les zones d'ouverture et d'obtenir une qualité cristalline de zone active compatible avec une exigence industrielle.
- De plus, l'utilisation du collage moléculaire permet de choisir des orientations différentes pour les deux niveaux de transistors.

Nous avons démontré la compatibilité de cette technique avec la réalisation de film très minces (10nm) et des épaisseurs de diélectrique inter-niveau faibles (60nm).

CHAPITRE II

CARACTERISATION ELECTRIQUE DE L'INTEGRATION COMPLETE

SOMMAIRE DU CHAPITRE II :
CARACTERISATION ELECTRIQUE DE L'INTEGRATION COMPLETE

A -	INTEGRATIONS REALISEES	58
B -	ETUDE DES CARACTERISTIQUES DU FET INFERIEUR.....	60
B. 1 -	Stabilité du siliciure inférieur.....	60
B. 2 -	La caractérisation de l'empilement de grille.....	63
C -	ETUDE DES CARACTERISTIQUES DES FETs SUPERIEURS	67
C. 1 -	pMOSFETs GeOI	67
C.1.1 -	<i>Résultats obtenus</i>	67
C.1.2 -	<i>Les voies d'améliorations possibles</i>	68
C.1.3 -	<i>Discussion préliminaire sur l'option Ge</i>	68
C. 2 -	MOSFETs Silicium à bas budget thermique	70
C.2.1 -	<i>Description des variantes du lot d'étude</i>	70
C.2.2 -	<i>Analyse des tendances : tests paramétriques</i>	72
C.2.3 -	<i>Présentation des caractéristiques unitaires</i>	75
C.2.4 -	<i>Caractérisation des résistances d'accès</i>	76
C.2.5 -	<i>Caractérisation de l'empilement de grille</i>	77
D -	DEMONSTRATION DE CELLULES EN 3D	80
D. 1 -	L'inverseur	80
D. 2 -	La SRAM 6T	82

A - Intégrations réalisées

La Figure 2.1 décrit les deux types d'intégrations réalisées à savoir une première option orientée pour des applications « basse consommation » avec un empilement de deux zones actives silicium et une deuxième option orientée pour des applications « haute performance » avec une zone GeOI au dessus d'une zone active inférieure en SOI. Dans l'option co-intégrée GeOI/SOI, seuls des transistors pMOS GeOI ont été fabriqués à l'exclusion des transistors nMOS GeOI.

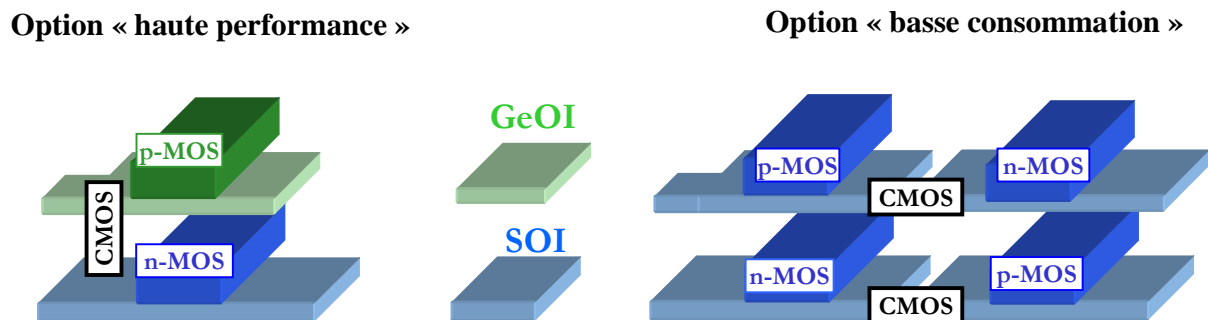


Figure 2.1 : Description des démonstrations réalisées. Dans l'option haute performance, les cellules CMOS sont nécessairement dessinées à cheval sur les deux niveaux empilés. Dans l'option basse consommation, où les deux actives sont en silicium, les cellules CMOS peuvent être dessinées sur les deux niveaux.

La Figure 2.2 récapitule l'enchaînement des principales étapes nécessaires à la réalisation des dispositifs 3D caractérisés dans cette partie.

Les transistors du niveau inférieur ont été fabriqués avec un procédé FDSOI classique (décrit dans le chapitre I partie A) à l'exception de la brique siliciuration. En l'occurrence, l'activation des dopants est donc réalisée par activation thermique à haute température (1050°C). Les transistors sont réalisés sur des films de silicium de 20nm environ avec un empilement de grille $\text{HfO}_2(5\text{nm})/\text{TiN}(10\text{nm})/\text{Poly-Si}(50\text{nm})$. Pour stabiliser le siliciure en vue du budget thermique du FET supérieur, une étape d'implantation de fluor a été effectuée après la formation du siliciure de nickel.

Les plaques de SOI ou de GeOI sont ensuite reportées sur le diélectrique inter-niveau planarisé et aminci. Plusieurs variantes d'épaisseurs de diélectriques inter-niveau ont été démontrées : une variante à diélectrique inter-niveau fin ($T_{\text{ILD}}=100\text{nm}$), et une variante à diélectrique inter-niveau ultra fin ($T_{\text{ILD}}=60\text{nm}$). Les épaisseurs des films sont aux alentours de 20 nm et 50 nm pour les zones actives de Si et Ge respectivement.

Les transistors supérieurs sont ensuite fabriqués avec un budget thermique limité à 600°C, l'empilement de grille restant identique à celui des transistors inférieurs.

Les contacts sont définis à partir d'un seul niveau lithographique, en utilisant un procédé de gravure sélectif. Cette intégration est démontrée sur des plaques de diamètre 200mm.

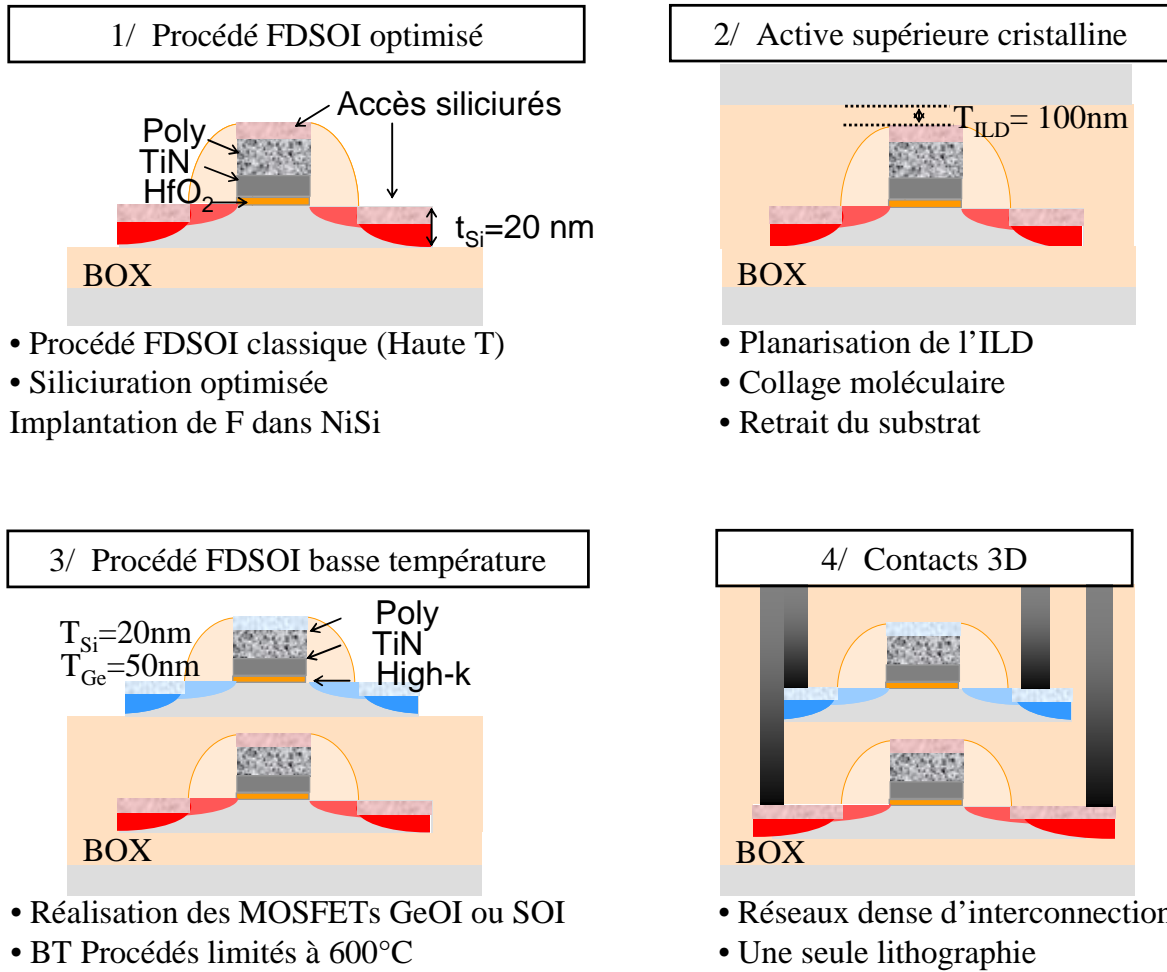


Figure 2.2 : Récapitulation des grandes étapes de fabrication des démonstrateurs réalisés de la technologie 3D séquentielle.

Les vues en coupe et du dessus des dispositifs réalisés sont présentées dans les figures suivantes.

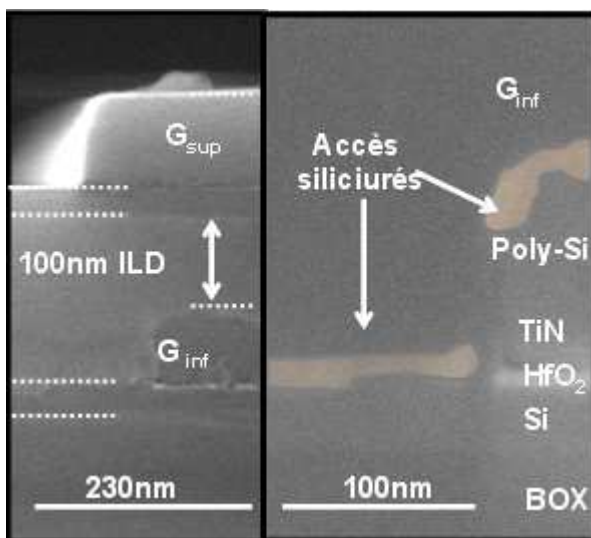


Figure 2.3: Photographies MEB d'une coupe d'une structure 3D.

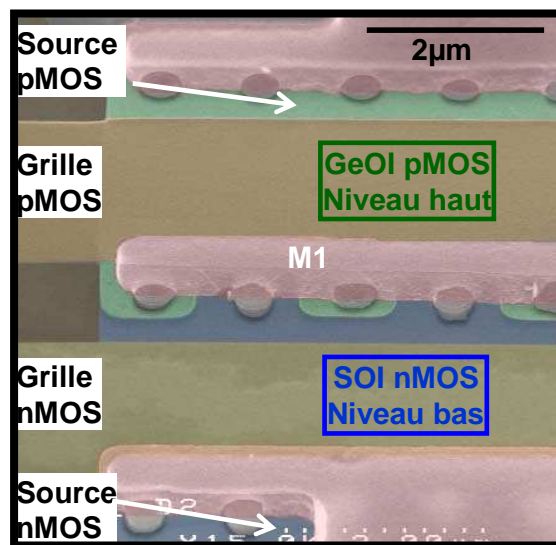


Figure 2.4 : Photographies MEB vue de dessus tiltée d'une structure 3D.

B - Etude des caractéristiques du FET inférieur

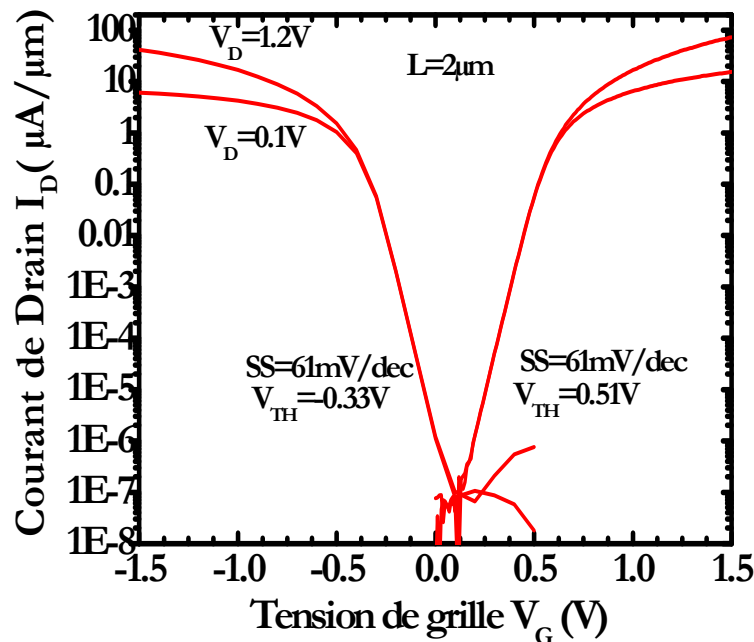


Figure 2.5 : Caractéristiques I_D - V_G du transistor inférieur après réalisation du FET supérieur.

les différentes briques technologiques susceptibles d'être dégradées par le budget thermique du transistor supérieur : à savoir les siliciures, et l'empilement de grille.

B. 1 -Stabilité du siliciure inférieur

Dans cette partie, nous allons caractériser la stabilité du siliciure du niveau inférieur dans l'intégration complète. Nous rappelons que l'étude présentée dans la partie I a été effectuée sur des plaques entières, sans motifs. Dans le cas d'un substrat avec motifs (mésas de zone active sur BOX), la stabilité morphologique du film peut se révéler différente.

D'autre part, dans l'étude réalisée sur substrat pleine plaque, le siliciure n'est pas encapsulé par une couche d'isolant. Dans le cas des transistors inférieurs, les transistors sont encapsulés par le diélectrique inter-niveau. Cette encapsulation peut entraîner des modifications dans le comportement du siliciure.

Enfin, le budget thermique appliqué réellement au transistor inférieur est plus complexe qu'un seul recuit à une température fixée. La récapitulation des BT du FET supérieur est présentée dans le tableau 3.1.

Nous rappelons que les résultats sur le lot électrique correspondent à un siliciure de NiSi avec implantation de fluor. Dans l'étude menée sur les substrats « pleines plaques », la variante avec implantation de fluor seul ne correspond pas à la meilleure option pour stabiliser le siliciure, mais nous ne disposons pas de tous les éléments de réponses apportés par l'étude pleine plaque au moment de la réalisation du lot électrique complet.

Les caractéristiques I_D - V_G des transistors inférieurs de type n et de type p sont présentées dans la figure ci-contre. Ces caractéristiques sont mesurées après la fabrication du transistor supérieur.

L'allure générale des caractéristiques montre que les transistors sont bien fonctionnels. Les pentes sous le seuil de 61mV/décade extraites indiquent que les qualités d'interface oxyde-semiconducteur sont bien contrôlées. Les rapports I_{ON}/I_{OFF} obtenus sont cohérents avec ceux attendus pour des FETs silicium.

Dans la suite de cette partie, nous analyserons plus en détails

La Figure 2.6 présente la valeur moyenne de la résistance de la zone active inférieure après la réalisation du FET supérieur. Elle est comparée aux valeurs mesurées sur substrats «splines plaques» et soumis à un recuit à 600°C en fonction du temps. Les mesures de résistance de l'active sont réalisées sur des motifs Vanderpaw.

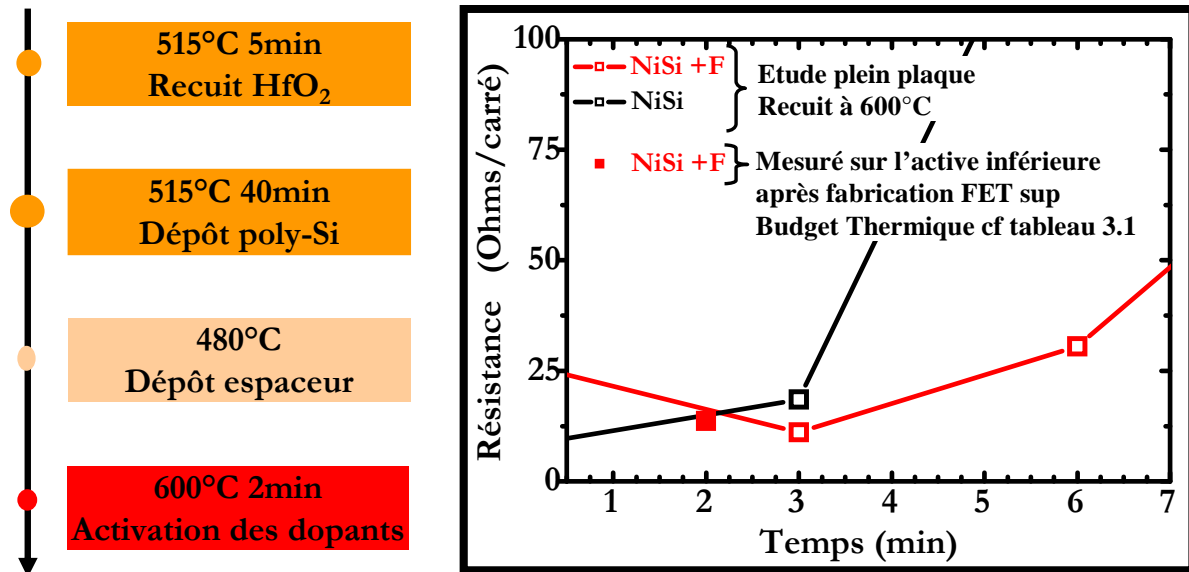


Tableau 3.1: Récapitulation des budgets thermiques principaux utilisés pour la fabrication du FET supérieur.

Figure 2.6 : Comparaison de la stabilisation du siliciure entre étude pleine plaque soumis à un recuit isotherme et mésas de zone actives encapsulés dans l'ILD soumis aux BT du FET supérieur récapitulés dans le tableau ci-contre.

Pour placer sur ce graphique le point de mesure expérimental sur zone active inférieure après réalisation du FET supérieur, nous n'avons pris en compte que le plus fort budget thermique à 600°C 2min. Nous observons alors que les valeurs pleines plaque et sur transistor inférieur sont très proches. L'encapsulation et la texturisation de l'active n'ont pas une influence forte. D'autre part, il semble que les budgets thermiques plus faibles (dépôt poly-Si, dépôt espaceur, recuit HfO₂) aient une influence moins importante que le budget thermique à 600°C.

Les Figure 2.7 et Figure 2.8 permettent de mettre en évidence l'évolution de la résistance du siliciure avant et après réalisation du FET supérieur. Elles sont aussi comparées à celles de la zone active supérieure.

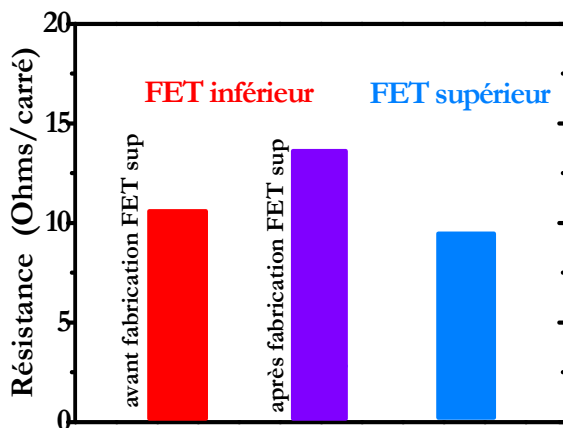


Figure 2.7: Moyenne des mesures des résistances d'actives inférieures (avec ou sans FET haut) et actives supérieures.

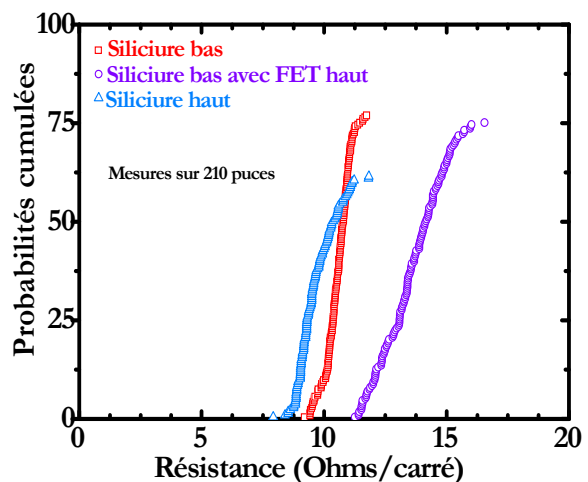


Figure 2.8 : Probabilités cumulées des résistances d'actives inférieures (avec ou sans FET haut) et actives supérieures.

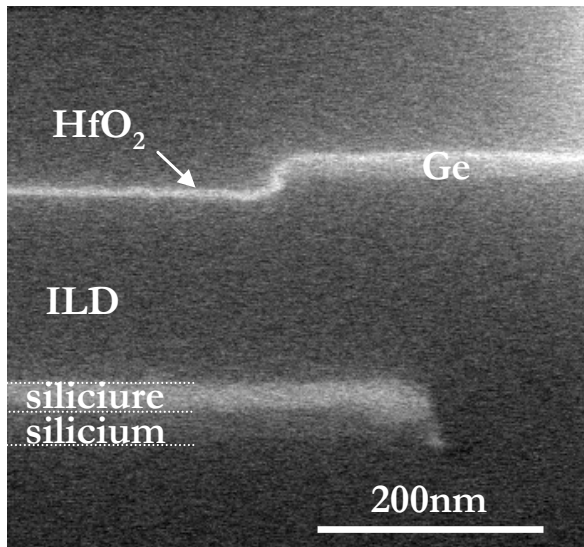


Figure 2.9 : Observation d'une coupe par MEB d'un transistor CMOS inférieur après fabrication du transistor supérieur

Nous observons que les valeurs de résistance des zones actives avec ou sans réalisation du FET supérieur sont très proches. Notons cependant qu'une dispersion plus importante des valeurs apparaît dans la variante avec transistor supérieur (cf. Figure 2.8). Néanmoins, la proximité de valeurs des résistances des zones actives inférieures avec ou sans transistor supérieur indique que la couche de siliciure n'est pas agglomérée, car dans ce cas, nous aurions mesuré des valeurs de résistance de silicium dopé (de l'ordre de la centaine d'Ohms/carré). La continuité du film est aussi confirmée par l'observation MEB présentée en Figure 2.9.

Ce résultat permet de démarquer l'étude réalisée au Leti par rapport à l'état de l'art. En effet, à notre connaissance, il s'agit de

la première démonstration de structure 3D séquentielle avec des accès inférieurs siliciurés avec le procédé auto-aligné SALICIDE. Cette étape de siliciuration est rendue possible grâce au développement conjoint d'un siliciure à stabilité optimisée et à la diminution du budget thermique du FET supérieur (budget thermique maximal : 600°C).

Au vue des développements présentés dans la partie I, le siliciure avec incorporation de platine associé avec l'implantation de tungstène et de fluor s'est révélé morphologiquement stable pour un budget thermique de 650° jusqu'à 40 minutes de recuit. Ces résultats pleine plaque nous permettent de penser que le siliciure inférieur pourrait supporter l'étape de surélévation des accès (épitaxie à 650°C). Il s'agit d'un premier point à étudier pour statuer sur la possibilité de réaliser des sources drains surélevés avec le procédé décrit en partie II-E.3.2. (Épitaxie SiGe 30% à 650°C).

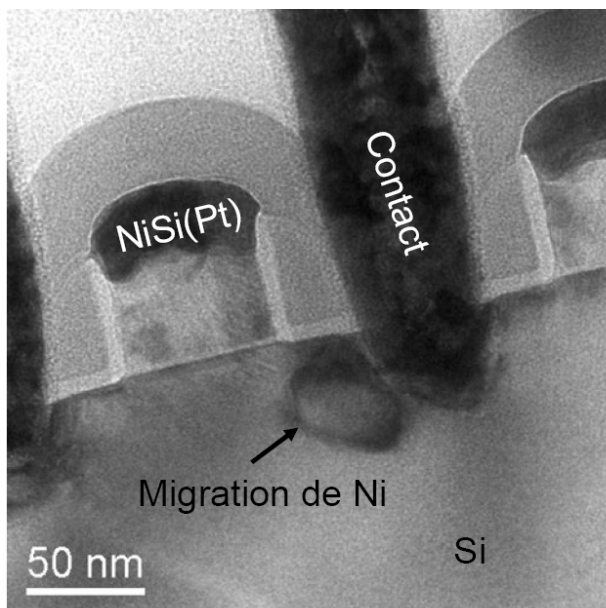


Figure 2.10 : Observation d'une coupe par TEM d'un transistor CMOS défectueux présentant une migration anormale du nickel [Imbert09]

De plus, nous n'avons pas pu étudier le phénomène de migration anormale du nickel (ou « encroachment »). Ce phénomène est activé par les différents budgets thermiques nécessaires à la réalisation du back-end. Il est responsable de l'apparition de fuite de jonctions et d'une diminution du rendement des dispositifs fonctionnels.

Comme ce phénomène est rare, il est nécessaire d'observer un nombre important de transistors dans une structure très dense. Nous noterons qu'aujourd'hui dans les filières industrielles, le budget thermique du back end a été drastiquement réduit pour lutter contre le phénomène d'« encroachment » [Bonnetier07].

Il a été montré par ailleurs, que l'introduction de Pt dans le NiSi limite le

phénomène d'encroachment [Strane07].

Pour statuer définitivement sur la possibilité d'intégrer les siliciures stabilisés que nous avons développés, il faudra réaliser une étude de rendement.

B. 2 -La caractérisation de l'empilement de grille

Le HfO_2 utilisé comme oxyde de grille est déposé après un nettoyage qui conduit à la formation d'un oxyde SiO_2 piédestal (cf. Figure 2.11). Au cours des étapes du procédé de fabrication du transistor supérieur, cet oxyde piédestal est susceptible de croître.

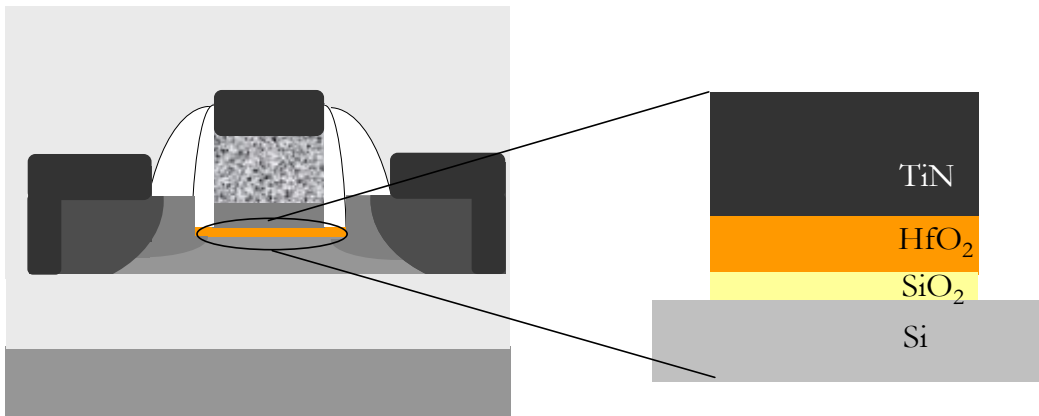


Figure 2.11 : Description de l'oxyde de grille, à l'épaisseur de HfO_2 déposée s'ajoute un oxyde de silicium susceptible d'être épaissi au cours du procédé de fabrication des transistors.

Dans l'intégration 3D, l'empilement de grille du transistor inférieur subira en plus du procédé de fabrication classique du transistor, les budgets thermiques nécessaires à la réalisation du transistor supérieur. Il est nécessaire de caractériser l'influence de ce budget thermique additionnel sur l'épaisseur de l'oxyde piédestal. .

La Figure 2.12 présente les mesures de capacité en fonction de la tension de grille ($C(V)$) sur les transistors de type n et p inférieurs après la fabrication du FET supérieur. L'extraction de l'EOT a été réalisée en comparant les mesures aux simulations résolvant les équations couplées de Poisson et de Shrödinger.

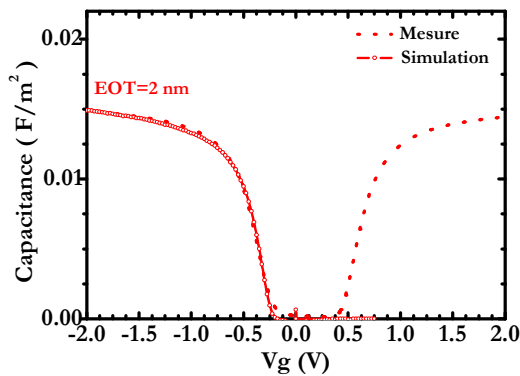


Figure 2.12: Mesures $C(V)$ du pFET inférieur et comparaison à la simulation. $L=2\mu$, $W=10\mu$, $f=500\text{kHz}$

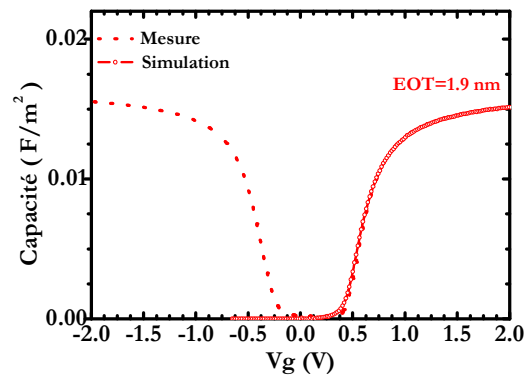


Figure 2.13: Mesures $C(V)$ du nFET inférieur et comparaison à la simulation $L=2\mu$, $W=10\mu$, $f=500\text{kHz}$

Nous observons en premier lieu que les mesures et les simulations se superposent de manière quasi idéale. Les valeurs extraites pour les EOT sont respectivement de 2 nm pour le pMOS et de 1.9 nm pour le nMOS.

Les plaques de référence permettant de quantifier précisément l'éventuelle recroissance de l'oxyde piédestal du transistor inférieur pendant la fabrication du transistor supérieur sont en cours de fabrication. Cependant, en faisant les hypothèses suivantes : $\epsilon_{\text{HfO}_2}=20$, et $\epsilon_{\text{SiO}_2}=3.9$, nous obtenons une EOT pour l'oxyde piédestal entre 0.9 et 1nm. Ces valeurs sont proches des valeurs obtenues classiquement sur des transistors avec un procédé standard à haute température. Au premier ordre, nous pouvons donc dire que s'il y a un phénomène de recroissance de l'oxyde piédestal pendant la réalisation du transistor supérieur, il est relativement limité.

La Figure 2.14 présente la figure de mérite de l'oxyde de grille du transistor inférieur après réalisation du transistor supérieur. Nous observons un gain de 3 décades en courant de fuite pour la même EOT en utilisant le HfO_2 à la place du SiO_2 comme isolant de grille. Ce gain correspond au gain obtenu de manière standard au Leti avec ce type d'empilement (empilement de grille de type « gate first » avec un recuit d'activation des dopants à haute température).

L'absence de structure de référence (sans réalisation de FET supérieur) ne nous permet pas de répondre avec précision quant à une éventuelle recroissance de l'oxyde piédestal ou à une dégradation de l'oxyde de grille avec la réalisation d'un FET supérieur à 600°C. Cependant par comparaison avec les résultats obtenus sur des procédés similaires au Leti, nous pouvons conclure que l'empilement de grille ne subit pas de modification majeure de ses caractéristiques électriques. Au premier ordre, le budget thermique du FET supérieur à 600°C peut être conservé sans dégrader l'empilement de grille.

Les figures suivantes présentent les extractions de la mobilité effective des trous et des électrons en fonction du champ effectif (E_{eff}), obtenues en appliquant la méthode « split-CV » [Koomen73, Sodini82] sur les transistors inférieurs (ayant subit la fabrication du FET supérieur). Des mobilités légèrement supérieures aux références internes avec le même empilement de grille sont obtenues (HfO_2 (3nm)/TiN (PVD)10nm).

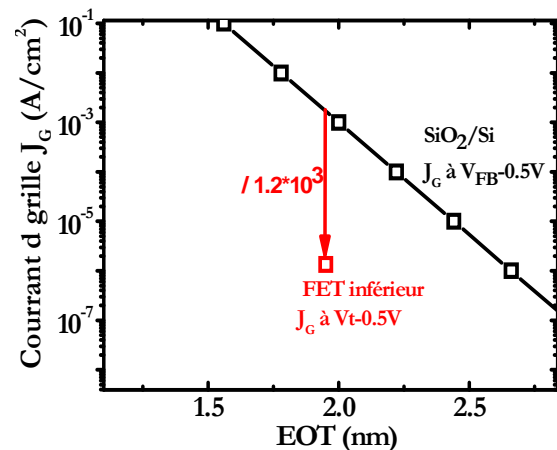


Figure 2.14 : Figure de mérite de l'empilement de grille du transistor inférieur.

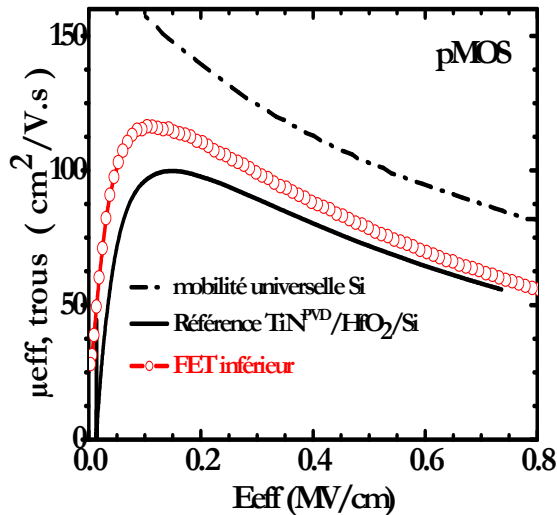


Figure 2.15: Extraction de la mobilité effective des trous en fonction du champ effectif.

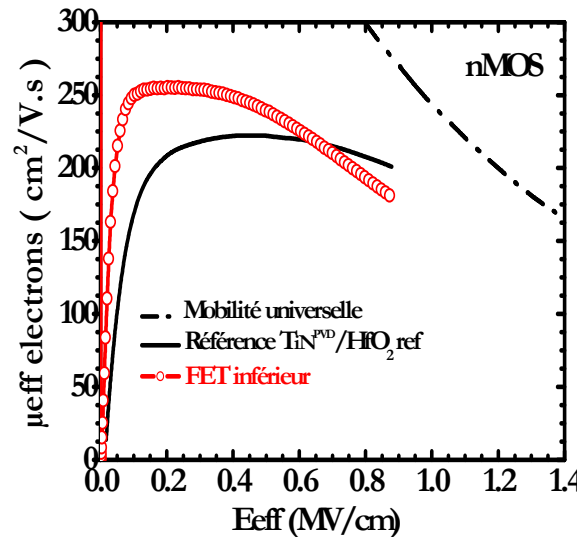


Figure 2.16: Extraction de la mobilité effective des électrons en fonction du champ effectif.

Ces valeurs de mobilités légèrement plus hautes pour le FET inférieur que pour la référence peuvent s'expliquer par l'épaisseur de HfO_2 plus grande dans le lot étudié que dans la référence (5 nm, contre 3nm). En effet, Garros et al [Garros08] ont montré que la présence d'azote proche de l'interface oxyde semi-conducteur dégrade les mobilités. Plus la source d'azote (le TiN) est éloignée de l'interface, moins la diffusion de l'azote vers l'interface est importante.

La réalisation du FET supérieur ne conduit pas à une dégradation des propriétés de transport dans le transistor inférieur. Ces mesures permettent de conclure que l'interface oxyde/semi-conducteur du transistor inférieur n'est pas dégradée par le budget thermique nécessaire à la réalisation du transistor supérieur, résultats cohérents avec les pentes sous le seuil montrées précédemment.

Concernant les éventuelles modifications dans le module de jonction (diffusion, désactivation), nous n'avons pas disposé au cours de la thèse de structures de tests ainsi que des variantes nécessaires à leur étude. Le masque utilisé ne possède que des transistors de longueur de grille $2\mu\text{m}$ et les plaques références (sans transistors supérieurs) sont en cours de fabrication.

Pour permettre ces études par la suite, des structures avec des longueurs de grille variable et plus avancées ont été dessinées dans un nouveau masque. Le suivi de l'évolution de la longueur effective des transistors peut être fait en utilisant par exemple la méthode décrite dans [Fleury09].

Dans le cas de la désactivation, la quantification sur transistor est plus complexe. L'étude du courant de conduction I_{ON} ne peut mener à une conclusion nette sur la désactivation car de nombreux paramètres peuvent contribuer à l'évolution du I_{ON} , comme par exemple une modification de l'épaisseur d'oxyde ou des travaux de sortie du silicium.

Pour étudier le phénomène de désactivation, il semble plus approprié de réaliser une étude sur substrats « pleines plaques » non siliciurés. Le niveau d'activation des dopants est alors déterminé à partir de la mesure de résistivité de la couche dopée (obtenue par la méthode Van Der Paw) et du profil SIMS (Secondary Ion Mass Spectrometry) de distribution des dopants dans la couche.

Ce qu'il faut retenir

III-A : Etude des caractéristiques du FET inférieur

- Les transistors du niveau inférieur restent fonctionnels après réalisation du transistor supérieur.
- La brique technologique siliciuration est essentielle pour pouvoir adresser des nœuds technologiques sub-45nm. Les transistors inférieurs sont siliciurés avec le procédé SALICIDE. La limitation du budget thermique du FET supérieur à 600°C et l'optimisation du siliciure de nickel permettent d'éviter l'agglomération de la couche.
- La réalisation du FET supérieur n'entraîne pas de modification majeure des caractéristiques du transistor inférieur en termes de recroissance d'oxyde piédestal, de qualité d'interface, de qualité de l'oxyde (fuite de grille) et de transport.

C - Etude des caractéristiques des FETs supérieurs

C.1 - pMOSFETs GeOI

C.1.1 - Résultats obtenus

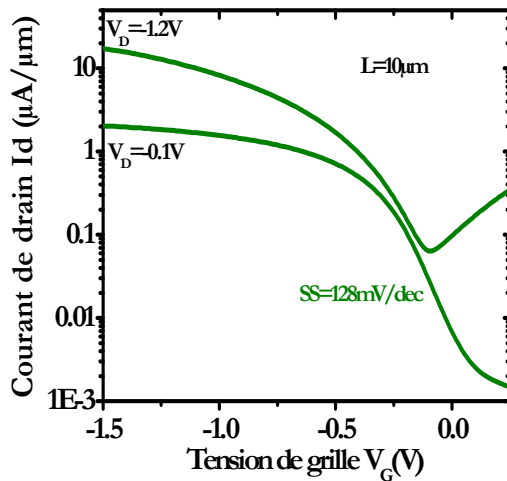


Figure 2.17: Caractéristiques I_D - V_G du pFET sur GeOI fabriqué sur le niveau supérieur.

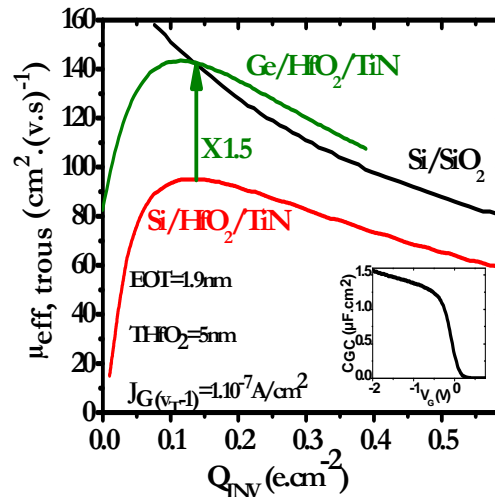


Figure 2.18 : Extraction de la mobilité des trous du p-FET GeOI supérieur par split CV.

La Figure 2.17 présente les caractéristiques I_D - V_G du transistor p-GeOI réalisé sur le niveau supérieur de l'empilement 3D. Les accès sont dopés bore (avec une dose de 6.10^{14} atomes. cm^2 et une énergie d'implantation de 5keV). Le budget thermique de l'épitaxie fine de silicium (capping) a été abaissé à 525°C afin de limiter l'agglomération en phase solide du film mince de germanium. Les dispositifs n'ont pas été germaniurés, car cette étape n'est pas technologiquement maîtrisée en interne.

Préalablement à la fabrication du transistor, la zone active a été contre-dopée avec de l'arsenic (avec une dose de 1.10^{13} atomes. cm^2 et une énergie d'implantation de 90keV) afin de compenser le dopage p non-intentionnel ou/et de bloquer le transistor parasite face arrière dû à une mauvaise interface face arrière [LeRoyer07][Romanjek08]. Cette étape permet d'obtenir un meilleur contrôle du rapport I_{ON}/I_{OFF} sans polarisation face arrière ainsi qu'une tension de seuil négative (cf. Figure 2.17).

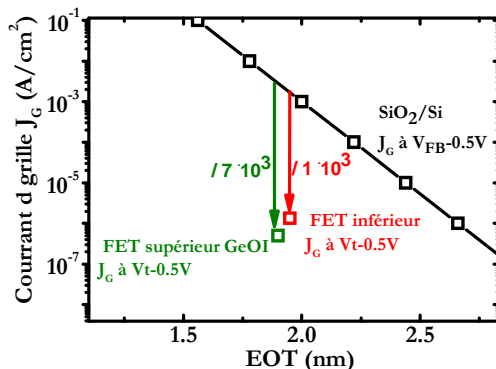


Figure 2.19 : Figure de mérite de l'empilement de grille sur GeOI (transistor supérieur)

Malgré la présence de dopants dans le canal, un gain en mobilité de 50% par rapport à une référence avec le même empilement de grille sur silicium est mesuré (cf. Figure 2.18).

Par rapport aux transistors sur Si, les transistors sur Ge présentent des courants de fuite de jonctions plus élevés à cause de son gap plus petit. D'autre part, la mobilité des trous dans les dispositifs y est plus élevée. Ces deux caractéristiques font du germanium un matériau plus adapté pour des applications hautes

performances.

L'épaisseur équivalente en SiO₂ de l'oxyde de grille est de 1.9nm pour une épaisseur de HfO₂ déposé de 5nm. Un courant de fuite de grille de l'ordre de 5.10⁻⁷A/cm² est obtenu. Par rapport à un empilement de grille SiO₂/Si un gain proche de 4 décades en courant de fuite de grille est obtenu (cf. Figure 2.19).

C.1.2 - Les voies d'améliorations possibles

Les caractéristiques du transistor Ge supérieur, sans être loin de l'état de l'art en termes de pente sous le seuil et de rapport I_{ON}/I_{OFF} [Kamata09, Mitard08], peuvent cependant être améliorées en travaillant sur les points suivants :

1-Amélioration de la qualité de l'interface oxyde-semiconducteur :

La réalisation d'un capping Si a pour but de se ramener à une interface oxyde-semiconducteur Si-SiO₂ classique bien contrôlée en termes de densité d'état d'interface. Or une diffusion de Ge dans le capping est observée [Hartmann09] [Mitard08].

Mitard et al montrent qu'une partie des défauts d'interface sont liés à la diffusion du Ge dans le capping Si jusqu'à atteindre l'interface Si-SiO₂. En réduisant le budget thermique de l'épitaxie (de 500°C à 350°C), ils parviennent à réduire la densité de défauts d'interface (d'environ d'un facteur 4), ainsi que la pente sous le seuil (de 20mV/dec pour atteindre 87mV/dec (L_g=10µm)). Cette amélioration de la qualité d'interface s'accompagne d'une variation de la tension de seuil vers des valeurs plus cohérentes avec le dopage réel des zones actives.

En conséquence, la dégradation de la pente sous le seuil de nos dispositifs peut en partie être imputée à la présence de germanium à l'interface Si/SiO₂. La réduction du budget thermique de l'épitaxie de silicium sur germanium peut alors permettre d'améliorer la pente sous le seuil de nos dispositifs, ainsi que leurs courants de conduction grâce à la suppression du contredopage responsable d'une chute de la mobilité des porteurs dans le canal [Romanjeck08].

2- Germaniuration des accès:

Cette étape permettrait de diminuer les résistances d'accès qui sont aux alentours de 900 Ohms.µm (mesure sur la filière Ge actuelle) et ainsi obtenir des courants de conduction plus élevés. L'état de l'art en termes de résistance d'accès sur transistors Ge germaniurés est de l'ordre de 140 Ohms.µm. [Mitard08]

3- Optimisation du module de jonction :

Les dispositifs germanium actuels présentent actuellement de forts courants de fuite de jonctions. Même si ces fuites sont intrinsèquement plus importantes sur germanium que sur silicium à cause de son faible gap, les courants de fuites peuvent être réduits (i) en travaillant sur le profil de jonction. La réduction de l'abrupté peut permettre de diminuer le courant tunnel bande à bande (BTBT, Band To Band Tunneling) (ii) en améliorant la qualité du matériau via la réduction du courant tunnel assisté par piège (TAT, Trap assisted Tunneling), (iii) en diminuant les tensions d'alimentation.

C.1.3 - Discussion préliminaire sur l'option Ge

A l'heure actuelle, la réalisation de transistors germanium compétitifs par rapport au transistor silicium nécessite de relever de nombreux défis technologiques. En premier lieu, la question du choix de l'empilement de grille permettant d'obtenir une bonne qualité d'interface et une pertinence pour des nœuds technologiques avancés (EOT<1nm) reste ouverte.

D'autre part, l'intérêt du germanium réside dans ses mobilités de porteurs plus importantes que dans le silicium (cas matériau massif). Or nous observons actuellement que les gains en mobilité obtenus dans les transistors sont relativement faibles (Gain de $\sim 100\%$ pour les trous [Chui03] [Yu04b] et gain inexistant dans le cas du nMOS (cf. Figure 2.20). De plus, les techniques de contraintes développées sur silicium montrent actuellement des bénéfices en mobilité supérieurs (par exemple une augmentation de mobilité de trous d'un facteur trois mesuré sur substrat (100) pour un stress uniaxial en compression selon $\langle 110 \rangle$ avec une contrainte de 2GPa et un facteur 2 sur la mobilité des électrons avec un stress uniaxial en tension selon $\langle 110 \rangle$ et une contrainte de 3GPa) [Weber07]. Pour apporter un réel gain en mobilité par rapport au silicium, le germanium doit être contraint. Weber et al démontre un gain d'un facteur neuf pour les trous pour un substrat de Ge contraint [Weber05].

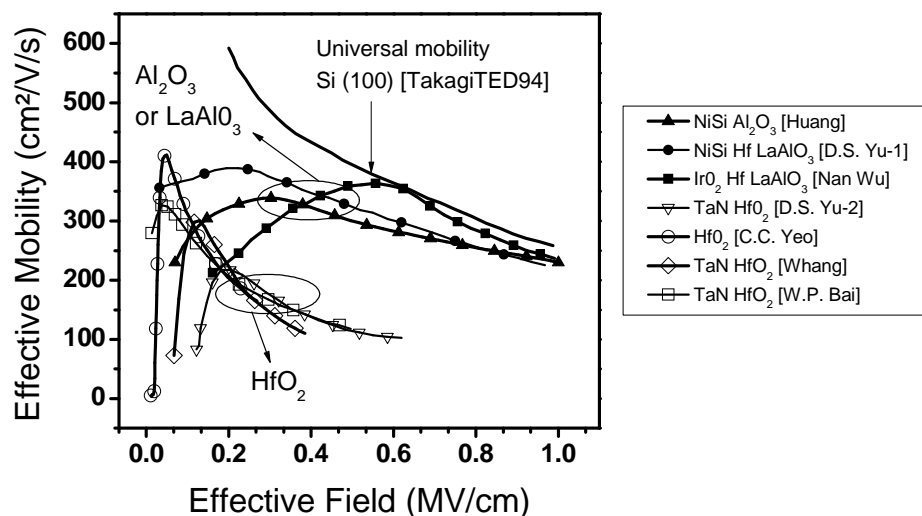


Figure 2.20 : Etat de l'art des extractions de mobilité dans les dispositifs nMOS Ge [Rafhay08]

D'autre part, dans le cas où les performances du transistor nMOS resteraient aussi faibles que démontrées actuellement, une co-intégration pFET Ge/ nFET Si s'impose. Cette possibilité est facilement offerte par l'intégration 3D séquentielle, où les cellules CMOS peuvent être dessinées à cheval sur les deux niveaux empilés.

Cependant, l'impossibilité de réaliser des cellules CMOS sur germanium peut apparaître comme une limitation importante par rapport à l'option SOI/SOI. L'influence de ce paramètre sera discutée dans l'étude de gain en densité présentée au chapitre III.

De plus, les courants de fuite intrinsèquement plus importants sur germanium que sur silicium semblent limiter le germanium à des applications hautes performances. Or dans des circuits réels, il est nécessaire de pouvoir aussi réaliser des transistors à basse consommation (pour la SRAM par exemple). *A contrario*, le transistor silicium, avec ces fuites de jonction plus faible permet de réaliser des zones « haute performance » et « basse consommation » en adaptant la tension de seuil des transistors. En conséquence, l'option SOI/SOI semble plus pertinente pour réaliser des circuits intégrés.

C. 2 - MOSFETs Silicium à bas budget thermique

C.2.1 - Description des variantes du lot d'étude

Les résultats présentés dans les parties suivantes ne sont pas des résultats obtenus sur un niveau de transistor intégré en trois dimensions. L'étude préliminaire sur un masque avec des dimensions plus avancées (L_G jusqu'à 30 nm) que sur le masque 3D (L_G limité à 0.5 μ m), nous permet d'étudier par exemple l'évolution des effets canaux courts avec le procédé bas budget thermique.

Dans le cas du transistor SOI, la modification majeure du procédé de fabrication par rapport au transistor inférieur est le remplacement du recuit d'activation à haute température (1050°C) par une activation de type SPE à 600°C.

Comme la diffusion des dopants pour ce faible budget thermique est très faible, l'étape de réalisation du premier espaceur avant l'étape d'implantation des accès faiblement dopés (LDD Low Doped Drain) est supprimée de manière à réaliser la jonction à l'aplomb de la grille. (cf. Figure 2.21).

Des poches de dopage de type opposé à celui du transistor sont introduites afin de lutter contre les effets canaux courts (l'épaisseur du film est dans cette étude de 30nm, ainsi les transistors avec des longueurs de grilles avancées (<100nm) ne seront pas totalement déplétés). Les zones d'accès fortement dopées (HDD High Doped Drain) sont réalisées après la fabrication des espaceurs.

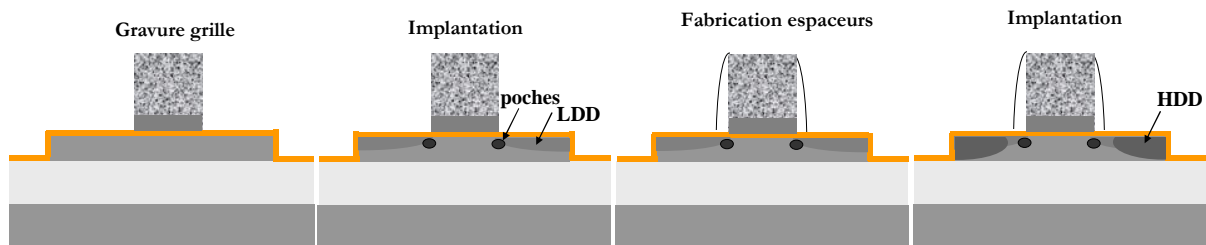


Figure 2.21 : Description du procédé de réalisation des jonctions dans le cas du transistor silicium basse température.

De manière à comparer le procédé SPER à une activation thermique classique, nous avons gardé la même architecture (sans premiers espaceurs). Ce schéma d'intégration n'est pas adapté pour des accès activés thermiquement, c'est pourquoi la référence aura un comportement détérioré en termes de contrôle des effets canaux courts par rapport à un transistor optimisé (avec les deux espaceurs).

Les variantes d'implantation sont résumées dans la table 3.1 : Une variante dans la profondeur des LDD a été réalisée.

		Références recuit à 1050°C		SPE LDD fine		SPE LDD profondes	
		Type N	Type P	Type N	Type P	Type N	Type P
LDD	Implantation préamorphisation				Ge 5keV 5.10^{14}		Ge 9keV 5.10^{14}
	Implantation dopants	As 5keV 1.10^{15}	BF ₂ 5keV 1.10^{15}	As 5keV 8.10^{14}	B 3keV 8.10^{14}	As 8keV 8.10^{14}	B 3keV 8.10^{14}
Poches	Implantation dopants	B 5keV 2.10^{13}	As 20keV 4.10^{13}	B 5keV 2.10^{13}	As 20keV 4.10^{13}	B 5keV 2.10^{13}	As 20keV 4.10^{13}
HDD	Implantation préamorphisation				Ge 13keV 5.10^{14}		Ge 13keV 5.10^{14}
	Implantation dopants	As 10keV 2.10^{15}	BF ₂ 8keV 2.10^{15}	As 10keV 2.10^{15}	B 3keV 2.10^{15}	As 10keV 2.10^{15}	B 3keV 2.10^{15}

Table 3.1 : Récapitulation des variantes d'implantations. Les doses sont données en atomes.cm² et les énergies en keV.

Les conditions d'implantations ont été calibrées à l'aide du logiciel de simulation Cristal Trim. [CTRIM]

Pour la réalisation des poches du nMOS, le bore est implanté directement sans étape d'amorphisation préalable, car pour de petites doses, le bore peut être activé dans de bonnes proportions à 600°C. Woodard et al montre par exemple que 60% d'une dose de bore à 8.10^{12} cm⁻² est activé par un recuit à 600°C sans étape de préamorphisation. [Woodard06].

Les Figure 2.22 et Figure 2.23 présentent les résultats de simulations CTRIM avec les paramètres utilisés sur ce lot.

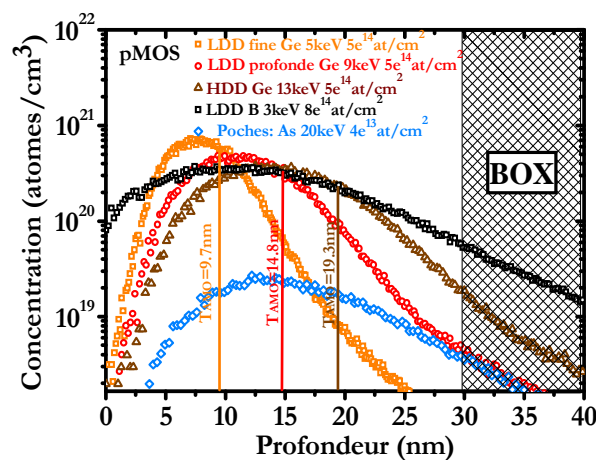


Figure 2.22 : Profils d'implantations simulés avec CTRIM- cas pMOS

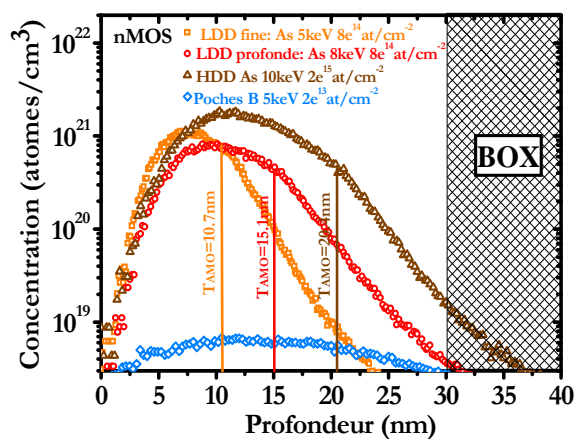


Figure 2.23: Profils d'implantations simulés avec CTRIM- cas nMOS

Sur ces simulations sont reportées les profondeurs de silicium amorphisées par les implantations.

Dans le cas des jonctions de type n, l'amorphisation est obtenue par l'implantation du dopant lui-même (l'arsenic). Les conditions d'implantations de l'As sont choisies pour amorphiser 10, 15, 20 nm de silicium dans le cas de l'implantation LDD fine, LDD profonde et HDD respectivement.

Dans le cas des jonctions de type p, l'amorphisation est réalisée par l'implantation préalable de germanium. De la même manière les conditions d'implantation du Ge sont prévues pour amorphiser 10, 15, 20nm de silicium dans le cas de l'implantation LDD fine, LDD profonde et HDD respectivement. Cependant, il faut noter que les conditions d'implantation du dopant (le Bore) sont identiques pour les variantes LDD fine et LDD profonde.

Le caractère fin ou profond des LDD est considéré comme lié à la profondeur de la zone amorphisée et non au profil de concentration chimique du dopant. Plus précisément nous supposons que seuls les dopants situés dans la zone amorphisée sont activés lors du recuit de recristallisation à 600°C.

C.2.2 - Analyse des tendances : tests paramétriques

Les Figure 2.24 et Figure 2.25 présentent les valeurs des compromis I_{ON}/I_{OFF} des variantes SPER comparés à la référence haute température. Les longueurs de grille des dispositifs varient de 10 μm à 30nm. Les tests électriques sont réalisés sur 11 puces par plaques. Les puces sont CMOS.

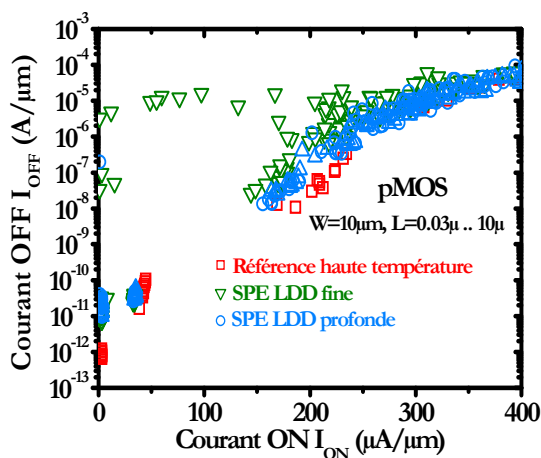


Figure 2.24 : Figure de mérite I_{ON}/I_{OFF} du procédé à basse température comparé à la référence haute température. Cas pMOS

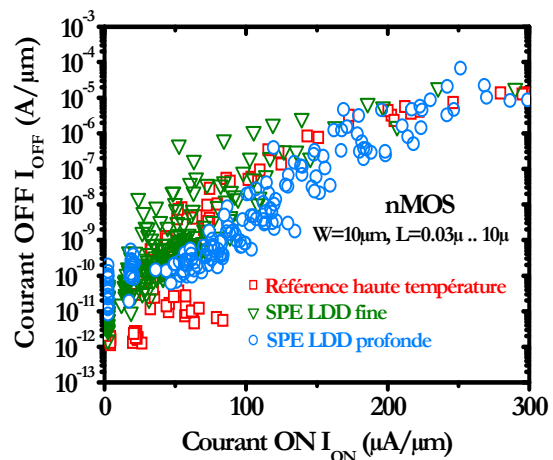


Figure 2.25 : Figure de mérite I_{ON}/I_{OFF} du procédé à basse température comparé à la référence haute température. Cas nMOS

Nous observons que les variantes SPE ont un comportement proche des variantes haute température. Ces courbes mettent en évidence que la SPER a permis d'activer les jonctions des transistors à bas budget thermique.

L'influence du faible budget thermique sur le contrôle des effets canaux courts a été étudiée et les résultats sont présentés dans les Figure 2.26 à Figure 2.30. Pour chaque longueur de dispositif, les mesures électriques des paramètres présentés sont moyennés (11puces/ variante)

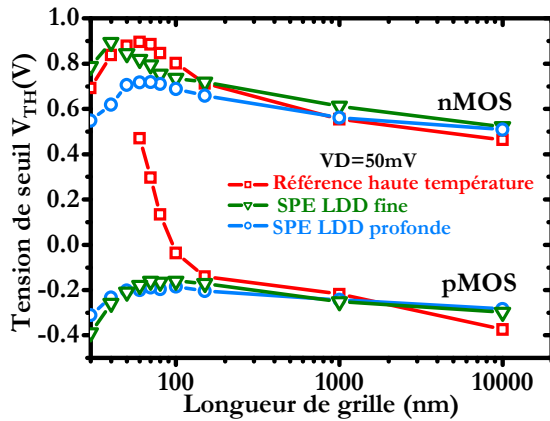


Figure 2.26 : Evolution de la tension de seuil à $|V_D|=50\text{mV}$ en fonction de la longueur de grille des transistors

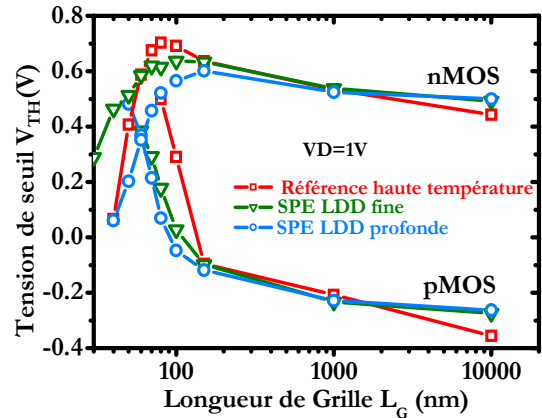


Figure 2.27 : Evolution de la tension de seuil à $|V_D|=1\text{V}$ en fonction de la longueur de grille des transistors

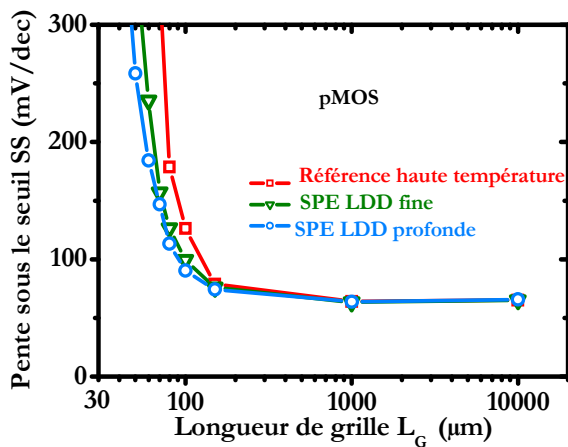


Figure 2.28 : Evolution de la pente sous le seuil des transistors pMOS en fonction de la longueur de grille.

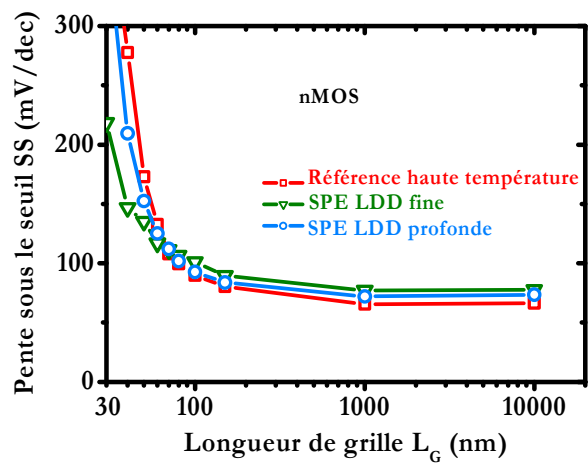


Figure 2.29 : Evolution de la pente sous le seuil des transistors nMOS en fonction de la longueur de grille.

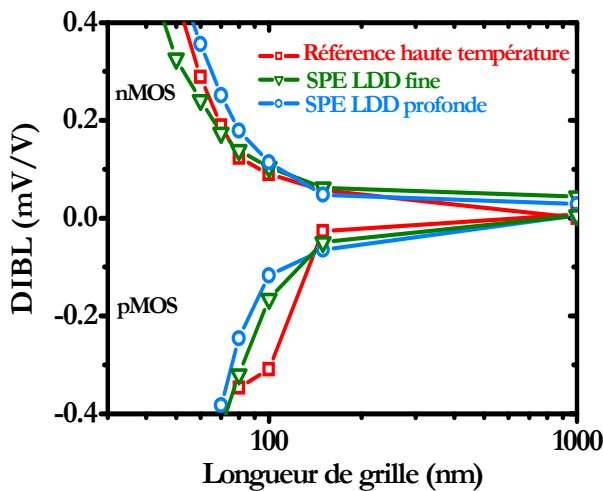


Figure 2.30 : Evolution du DIBL (Drain Induced Barrier lowering) en fonction de la longueur de grille.

Toutes ces figures de mérite confirment que la réalisation du FET à basse température est bénéfique en termes de contrôle des effets canaux courts. Cette amélioration est attribuée à la réduction de la diffusion des dopants grâce à l'abaissement du budget d'activation des dopants qui passe de 1050°C à 600°C . En effet, la réduction (voire suppression de la diffusion) permet la réalisation de jonctions plus fines permettant un meilleur contrôle des effets canaux courts [Skotnicki88].

Dans le cas du nMOS, nous observons une nette influence positive de la réduction de la profondeur des LDD sur le contrôle des effets canaux courts. Cette influence est visible sur l'ensemble des figures de mérite de contrôle

des effets canaux courts présentées. Par exemple, sur le graphe d'évolution de la tension de seuil à V_D fort, nous observons que pour une longueur de 50nm, la variante LDD fine permet de garder une tension de seuil quasi égale à la valeur obtenue pour $L_G=1\mu m$, alors dans le cas de la variante LDD profonde la tension de seuil diffère d'environ 300mV par rapport à la valeur à $1\mu m$.

Dans le cas du pMOS, la différence entre les variantes LDD fine et profonde est moins nette. Une explication possible serait que dans le cas du bore, l'activation puisse avoir lieu de manière non négligeable en dehors de la zone amorphisée. Dans ce cas, puisque le profil de bore est constant dans les deux variantes, les profils des jonctions pourraient être similaires.

Dans la littérature, les études sur la possibilité d'activer les dopants avec de fortes concentrations sans avoir préalablement amorphisé conduisent à des résultats divergents. Une méthode permettant de connaître la concentration de dopants actifs en fonction de sa position dans le film de silicium est la SRP (Spreading Resistance Profile). Dans [Lindsay04], les mesures SRP tendent à montrer que seuls les dopants placés dans la zone amorphisée sont activés (cf. Figure 2.31), alors que dans [Pawlak04] (cf. Figure 2.32), nous n'observons pas de décrochement dans la concentration de bore actif au niveau de l'interface Si amorphe-Si cristallin.

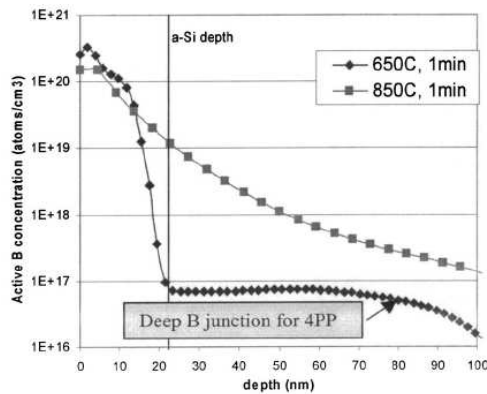


Figure 2.31 : Mesure SRP sur jonctions SPER donnant les concentrations de dopants actifs pour un recuit à 650°C par rapport à la région amorphisée. Cas du Bore [Lindsay04]

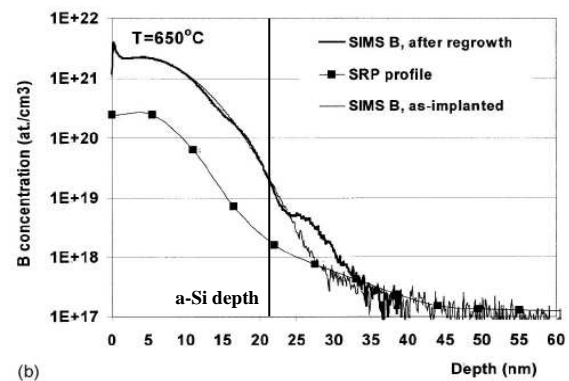


Figure 2.32 : Mesure SRP sur jonctions SPER donnant les concentrations de dopants actifs pour un recuit à 650°C par rapport à la région amorphisée. Cas du Bore [Pawlak04]

D'autre part, Woodard et al montrent qu'une partie du bore implanté peut être activé à bas budget thermique (600C°) sans préamorphisation [Woodard06] (Pour une dose de 3.10^{15} atomes/cm², 5% des atomes de bore sont activés).

En conclusion, il nous est donc difficile de trancher sur ce point. Cependant, l'hypothèse d'une activation du bore en dehors de la zone amorphisée est une explication plausible aux résultats similaires en termes de contrôle d'effet canaux courts, obtenus pour les variantes LDD fines et profondes de type p.

Par la suite, plusieurs améliorations peuvent permettre d'augmenter le contrôle des effets canaux courts.

En premier lieu, un travail d'ingénierie de jonction avec un plan d'expérience complet permettra d'optimiser leur profil. Rappelons que les résultats présentés correspondent à une première démonstration et que les conditions du procédé méritent d'être affinées.

En second lieu, les épaisseurs des films d'active sont d'environ 30nm. Avec de telles épaisseurs, et pour des dispositifs avec des longueurs de grilles <100nm, nous ne pouvons pas profiter du meilleur contrôle électrostatique permis par l'utilisation de l'architecture FDSOI.

C.2.3 - Présentation des caractéristiques unitaires

Les Figure 2.33 et Figure 2.34 présentent les caractéristiques I_D-V_G d'un FET de longueur 150 nm en régimes linéaire et saturé respectivement.

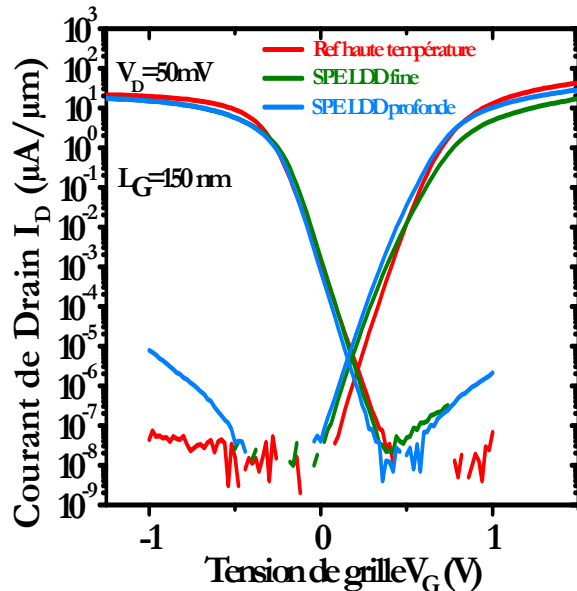


Figure 2.33 : I_D-V_G en régime linéaire des transistors SPER et comparaison à la référence haute température.

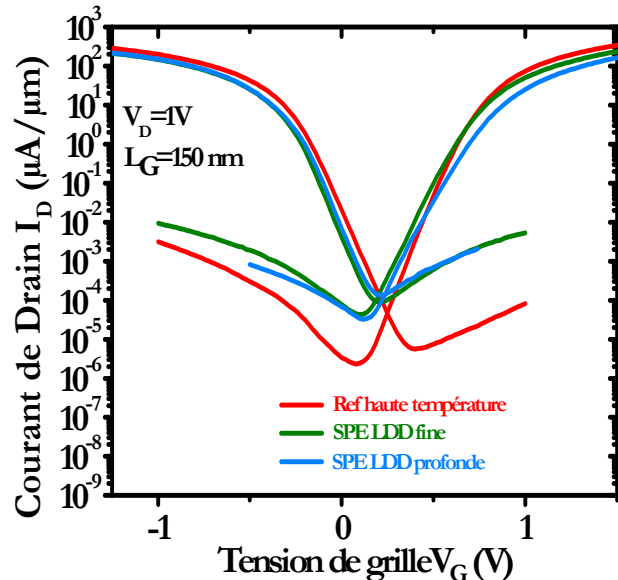


Figure 2.34 : I_D-V_G en régime saturé des transistors SPER et comparaison à la référence haute température.

Nous constatons que les fuites de jonctions sont plus importantes dans le cas SPE que dans la référence haute température. Cette dégradation s'explique par deux phénomènes : Tout d'abord, la préamorphisation conduit à la formation de défauts résiduels (End Of Range : EOR) liés à un excès d'interstitiels situés au delà de l'interface Si amorphe/Si cristallin et dont la dissolution est incomplète avec un budget thermique de (600°C) [Pawlack04, Lindsay04]. Une réduction de la concentration de ces défauts dans la zone de déplétion de la jonction peut être obtenue en réduisant la distance entre l'interface Si amorphe/Si cristallin et l'interface Si/BOX. [Hamilton06][Fazzini08].

D'autre part, la SPE conduit à la réalisation de jonctions très abruptes puisque la majeure partie des dopants activés sont placés dans la zone amorphisée, et que la diffusion des dopants à 600°C est réduite (voire supprimée). Le caractère des jonctions obtenues avec la technique de SPER, plus abrupt qu'avec une activation thermique à haute température, conduit à une augmentation du courant tunnel bande à bande.

Il se peut que cette spécificité rende ce procédé incompatible avec des applications à très basse consommation. Les études réalisées ne nous permettent pas de conclure sur ce point. Pour la suite des études, il serait intéressant de statuer sur l'origine de l'élévation du courant I_{OFF} avec le procédé basse température. Résulte-t-il d'une augmentation du BTBT ou d'un TAT plus fort à cause de la présence des défauts EOR dans la zone de déplétion des jonctions? Dans le second cas, cette contribution peut être éventuellement réduite en optimisant la profondeur d'amorphisation pour qu'elle soit suffisamment proche du BOX, tout en laissant suffisamment de zone active cristalline pour servir de germe à la recroissance.

Grâce à l'intégration de poches, nous pouvons présenter des caractéristiques de FET avec des longueurs de grille de nœuds technologiques plus avancés. (cf. Figure 2.35 et Figure 2.36)

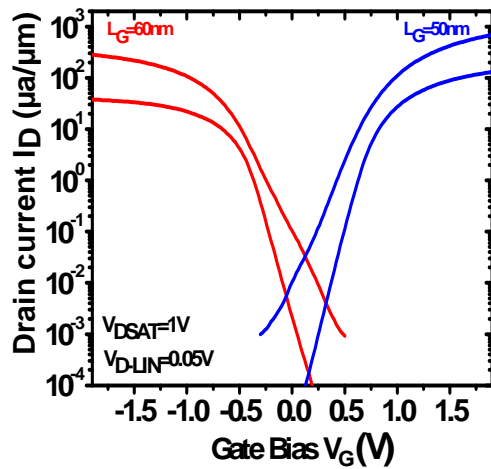


Figure 2.35 : Courbes I_D - V_G de transistors n et pFET à bas budget thermique. $L_{Gp}=60\text{nm}$ et $L_{Gn}=50\text{nm}$

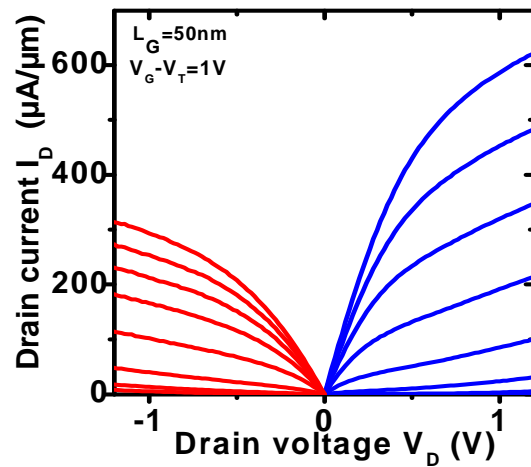


Figure 2.36 : Courbes I_D - V_D de transistors n et pFETs à bas budget thermique $L_{Gn}=L_{Gp}=50\text{nm}$

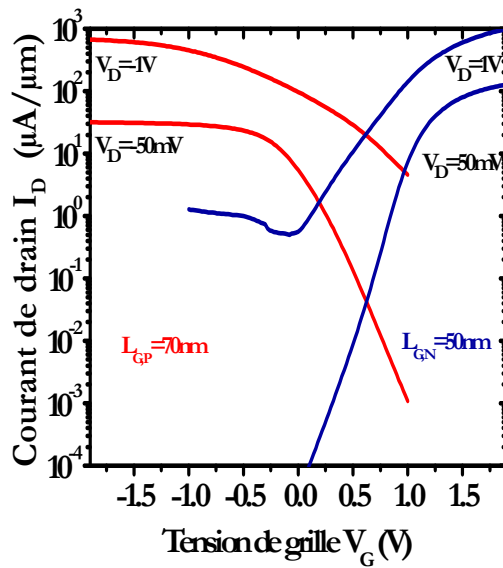


Figure 2.37 : Courbes I_D - V_G de transistors n et pFET à haut budget thermique. $L_{Gp}=70\text{nm}$ et $L_{Gn}=50\text{nm}$

La Figure 2.37 présente les caractéristiques I_D - V_G obtenues avec le procédé haute température pour des longueurs de grilles similaires à celles de la Figure 2.35. En comparant les deux figures, nous observons nettement l'influence positive du procédé basse température sur le contrôle des effets canaux courts.

C.2.4 - Caractérisation des résistances d'accès

L'extraction des résistances d'accès a été réalisée avec la méthode proposée par Fleury et al [Fleury09] à partir de mesures faites sur l'intégration complète : à la différence de la partie précédente, les transistors bas budget thermique sont empilés au dessus d'un premier niveau de transistor.

Nous extrayons des résistances d'accès totales de $390 \pm 50 \Omega \cdot \mu\text{m}$ pour le pFET et de $375 \pm 50 \Omega \cdot \mu\text{m}$ pour le nFET, i.e. $R_s \sim R_d \sim 200 \Omega \cdot \mu\text{m}$ pour les deux types de transistors. Ces valeurs

témoignent de l'activation des dopants via le procédé SPER à faible température. Cependant, elles restent un peu élevées par rapport à l'état de l'art [Yako08], surtout en ce qui concerne le nMOS. Habituellement, les résistances séries du nMOS sont environ deux fois plus faibles que celles du pMOS. Ces fortes valeurs de résistance d'accès expliquent les faibles valeurs du courant de conduction I_{ON} du nMOS (cf. Figure 2.25). Ces résultats sont néanmoins encourageants pour un premier essai.

En conclusion, l'utilisation du procédé SPER pour doper les accès nous permet de fabriquer avec succès des transistors avec un budget thermique limité à 600°C.

Pour bénéficier du meilleur contrôle électrostatique apporté par l'architecture FDSOI, il faut que le rapport entre l'épaisseur du film d'active et la longueur de la grille soit au moins d'un facteur 3 [Lolivier05]. Pour des dispositifs avancés (à partie du nœud 45nm), la zone active devrait alors mesurer 10 nm. Pour cette gamme d'épaisseur, l'étape d'amorphisation devient complexe car il faut conserver une partie du film cristallin qui servira de germe pour la recristallisation.

Pour permettre la réalisation des jonctions pour des films minces, une première option pourrait consister à surélever les sources et drains. Pour éviter de réaliser un court circuit entre les accès et la grille, il est nécessaire de réaliser un espaceur. Sa largeur doit être optimisée avec attention, à cause du caractère non diffusant du procédé SPE. Si la zone d'active située en dessous de l'espaceur se retrouve non dopée, les résistances d'accès seront alors nettement dégradées. L'inclinaison des implantations peut permettre de doper sous les espaceurs quand ceux-ci sont de petite taille.

Une seconde option serait d'amorphiser l'ensemble des accès puis d'utiliser la zone cristalline sous la grille comme germe pour une recristallisation latérale. Cette option a été expérimentée sur une variante de lot. La Figure 2.38 présente une observation TEM de zone active en bord de grille après le recuit de recristallisation à 600°C, 2min. Nous n'observons pas de recristallisation latérale

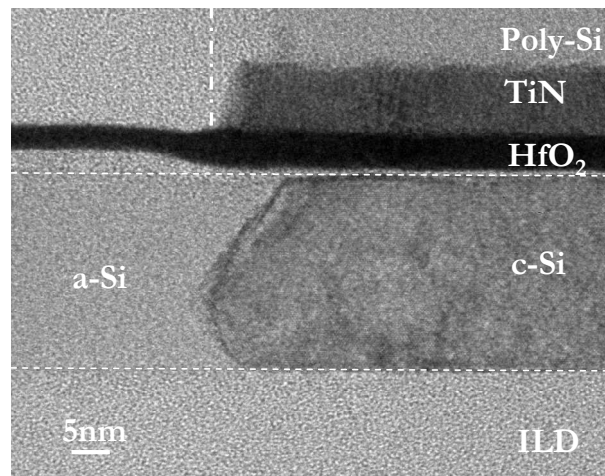


Figure 2.38 : Photo TEM en coupe d'un transistor avec accès totalement amorphisé après recuit d'activation (600°C, 2min)

des accès suffisante même si les contacts sont positionnés très proches de la grille comme c'est le cas dans des technologies avancées (~40 nm valeur DRM 45, ~25 nm valeur DRM 22 nm).

C.2.5 - Caractérisation de l'empilement de grille

Par rapport au procédé du transistor inférieur, les budgets thermiques appliqués pour réaliser le transistor supérieur ont été réduits. Ces modifications de budget thermique (récapitulés dans le chapitre I partie E) sont susceptibles de modifier les propriétés électriques de l'empilement de grille, en particulier au niveau de l'oxyde piédestal.

La Figure 2.39 présente les mesures C(V) des transistors haut et bas. Les valeurs des EOT ont été extraites avec précision en comparant les résultats de simulation résolvant les équations couplées de Poisson et de Shrodinger.

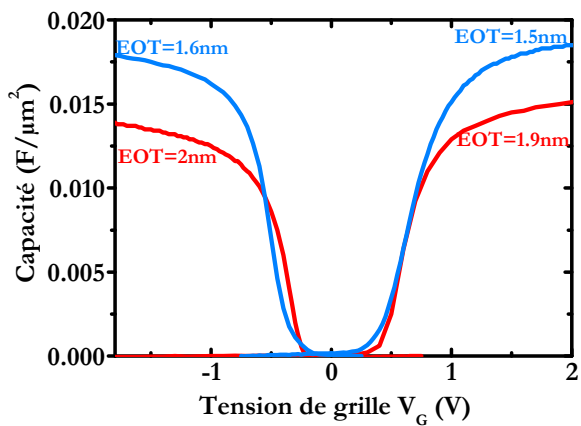


Figure 2.39 : Mesures C(V) des transistors inférieurs et supérieurs.

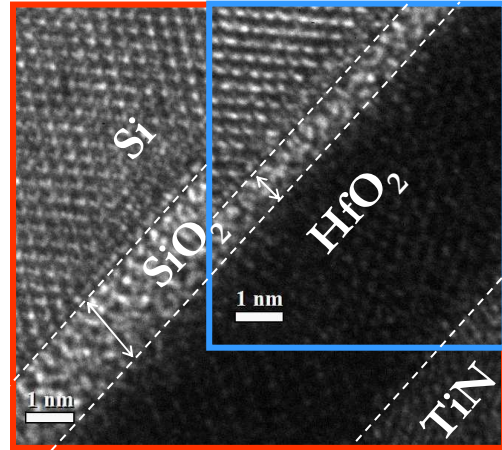


Figure 2.40 : Observation en coupe des empilements de grille des transistors inférieurs et supérieurs.

Nous observons que les EOT des transistors supérieurs sont 0.4nm plus faibles que celles du transistor inférieur. Comme les épaisseurs de HfO₂ sont égales dans les deux variantes, ce résultat indique que la réalisation du transistor supérieur à basse température permet de limiter la recroissance de l'oxyde piédestal au cours du procédé de fabrication. Cette hypothèse est renforcée par l'observation de la photo TEM présenté en Figure 2.40 confirmant que le transistor supérieur (à bas budget thermique) possède un oxyde piédestal plus mince que le transistor inférieur (à haut budget thermique).

Cette réduction d'épaisseur de l'oxyde piédestal peut être considérée comme un atout majeur du procédé basse température. En permettant de réduire la contribution de l'épaisseur de l'oxyde piédestal dans l'EOT globale des dispositifs, le procédé à bas budget thermique permet d'obtenir des EOT suffisamment faibles pour adresser des nœuds technologiques avancés.

Les Figure 2.41 et Figure 2.42 présentent les extractions de mobilité effective en fonction du champ effectif, extraites par la méthode split-CV.

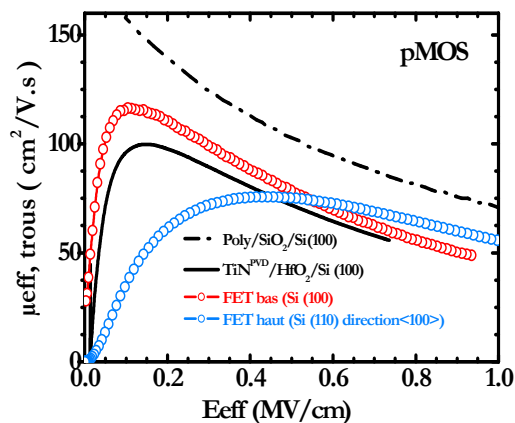


Figure 2.41 : Extractions des mobilités effectives des trous dans le transistor supérieur par split CV.

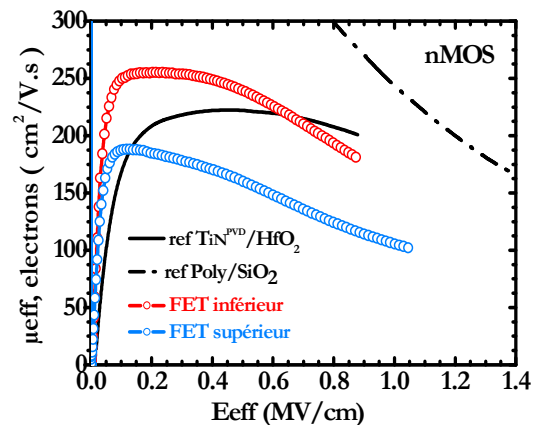


Figure 2.42 : Extractions des mobilités effectives des électrons dans le transistor supérieur par split CV.

Dans le cas du nMOS, nous mesurons une mobilité clairement réduite par rapport au transistor inférieur. Cette réduction de mobilité peut être expliquée par la réduction de l'oxyde piédestal et donc par la plus grande proximité entre le high-k et le canal. [Mitard08].

Dans le cas du transistor pMOS, la comparaison entre le FET supérieur et inférieur ne peut être faite car les transistors sont fabriqués sur des substrats avec des orientations de surface différentes (le pMOS est réalisé sur substrat Si (110) dans la direction de conduction $\langle 100 \rangle$ sur le niveau supérieur).

Ce qu'il faut retenir

III-C : Etude des caractéristiques du FET supérieur

- Un budget thermique maximum de 650°C a été fixé pour préserver les caractéristiques du transistor inférieur.
- Des procédés complets de fabrication de transistor SOI et GeOI ont été développés avec un budget thermique maximal de 650°C. Les transistors supérieurs sont réalisés sur des zones actives reportées par collage moléculaire. Les zones actives sont alors monocristallines et leur épaisseur vaut 30 nm dans le cas du MOS SOI et 50 nm dans le cas du MOS GeOI.
- Le FET sur germanium est particulièrement adapté à la réalisation du niveau supérieur d'une intégration 3D séquentielle car l'activation des dopants est réalisée par recuit thermique à 600°C seulement.
- Le gain en mobilité obtenu dans les transistors pMOSFET Ge par rapport au Si, et les courants de fuite intrinsèquement plus importants sur Ge que sur Si, font du germanium, un semiconducteur plus adapté pour des applications hautes performance.
- D'autre part, dans le cas où les performances du transistor nMOS restent aussi faibles que démontrées actuellement, une co-intégration pFET Ge/ nFET Si est une option intéressante. Cette possibilité est facilement offerte par l'intégration 3D séquentielle, où les cellules CMOS peuvent être dessinées à cheval sur les deux niveaux empilés.
- Dans le cas des transistors SOI, l'étape d'activation des dopants par recuit thermique à ~1000°C doit être remplacée par une technique d'activation alternative à basse température.
- La recroissance en phase solide (SPER : Solid Phase Epitaxy Regrowth) a été choisie pour activer les dopants à 600°C. La recristallisation du silicium amorphisé permet de placer les dopants en site substitutionnel sans avoir à utiliser des budgets thermiques élevés.
- Cette technique de dopage alternative a été démontrée avec succès sur des transistors. L'extraction des résistances d'accès (~200 Ohms.µm pour le p et le nMOS) démontre une activation satisfaisante des dopants.
- En plus de rendre compatible le transistor silicium avec l'intégration 3D séquentielle, ce mode d'activation présente des avantages par rapport à une activation thermique à haute température : Elle conduit non seulement à la formation de jonctions fines permettant d'améliorer le contrôle des effets canaux courts (via une diffusion réduite voire supprimée) mais aussi à des EOT plus faibles (via une recroissance de l'oxyde piédestal moins importante (0,4 nm)).

- Une augmentation du courant I_{OFF} est observée, cette augmentation peut être partiellement imputée aux défauts EOR apparaissant au delà de l'interface Si amorphe/ Si cristallin non guéris avec des recuits à 600°C. Dans une intégration de type FDSOI, cette contribution peut éventuellement être minimisée ou supprimée en plaçant l'interface Si amorphe/Si cristallin suffisamment proche du BOX.

D - Démonstration de cellules en 3D

D. 1 - L'inverseur

L'inverseur est la cellule CMOS la plus simple et aussi la plus commune dans un circuit électronique. Son schéma électrique est représenté en Figure 2.43, et les différentes variantes fabriquées sont récapitulées en Figure 2.44 .

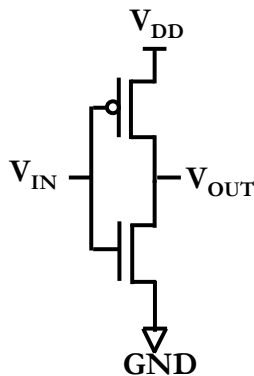


Figure 2.43 : Description de l'inverseur

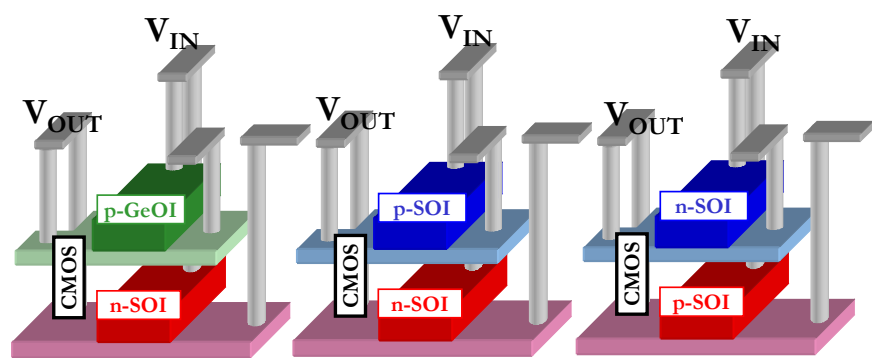


Figure 2.44 : Différentes variantes fabriquées. Option co-intégrée pMOS GeOI au dessus d'un nMOS SOI et options SOI sur SOI, n MOS en haut ou en bas.

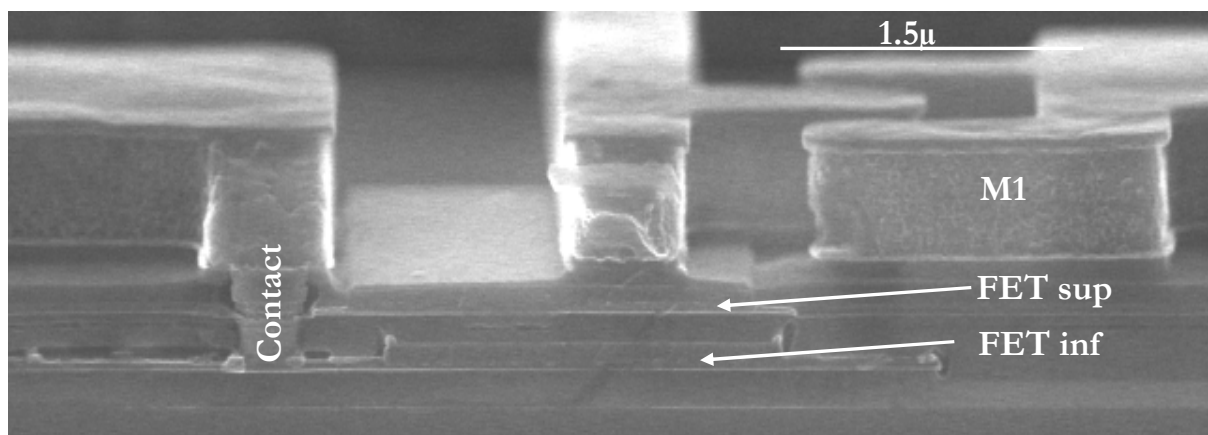


Figure 2.45 : Vue en coupe d'un inverseur « à cheval » sur deux niveaux observé par MEB.

Les cellules CMOS sont bien « à cheval » sur les deux niveaux. Cette possibilité est un atout particulier de l'intégration 3D séquentielle, où les performances d'alignement sont telles que les interconnexions 3D peuvent être réalisées à l'échelle du transistor. *A contrario*, dans l'intégration

3D parallèle, les performances d'alignement ne permettent pas d'interconnecter avec une densité suffisante les transistors du niveau haut et ceux du niveau bas. Actuellement, il existe au minimum un facteur 100 entre la densité de vias 3D reliant les deux niveaux et la densité de contacts dans un niveau (cas nœud 45nm).

La fonctionnalité d'un inverseur est caractérisée par la mesure de la courbe de transfert en tension. Le signal doit être inversé entre l'entrée et la sortie, c'est-à-dire que lorsque l'on applique 0 à l'entrée, on mesure V_{DD} à la sortie et à l'inverse, lorsque l'on applique V_{DD} à l'entrée, on mesure 0 à la sortie.

La Figure 2.46 et la Figure 2.47 présentent les caractéristiques de transfert en tension des options GeOI sur SOI et SOI sur SOI.

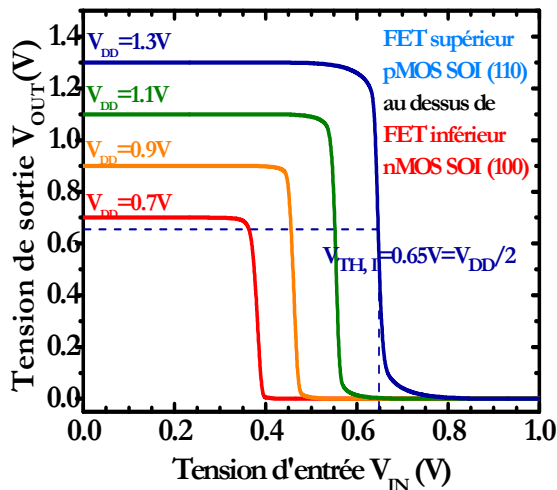


Figure 2.46 : Caractéristiques de transfert en tension d'un inverseur constitué d'un pMOS SOI empilé au dessus d'un nMOS SOI. ($L_P=2\mu\text{m}$, $W_P=10\mu\text{m}$; $L_N=2\mu\text{m}$, $W_N=5\mu\text{m}$)

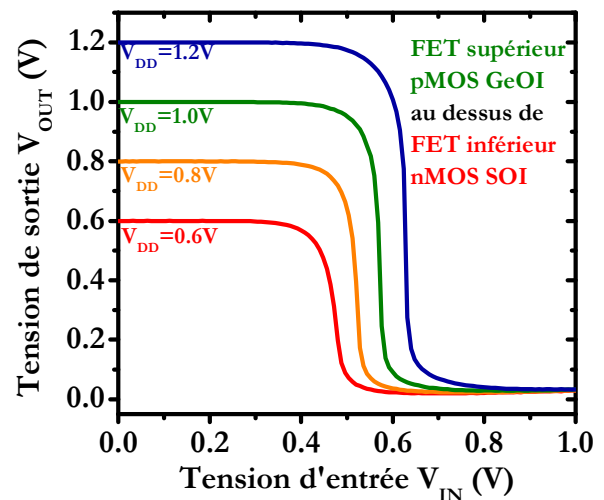


Figure 2.47 : Caractéristiques de transfert en tension d'un inverseur constitué d'un pMOS GeOI empilé au dessus d'un nMOS SOI. ($L_P=L_N=2\mu\text{m}$, $W_P=W_N=10\mu\text{m}$)

Ces caractéristiques montrent que les inverseurs sont fonctionnels. La tension de seuil de l'inverseur SOI/ SOI (i.e. la valeur de la tension de sortie quand la tension d'entrée vaut $V_{DD}/2$) est particulièrement bien équilibrée, même pour des faibles tensions d'alimentation. Equilibrer la tension de l'inverseur permet d'obtenir des temps de montée et de descente égaux. L'inverseur est équilibré lorsque les niveaux de courant des transistors nMOS et pMOS sont égaux pour une tension de grille valant $V_{DD}/2$.

A notre connaissance, ces cellules correspondent aux premières démonstrations de cellules en 3D avec zone active supérieure réalisée par collage moléculaire.

D. 2 - La SRAM 6T

La cellule SRAM à 6 transistors (SRAM 6T) est la cellule SRAM la plus utilisée dans les circuits intégrés. Actuellement, les plans mémoires SRAM occupent environ 60% de la surface de la puce. En conséquence, un gain en densité sur la cellule SRAM entraîne un gain de place sur une puce non négligeable. La cellule SRAM est une application phare de l'intégration 3D séquentielle qui permet de la rendre plus intéressante en termes de densité. Cet aspect sera détaillé dans le chapitre IV-Applications.

La Figure 2.48 présente le schéma électrique de la cellule SRAM 6T. Elle est composée de deux inverseurs tête bêche, ainsi que de deux transistors d'accès qui permettent d'accéder aux nœuds de stockage de l'information « L » et « R » de la cellule. La cellule est donc composée de 6 transistors allant par paire : les transistors d'accès notés TA, les transistors de charge notés TL (pour load en anglais), et les transistors de conduction notés TD (pour drive en anglais).

La Figure 2.49 présente le layout de la cellule réalisée. Les transistors nMOS de drive et d'accès ont été réalisés sur le niveau inférieur (SOI) et les transistors pMOS de load ont été réalisés sur le niveau supérieur. Dans le cas des transistors pMOS, les variantes GeOI et SOI ont été adressées.

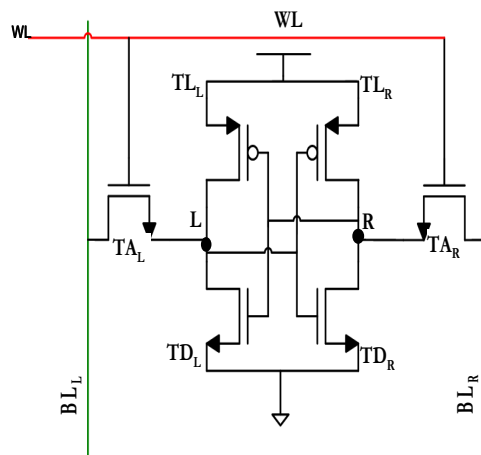


Figure 2.48: Schéma électrique d'une cellule SRAM à 6 transistors.

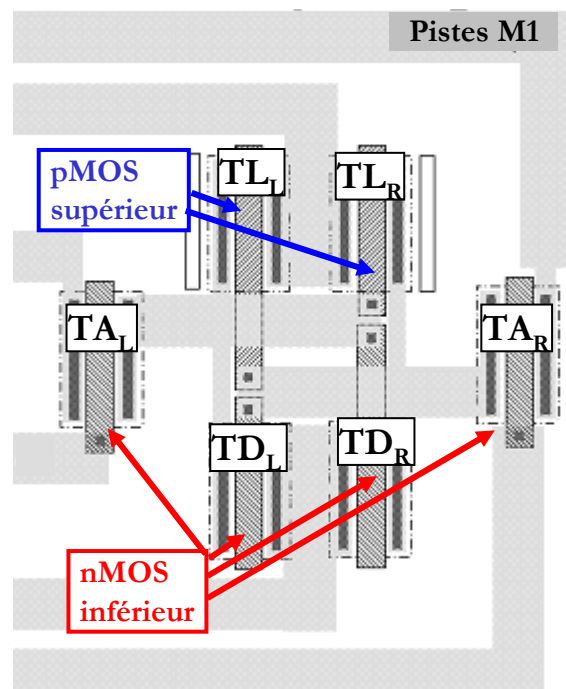


Figure 2.49: Dessin des masques des cellules SRAM testées.

En général, la fonctionnalité des cellules SRAM est démontrée par la courbe communément dénommée courbe « papillon ». Sur cette courbe sont présentées l'évolution du potentiel du nœud « L » lorsque la tension du nœud « R » passe de 0 à V_{DD} , ainsi que l'évolution du potentiel du nœud « R » lorsque le potentiel du nœud « L » passe de 0 à V_{DD} . (cf. Figure 2.50)

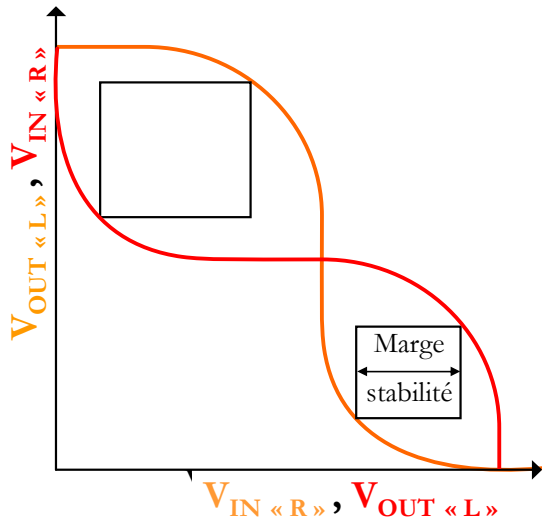


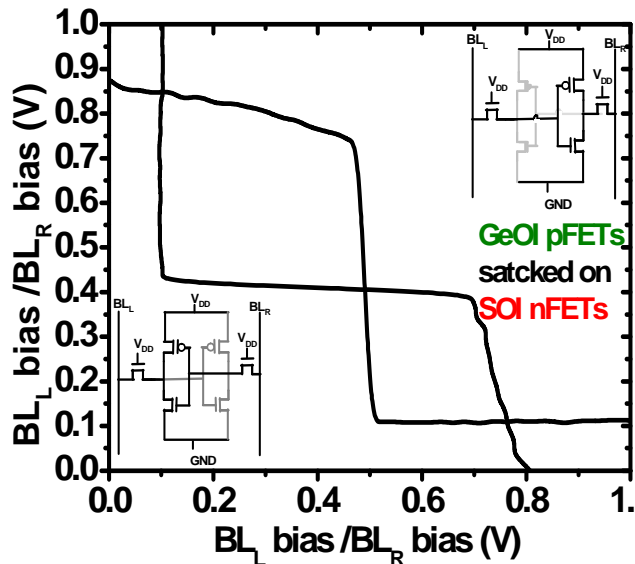
Figure 2.50 : Description de la courbe papillon

p sur GeOI empilés au dessus de transistors de type n sur SOI.

Cette caractéristique permet de vérifier la capacité de la cellule à maintenir l'information pendant qu'elle est lue. Dans notre cas, le dessin des masques (cf. Figure 2.49) ne nous permet pas de mesurer le potentiel aux nœuds internes «L et R».

La vérification du fonctionnement des transistors à été contrôlée en polarisant la BT_L (Ligne de Bit gauche) et en mesurant le potentiel dans Bl_R (Ligne de Bit gauche) et vice et versa.

La Figure 2.51 présente cette caractéristique pour des transistors de type


 Figure 2.51 : Left bit line (BL_L) potential vs. right bit line (BL_R) potential and vice versa for a 3D SRAM (stacked GeOI p-MOSFET over bottom SOI n-MOSFET).

passants les transistors d'accès en polarisant la ligne de mot (WL) à V_{DD} , et nous polarisons la ligne de bit gauche (BLL) à 0 et la ligne de bit droite à V_{DD} . Après une phase de rétention de l'information (passage de WL , BLL et BLL à 0) puis une phase de précharge des BL à 1, nous lisons la cellule en rendant passante la ligne de mot à V_{DD} . Nous mesurons alors le courant dans les lignes de bit : nous mesurons un courant fort ($2 \cdot 10^{-4}$ A) dans la BLL et un courant faible ($4 \cdot 10^{-11}$ A) dans la BLR . Ces valeurs de courant des Bit line montre que la cellule a bien mémorisé l'information et que celle ci n'est pas perdue pendant la phase de lecture.

Ensuite nous polarisons la BLR à des valeurs de tension décroissante V_{DD} , $3V_{DD}/4$, $V_{DD}/2$, $V_{DD}/4$ et 0. Nous observons que l'information contenue dans la cellule bascule puisque le niveau de courant dans la Bit line passe d'une valeur forte, à une valeur faible. Nous avons réalisé avec succès l'écriture du « 1 » (nœud $R=0$) dans la cellule.

Cette caractéristique ne permet pas de savoir si la cellule est stable en lecture, par contre elle apporte une information supplémentaire par rapport à une courbe papillon: elle démontre que les transistors d'accès sont bien, eux aussi, fonctionnels.

Pour vérifier la capacité de la cellule à mémoriser une information puis à être lue, nous avons réalisé le cycle de polarisation récapitulé dans la Figure 2.52. Durant l'ensemble du cycle, la cellule sera alimentée puisqu'il s'agit d'une mémoire statique, elle doit être alimentée pour retenir l'information.

Nous commençons par écrire un « 0 » dans la cellule, par convention cela signifie que le potentiel du nœud left vaudra 0. Pour cela nous rendons

Après une phase de rétention et de précharge des BL, nous lisons la cellule et nous observons que la cellule est bien stable en lecture lorsqu'un « 0 » est mémorisé puisque le niveau de courant est fort du côté R et faible du côté L.

En polarisant la BLL à des valeurs de tension décroissante V_{DD} , $3V_{DD}/4$, $V_{DD}/2$, $V_{DD}/4$ et 0, nous observons le basculement de l'information contenue dans la cellule. Dans chacune des phases les polarisations des WL, BLL et BLR sont appliquées pendant 70 ms, la mesure de courant sur les BL se faisant pendant les 20 dernières ms.

La caractéristique présentée dans la Figure 2.52 démontre que la cellule SRAM 3D réalisée est fonctionnelle en rétention, en lecture, et les potentiels des Bit line mesurés démontrent la capacité de la cellule à être écrite, lue et à retenir l'information entre ces différentes étapes. Il ne s'agit pas d'un résultat isolé, sur 40 structures testées, 19 sont fonctionnelles.

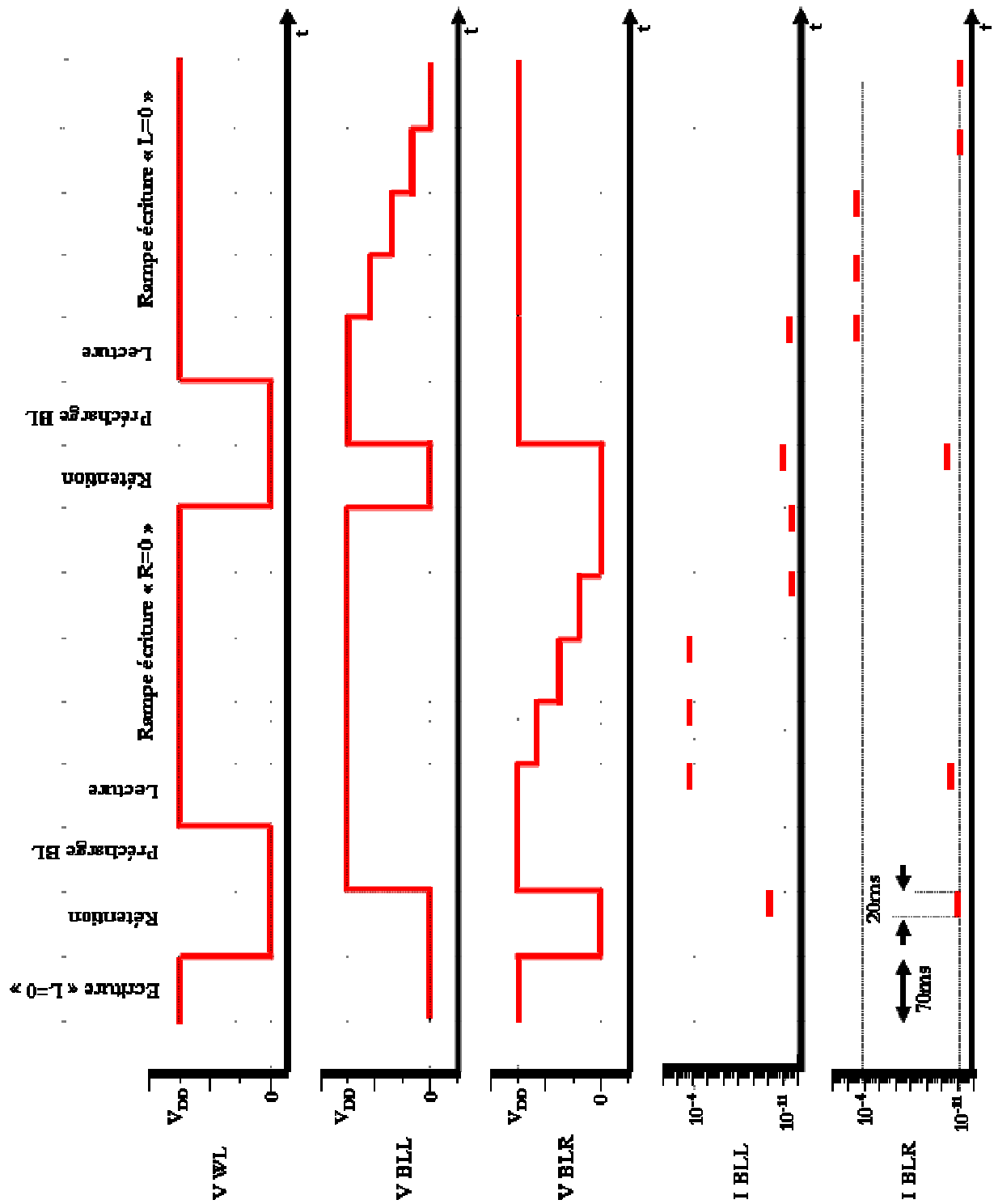


Figure 2.52 : Caractérisation de la cellule SRAM par un cycle composé de phases d'écritures, de lectures et de rétentions dans le cas de la mémorisation d'un « 1 » et d'un « 0 » dans la cellule.

Ce qu'il faut retenir

III-D : Démonstrateur de cellules en 3D

- L'ensemble des développements réalisés (développement d'un siliciure adaptée à l'intégration 3D, réalisation d'une zone active monocristalline à bas budget thermique, FET supérieur à bas budget thermique (Ge ou Si), développement d'interconnexions 3D), nous permettent de démontrer des cellules en trois dimensions.
- Nous démontrons des inverseurs et des SRAM 6T fonctionnelles.
- L'intégration 3D séquentielle nous permet de faciliter la cointégration de pFET GeOI et de nFET SOI.
- A notre connaissance, il s'agit de la première démonstration de cellules 3D avec une zone active supérieure réalisée par collage moléculaire.

CHAPITRE III

LES PERSPECTIVES DE GAINS DE L'INTEGRATION 3D SEQUENTIELLE

Nous l'avons vu, l'intégration 3D suscite actuellement beaucoup d'intérêt. Selon les références, le gain en densité varie de 10 à 60% environ lorsque l'on empile deux niveaux de transistors. Cependant la réalisation de cette intégration va entraîner un surcoût par plaque car le procédé front end est réalisé deux fois, que l'intégration utilise deux substrats ainsi qu'une étape de collage.

Dans cette partie, nous avons cherché à étudier conjointement les perspectives de gain (ou perte) en densité, performance et coût. Chacun des ces points fera l'objet d'une partie de ce chapitre.

Toutes ces questions sont extrêmement complexes. Cependant, nous verrons que certains arguments et études permettent de définir un faisceau de présomptions sur les perspectives de gain de l'intégration 3D séquentielle. Ce chapitre est entièrement consacré aux applications logiques pures. En particulier, les études concernant la réalisation de SRAM avec la technologie 3D séquentielle seront abordées dans le dernier chapitre.

SOMMAIRE DU CHAPITRE III :

LES PERSPECTIVES DE GAIN DE L'INTEGRATION 3D SEQUENTIELLE

A - ETUDE DU GAIN EN DENSITE.....	89
A. 1 - Etat de l'art	89
A. 2 - Méthodologie de conception	90
A. 3 - Mise au point du DRM de l'intégration 3D séquentielle	91
A. 4 - Réalisation de la bibliothèque de cellules standard.....	95
A. 5 - Synthèse logique.....	98
B - GAIN EN PERFORMANCES	100
B. 1 - Gain en performances sur les transistors unitaires	100
<i>B.1.1 - Intégration SOI / SOI.....</i>	<i>100</i>
<i>B.1.2 - Co-intégration SOI/ GeOI.....</i>	<i>103</i>
B. 2 - Gain en délai dans les interconnexions.....	103
C - ETUDE DE COUT	109
C. 1 - Présentation du cadre de l'étude.....	109
C. 2 - Résultats	111
D - CONCLUSION GENERALE SUR LES PERSPECTIVES DE GAIN DE L'INTEGRATION 3D SEQUENTIELLE	115

A - Etude du gain en densité

A. 1 -Etat de l'art

Le tableau 3.1 récapitule les études de densité menées dans le cas d'intégrations 3D séquentielles. Dans ce tableau, nous ne citerons pas les études où seul le gain sur de petites cellules standard (<10 transistors) est quantifié. Nous ne récapitulons que les études réalisées sur des cellules relativement complexes (~1000 transistors), plus représentatives des perspectives de gain en densité sur un circuit complexe.

Groupe [Référence]	Année Publication	L_G	Cellules standard	Cellules complexes	Niveaux d'actives	Niveaux de métaux entre active	Configuration	Gain cellule standard	Gain cellule complexe (~1000transistors)
Stuttgart [Roos92]	1992	2.5 μ	NOT, NAND2, NOR2, MUX2	Multiplieur 8X8	2	0	pMOS double grille sur NMOS		46%-65%
TIMA [Abou Samra98]	1998	0.1 μ m	NVERSEUR, NAND2, AND2, NOR2, OR2, AOI3, XOR2, MAJ	Multiplieur 16X16 bits	2	0	pMOS au dessus de nMOS	30 à 40%	13%
Intel +Université de Purdue [Zhang99] [Wei02]	1999	0.35 μ		Multiplieur 4X4 bits	2 ou 4	1,2	N/P 2niveaux		~50%
							N/P 4niveaux		~75%
							CMOS2niveaux		~39%
							CMOS4niveaux		~70%

Tableau 3.1 : Récapitulation des études de gain en densité menées avec des technologies 3D séquentielles.

Nous observons que les valeurs de gain en densité estimées par ces références sont très divergentes. De 13% [Abou Samara98] à 65% [Roos92] pour des architectures avec deux niveaux empilés. Ces divergences peuvent s'expliquer principalement par trois points :

1/ Tout d'abord par les règles de dessin très variables utilisées pour réaliser ces études. Les longueurs de grille varient entre 2.5 μ et 0.1 μ m pour l'étude la plus avancée. Selon le nœud technologique pour lequel l'étude a été menée, l'influence de l'empilement va être nécessairement modifiée. Dans les années 80 ($L_G \sim \mu$ m), la surface occupée par les zones actives était largement prédominante par rapport à celle occupée par les contacts. Cette tendance s'est inversée pour les nœuds plus avancés (fin des années 90 $L_G=0.1\mu$), où c'est la densité des contacts qui est le facteur limitant.

2/ Aucune information n'est donnée sur la définition des règles de dessin. Comment ont-elles été générées ? Prennent-elles en compte les particularités de l'intégration 3D telles que les nouvelles stratégies d'alignement ainsi que les rapports de forme des contacts 3D ?

3/ Enfin, certaines études sont menées sur des architectures différentes de celle que nous étudions ; avec par exemple, un scénario où des niveaux de lignes métalliques sont présents entre deux niveaux de transistors¹, ou un scénario avec quatre niveaux d'actives empilées [Zhang99].

En conséquence, aucune des études citées ne nous permet de conclure sur les perspectives de gain en densité avec la technologie développée au Leti (deux niveaux de transistors empilés sans niveau métallique entre les deux) pour les prochaines générations technologiques (45 et sub-45 nm). De plus, il nous est nécessaire de connaître le mode de définition des règles de dessin pour connaître le degré de confiance que nous pouvons ensuite placer dans le résultat des études de densité.

Dans l'étude de densité proposée, nous chercherons à comparer deux schémas d'intégration. Un premier schéma, nommé intégration N/P, où des transistors de type n sont empilés au dessus de transistors de type p. Un deuxième schéma, nommé intégration CMOS/CMOS où les deux types de transistors n et p peuvent être dessinés sur chaque niveau.

A. 2 -Méthodologie de conception

La Figure 3.1 présente l'enchaînement des étapes de conception d'un circuit. Pour réaliser un circuit complexe, les concepteurs ont tout d'abord besoin d'une bibliothèque de cellules standard. Elle comporte des fonctions logiques de base (e.g. NOT, AND, OR, NAND, NOR) et des fonctions de mémorisation (e.g : FLIP-FLOP) à partir desquelles sont construites des fonctions plus complexes (étape de synthèse logique).

L'étape de synthèse physique consiste ensuite à placer physiquement ces différentes cellules et à les router (c'est-à-dire à les interconnecter).

L'utilisation de bibliothèques de cellules standard permet de diminuer les temps de conception (cellules réutilisables) mais, en contrepartie, le circuit final aura une surface plus importante qu'en méthodologie de conception « *full custom* » où toutes les fonctions logiques ne sont pas implémentées dans les bibliothèques de cellules standard.

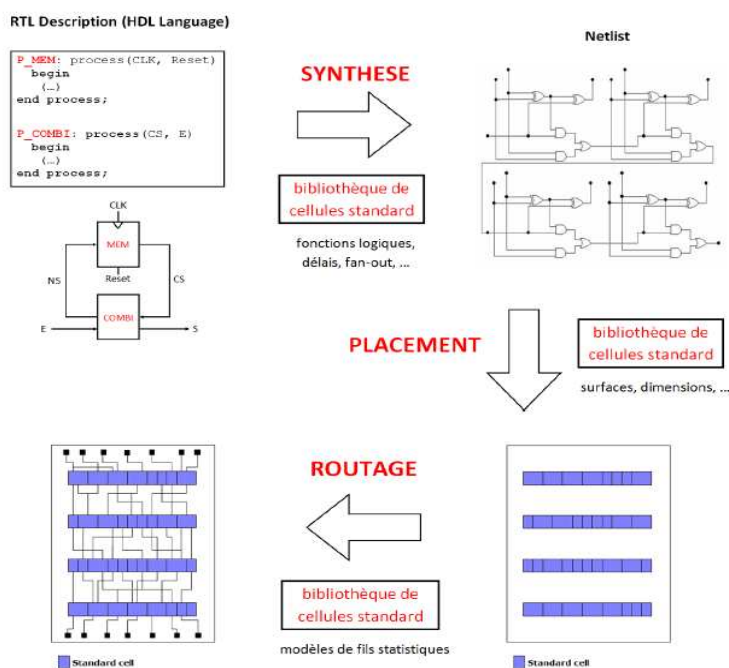


Figure 3.1 : Place d'une bibliothèque de cellule standard dans le flot de conception.

¹ Cette étude est purement théorique, il n'y a pas de démonstrateur réalisé avec des niveaux de métaux entre les niveaux de transistors empilés.

A. 3 - Mise au point du DRM de l'intégration 3D séquentielle

Le Design Rule Manual (DRM) est le document contenant toutes les règles de dessin d'une technologie devant être respectées dans le dessin des masques (layout) d'un circuit intégré. Ce jeu de règles est donné par les technologues qui doivent déterminer les dimensions minimales réalisables dans la nouvelle technologie.

En général, les règles du DRM sont obtenues, dans un premier temps, par extrapolation du jeu de règles développé pour le nœud antérieur. Selon la loi de Moore, les dimensions sont réduites d'un facteur racine de 2 à chaque génération. Dans un deuxième temps, des démonstrations de circuits avec la nouvelle technologie permettent d'affiner ces règles pour optimiser rendement, densité et performance. En conclusion, un jeu de règles se base sur l'expérience développée au nœud antérieur et sur les premières réalisations avec les dimensions du nœud plus avancé.

Pour le développement d'une nouvelle technologie, telle que la technologie 3D séquentielle, il faut donc faire des hypothèses pour pallier l'absence de données et identifier les paramètres critiques principaux pour élaborer de nouvelles règles de dessin. Dans ce paragraphe, nous allons décrire comment nous avons extrapolé le DRM de la technologie planaire 45 nm FDSOI pour l'adapter à la technologie 3D séquentielle.

Règle de distance entre deux niveaux :

De manière générale, une règle de distance entre un élément A et un élément B de niveaux de masques distincts est composée de deux termes, un terme μ qui correspond à une valeur fixe et un deuxième terme σ correspondant à un écart type. La Règle de Distance (RD) est alors la somme de ces deux termes : $RD = \mu + \sigma$.

La variance totale (σ) de la règle fait intervenir principalement deux paramètres : les performances d'alignement (overlay entre deux niveaux lithographiques : σ_{OV}) et le contrôle de dimension qui peut être obtenu après photo-lithographie et gravure (σ_{COTE}). En considérant ces deux écarts type comme des écarts types de lois gaussiennes d'événement indépendants, la variance totale est donnée par l'expression :

$$\sigma = \sqrt{(\sigma_{OV})^2 + (\sigma_{COTE})^2} \quad (\text{Eq 3.1})$$

La valeur fixe μ est basée sur l'expérience de la technologie et tient compte de nombreux paramètres. Par exemple, elle prend en compte les distances nécessaires à la limitation du couplage entre les deux niveaux concernés par la règle. Elle sera extraite du DRM planaire.

Pour comprendre le raisonnement suivi pour établir de nouvelles règles, prenons l'exemple de la règle « distance du contact sur grille par rapport à active » dans le DRM planaire et la même règle mais pour le niveau supérieur dans le DRM 3D. Les Figure 3.2 et Figure 3.3 présentent ces règles de dessin.

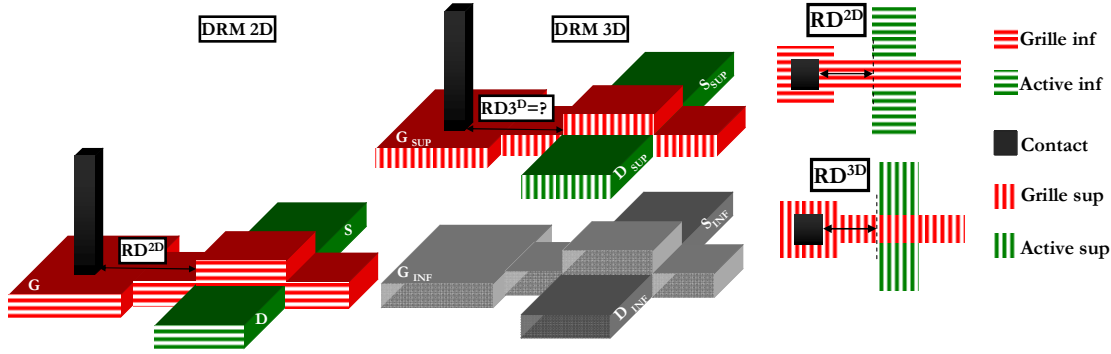


Figure 3.2 : Description de la règle « distance contact sur grille par rapport à active » cas du DRM planaire et de la règle pour le niveau supérieur du DRM 3D.

Figure 3.3 : Vue layout des règles représentées dans la Figure 3.2

Il nous apparaît clairement que la partie fixe μ de la règle de dessin doit rester constante pour ces deux règles puisque cette règle concerne le positionnement du niveau contact par rapport au niveau grille donc $\mu_{2D} = \mu_{3D}$. Par contre au niveau de la partie statistique, la partie σ_{OV} va être modifiée car la stratégie d'alignement diffère par rapport à la 2D (cf Figure 3.4).

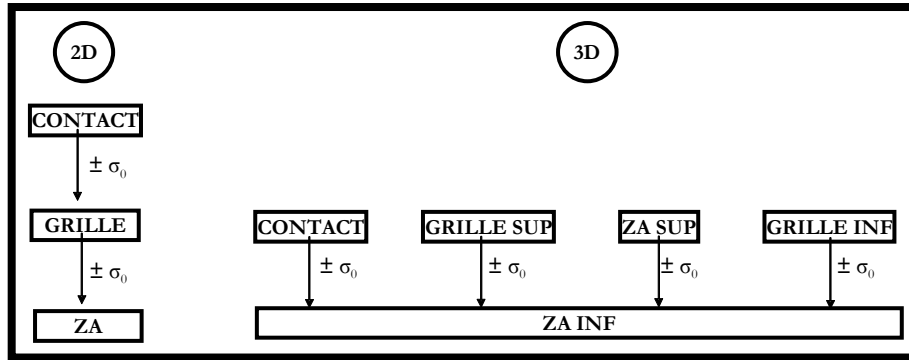


Figure 3.4 : Description de la stratégie d'alignement en 2D et en 3D

En effet, dans l'intégration planaire, le niveau contact est aligné sur la grille, qui lui-même est aligné sur la zone active, alors que dans le cas 3D, l'alignement de chaque niveau front end se fait

par rapport à l'active inférieure. Les performances d'alignement d'un niveau lithographique quelconque sur une référence sont considérées constantes et la variance est notée σ_0 . On en déduit les variances dans le cas planaire et le cas 3D :

Calcul de la variance σ_{2D}

$$\sigma_{2D} = \sqrt{(\sigma_{ov,2D})^2 + (\sigma_{COTE,ACT})^2 + (\sigma_{COTE,CONT})^2}$$

$$\sigma_{ov,2D} = \sqrt{(\sigma_{ov,PO \rightarrow ZA} + \sigma_{ov,CO \rightarrow PO})^2}$$

$$\sigma_{ov,2D} = \sqrt{(2\sigma_0)^2}$$

(overlay dépendants)

Calcul de la variance σ_{3D}

$$\sigma_{3D} = \sqrt{(\sigma_{ov,3D})^2 + (\sigma_{COTE,ACT})^2 + (\sigma_{COTE,CONT})^2}$$

$$\sigma_{ov,3D} = \sqrt{(\sigma_{ov,PO-SUP \rightarrow ZA-INF})^2 + (\sigma_{ov,ZA-SUP \rightarrow ZA-INF})^2 + (\sigma_{ov,CO \rightarrow ZA-INF})^2}$$

$$\sigma_{ov,3D} = \sqrt{3 * (\sigma_0)^2}$$

(overlay indépendants)

La valeur de la règle 3D est alors obtenue en appliquant les expressions suivantes.

$$RD_{3D} = (RD_{2D} - \sigma_{2D}) + \sigma_{3D}$$

$$RD_{3D} = \mu + \sigma_{3D}$$

Les valeurs des performances d'alignement en fonction des nœuds technologiques sont issues de l'ITRS et/ou de l'expérience interne. La partie variable σ_0 est prise égale à 20 nm pour

le nœud 65nm et 13 nm pour le nœud 45 nm et le $\sigma_{COTE} = 10\%$ de la dimension minimale du niveau dans le DRM planaire.

Nous avons vérifié la pertinence de ce raisonnement et des valeurs numériques considérées en extrapolant le DRM planaire 65 nm et nous l'avons comparé au DRM 45 nm. L'écart entre les règles obtenues et les règles du DRM 45 nm sont autour de 10%.

Règles de garde d'un contact

Pour s'assurer qu'un contact débouchera bien sur le niveau à contacter en cas de désalignement, une garde est dessinée autour du contact.

Pour la réalisation d'un contact tombant sur la couche supérieure et lorsque l'alignement se fait par rapport à l'active inférieure, la règle de design va faire intervenir les règles planaires usuelles et un alignement supplémentaire entre la couche supérieure et l'active inférieure. La garde 3D (G^{3D}) sera la somme de la garde planaire G^{2D} et du désalignement $G^{3D} = G^{2D} + \sigma_0$ (cf Figure 3.5.)

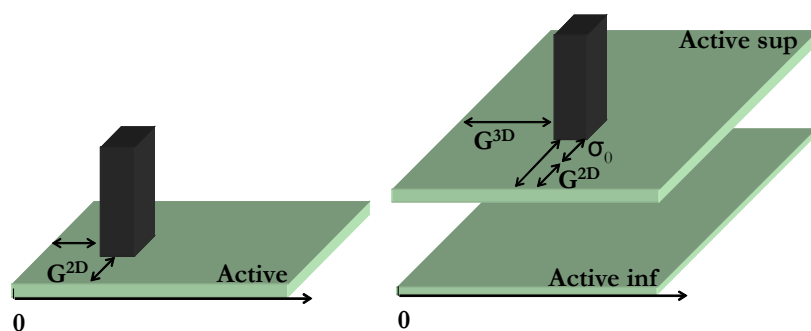


Figure 3.5 : Garde contact sur active en planaire et garde sur active supérieur dans le DRM 3D

Dimensionnement du contact 3D

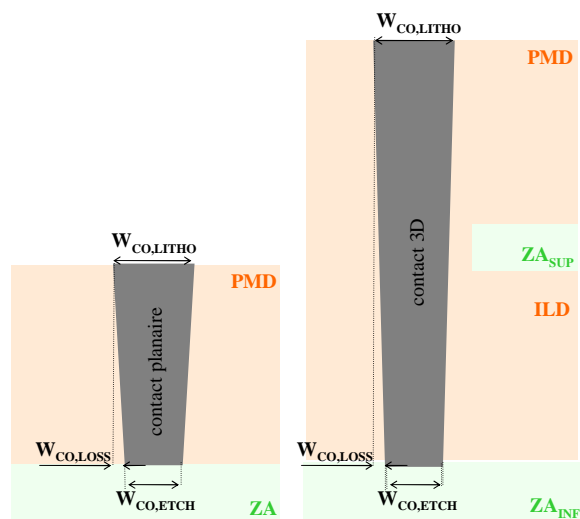


Figure 3.6 : Dimensionnement du contact sur niveau inférieur.

sur le niveau inférieur sera donc pris égal à la taille du contact en planaire W_{CO} à laquelle nous rajouterons le $\Delta_{CO, ETCH}$ supplémentaire pour tenir compte de la gravure du diélectrique inter-niveau.

Lors de la gravure des contacts, une perte de cote $\Delta_{CO, ETCH}$ apparaît entre le haut et le bas du contact (cf Figure 3.6). Afin que la surface de contact soit la même sur niveau haut et bas, nous avons redimensionné le contact sur niveau inférieur en tenant compte de l'épaisseur de diélectrique supplémentaire à graver. Nous nous sommes placés dans le cas, où l'épaisseur de diélectrique à graver pour atteindre le niveau inférieur est le double de l'épaisseur pour atteindre le niveau supérieur. On considérera alors que la perte de cote d'un contact 3D est le double de celle obtenue pour une structure planaire (ce qui est une hypothèse technologique raisonnable voire pessimiste) Pour le nœud 45nm, la perte de cote $\Delta_{CO, ETCH}$ vaut 7 nm, la taille du contact 3D atterrissant

La récapitulation des tailles de tous les contacts décrits dans la partie contact 3D du chapitre brique technologique est donnée dans la Figure 3.7:

Type contact	Contact sélectif supérieur	Contact sélectif inférieur	Contact sélectif chevauchant	Contact traversant	Contact interne
Schéma					
Taille contact	60x60 nm ²	67x67 nm ²	160x67 nm ²	67x67 nm ²	60x60 nm ²

Figure 3.7: Description des contacts et dimensionnements retenus dans le DRM 3D 45nm.

Nous observons que la taille du contact chevauchant correspond à plus du double du contact traversant. Pour des applications très hautes densités ce type de contact est donc inadapté. Cependant son utilisation permet de réaliser l'ensemble des contacts en une seule lithographie, rendant cette option moins couteuse. Il pourra être utilisé dans des applications où le critère de densité n'est pas essentiel.

Le DRM 3D 45nm ainsi défini comporte plus d'une centaine de règles régissant les règles de dimensionnements des différentes couches représentées schématiquement dans la Figure 3.8.

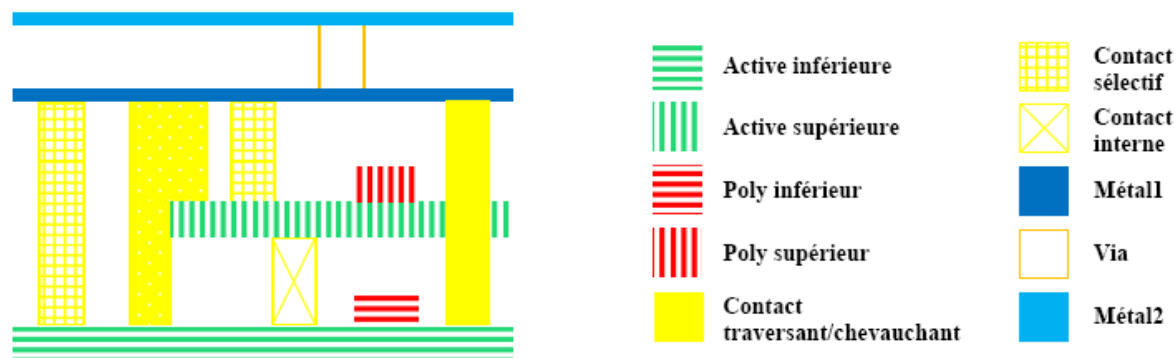


Figure 3.8 : Couches disponibles dans le design Kit.

A. 4 - Réalisation de la bibliothèque de cellules standard

La réalisation des bibliothèques de cellules en 2D et 3D a été réalisée par Stéphane Michaud pendant son stage de 2^{ème} année. Il a été encadré par Alexandre Valentian, concepteur dans l'équipe de conception « proche technologie » du Leti et moi même.

Pour mener l'étude de densité comparative 2D/3D, deux bibliothèques de cellules standard ont été développées : une bibliothèque 2D, une bibliothèque 3D en configuration nMOS sur pMOS (N/P). La bibliothèque comporte les portes suivantes : INVERSEUR, NOR2, NAND2, BUFFER 3 ETATS, LATCH et FLIP FLOP. Chaque layout de cellule a été décliné en plusieurs versions pour alimenter à sa sortie de 1 à 16 portes identiques en parallèle. (On parle de fan out noté $(X)_i$ allant de 1 à 16).

Les cellules d'une bibliothèque standard doivent partager des caractéristiques communes. En particulier leurs hauteurs et leurs largeurs sont des multiples d'une même grandeur appelée « pas » de la cellule (cf Figure 3.9). En général, ce pas est égal au pas (ou « pitch ») du niveau de métal utilisé pour le routage global. La hauteur des cellules est fixe, c'est la première grandeur à déterminer lorsque l'on développe une bibliothèque de cellules standard. La largeur quant à elle est variable et dépend de la complexité de la cellule à réaliser et de la taille des transistors. En technologie planaire, la surface de la cellule doit être divisée en deux parties : zone d'implantation des transistors de type n et zone d'implantation des transistors de type p. Le rapport entre ces deux surfaces sera choisi comme étant égal au rapport entre les largeurs du nMOS et du pMOS permettant d'avoir des courants identiques dans les deux transistors. Les rails d'alimentation V_{dd} et Gnd (Métal1) sont finalement placés respectivement en haut et en bas de la cellule. On définit ainsi un gabarit commun à toutes les cellules standard. L'outil de placement, une fois la synthèse logique effectuée, peut alors générer des lignes de cellules, chacune aboutée avec ses voisines, pour former le layout du circuit final (cf Figure 3.10).

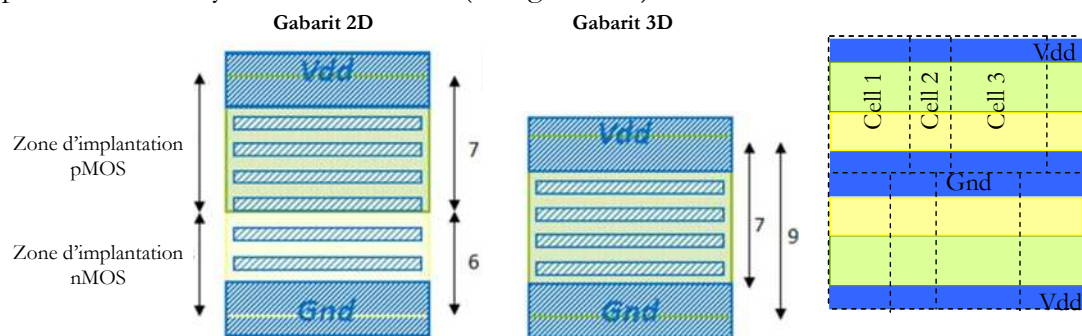


Figure 3.9: Dimensionnements choisis du gabarit de cellule en 2D et en 3D N/P. La hauteur des cellules est exprimée en nombre de pas, le pas valant $0.98 \mu\text{m}$.

Figure 3.10 : Utilisation du gabarit pour abouter les cellules des bibliothèques et former des circuits intégrés complexes

Le gabarit de la cellule 3D est choisi plus petit que celui de la cellule planaire, en conséquence la surface disponible pour le routage des cellules est plus faible. Il s'agit en premier lieu de vérifier que l'espace disponible entre les deux rails d'alimentation est suffisant pour pouvoir router la cellule la plus complexe de la bibliothèque : la cellule FLIP-FLOP. Le layout de cette cellule est présenté en Figure 3.11 et le gabarit choisi permet bien de router la cellule. Cependant il faut

noter que par rapport à la cellule planaire, la longueur du routage en polysilicium dans la cellule 3D est plus importante (de 12%). Cette différence devra être évaluée électriquement, cependant les outils nécessaires n'ont pas été développés au cours de cette thèse.

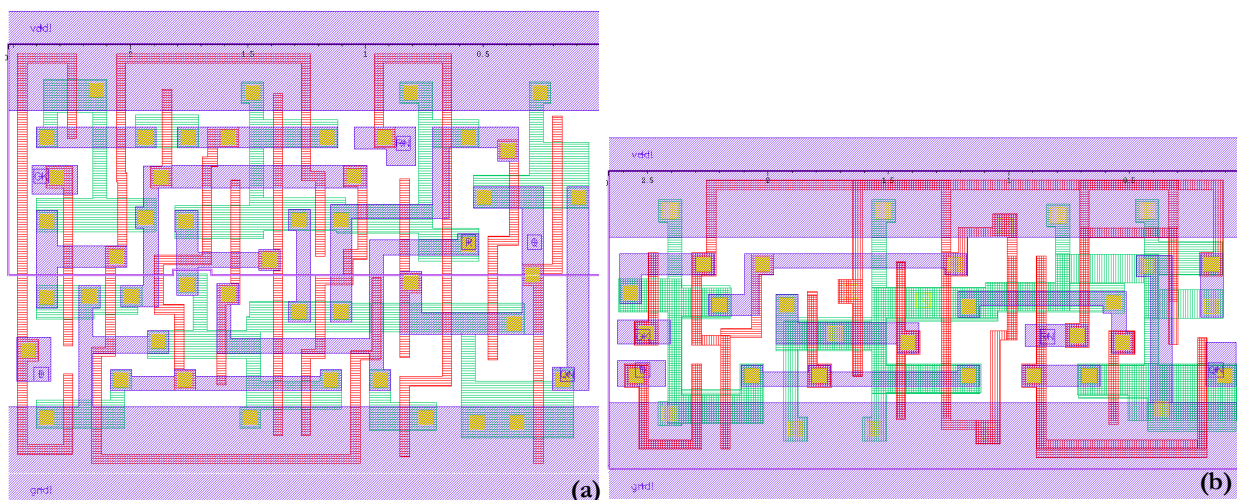


Figure 3.11 : Layout de la cellule FLIP FLOP (a) en technologie planaire et (b) avec l'intégration 3D séquentielle.

Dans la partie logique d'un circuit, les transistors seront toujours dimensionnés avec la longueur de grille minimum permise par la technologie. La largeur W des transistors sera toujours orientée dans le sens de la hauteur de la cellule. Lorsque le W dépasse la hauteur de la zone d'implantation, on augmente le nombre de doigts n du transistor (cf Figure 3.12), ce qui consiste à mettre en parallèle des transistors de $W' = W/n$.

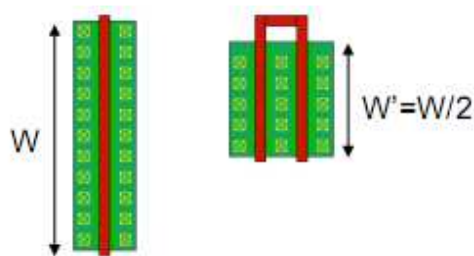


Figure 3.12 : Transistor à deux doigts.

Les largeurs des transistors dans chaque cellule seront dimensionnées à l'aide de simulation utilisant le modèle compact FDSOI développé au Leti pour le nœud 45nm. Ce modèle compact est développé pour des transistors avec actives silicium uniquement. La possibilité d'améliorer les performances des transistors en réalisant par exemple les transistors pMOSFET sur Germanium ou/et en introduisant des contraintes ou encore en modifiant les orientations cristallines des différents niveaux n'a pas été prise en compte.

Les gains en surface entre les cellules de la bibliothèque planaire et de la bibliothèque 3D N/P sont récapitulés dans la table 3.2.

Porte	Fan out Xi	Gain en surface (%)
Inverseur	1	30.8
	2	30.8
	4	30.8
	8	30.8
	16	30.8
NAND 2	1	44.6
	2	30.8
	4	30.8
	8	13.5
	16	1.1
NOR 2	1	44.6
	2	30.8
	4	30.8
	8	13.5
	16	23.1
BUFFER 3 ETATS	5	30.8
	10	30.8
	15	30.8
LATCH	4	30.8
FLIPFLOP	4	26.9

Table 3.2 : Gain en densité avec la bibliothèque 3D N/P

Porte	Xi	Gain en densité	
		N/P	CMOS / CMOS
NAND 2	1	44.6	58.5
	2	30.8	48
	4	30.8	48
	8	13.5	30.8
	16	1.1	30.8
NOR 2	1	44.6	58.5
	2	30.8	48
	4	30.8	30.8
	8	13.5	30.8
	16	23.1	46.2

Table 3.3: Gain en densité avec la bibliothèque 3D CMOS/ CMOS.

réalisation de layout des portes buffer 3 états, Latch et FLIPFLOP en CMOS/CMOS pourra conduire à des gains plus importants pour cette option.

En moyenne, sur l'ensemble des cellules de cette bibliothèque nous obtenons un gain en densité de 28,4%. Notons que les gains pour les portes NAND2 ET NOR 2 dans les versions X8 et X16 sont très faibles. Cette différence est due notamment au fait que pour ces versions de cellules, la configuration 3D passe en multi-doigt alors qu'elle reste en mono-doigts en planaire. Le gabarit de la cellule n'est pas favorable pour un gain en densité sur ces cellules. Cependant le gabarit doit rester constant dans tout le layout, puisqu'il permet ensuite le positionnement régulier des cellules dans un circuit complexe.

Pour augmenter le gain en densité apporté par la 3D sur ces cellules, nous nous autorisons à réaliser des n et des pMOS sur chacun des niveaux (configuration CMOS/CMOS). La table 3.3 présente les résultats obtenus.

Nous observons que la configuration CMOS/ CMOS est plus favorable au gain en densité sur ces portes. La bibliothèque hybride comportant les versions CMOS/CMOS pour les portes NAND2 et NOR2 et les autres portes de la bibliothèque N/P conduit à un gain moyen en densité de 36,7%. Nous voyons que l'architecture 3D CMOS/CMOS, en apportant un degré de liberté supplémentaire au niveau du layout, est bénéfique au niveau gain en densité.

Notons que pour la bibliothèque CMOS/CMOS, seule les portes NAND2 et NOR2 utilisent ce degré de liberté supplémentaire apporté par la possibilité de faire du CMOS sur chaque niveau. La

A. 5 -Synthèse logique

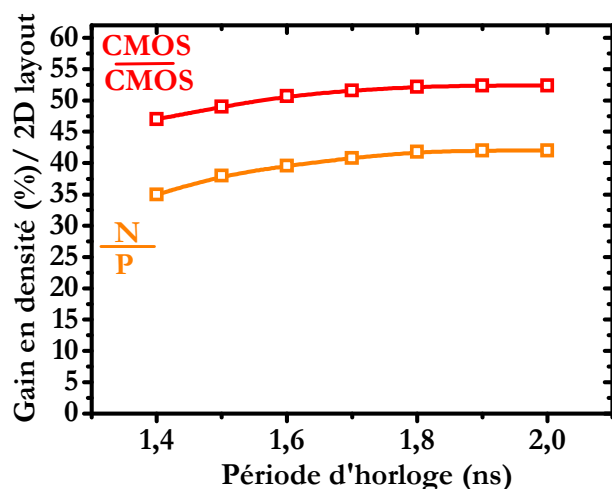


Figure 3.13 : Evolution du gain en densité en fonction de la fréquence d'horloge.

le cas 3D CMOS/ CMOS respectivement.

L'étude de gain en densité réalisée, basée sur la comparaison des surfaces moyennes de cellules standard des bibliothèques permet une première estimation des perspectives de gains en densité. Cependant, dans un circuit complexe, les cellules dessinées ne sont pas toutes utilisées avec la même fréquence. Pour obtenir une perspective du gain en densité plus proche de celle obtenue dans un circuit, une synthèse logique d'une cellule 16 bits Multiply ACumulate gate (MAC) a été réalisée. Cette cellule comporte environ 1000 transistors. La Figure 3.13 présente les résultats de la synthèse logique. Le gain en densité est aux alentours de 40% et de 50% dans le cas de l'intégration 3D N/P et dans

Le gain additionnel apporté par le degré de liberté supplémentaire dans un layout 3D CMOS/CMOS par rapport à un layout 3D N/P est donc confirmé dans une cellule de plus grande taille, plus représentative de la perspective de gain en densité dans un circuit complexe.

Bien que l'étape de synthèse logique donne une perspective de gain plus réaliste qu'une simple moyenne du gain en densité des cellules d'une bibliothèque, ce résultat reste partiel. En effet, cette étape consiste à définir les associations de cellule standard pour réaliser le circuit complexe. La surface est calculée en faisant la somme des surfaces des cellules standard utilisées et les interconnexions sont simulées en utilisant des logiciels de routage aléatoire.

Pour avoir une information totalement fiable sur le gain en densité sur la cellule MAC 16 bits, il faudrait aller jusqu'à l'étape de placement routage. Il est possible lors de cette étape, que la surface de routage s'avère insuffisante, on parle alors de congestion du routage. Si tel était le cas, le gabarit des cellules standard devrait être modifié.

Cette étude quantitative préliminaire nous permet donc de penser qu'un gain en densité de 40 à 50 % sur des cellules relativement complexes peut être obtenu. Cependant la situation est nettement moins claire pour un circuit intégré complet avec ses nombreux niveaux de métaux permettant d'interconnecter tous les transistors entre eux. Si le nombre de transistors par unité de surface est doublé, le nombre d'interconnexions nécessaires pour les relier va obligatoirement augmenter.

En conclusion, il semble possible de gagner un nœud en densité au niveau du Front end sans modifier les règles technologiques (densité du 32 nm avec une technologie 45 nm), mais il est impossible de réussir à connecter toutes ces cellules avec la même technologie 45 nm sans augmenter de manière drastique le nombre de niveaux de métaux. Pour que le gain en densité observé jusqu'au Metal 1 puisse se retrouver au niveau du circuit, il semblerait nécessaire de

réduire les dimensions du Back end. L'intégration 3D permettrait alors de relâcher les contraintes de la réduction des dimensions au niveau front end ; une réduction des dimensions du back end semble *a priori* nécessaire.

Dans l'optique d'une congestion du routage avec le gabarit choisi pour les cellules 3D, nous avons suivi le raisonnement inverse: si on fixe une surface constante de la cellule en 2D et 3D, quel gain en courant peut être obtenu ? Les résultats pour une cellule NAND X8 sont présentés dans la table 4.4 :

NAND X8	3D N/P	3D CMOS/CMOS
Gain en courant 3D/ 2D	20%	100%

Table 4.4 : Gain en courant en passant en 3D pour une porte NAND X8 avec une surface de cellule constante.

Le gain en courant à surface constante est de 20% dans le cas 3D N/P et 100% dans le cas 3D CMOS/CMOS.

Là encore, l'intégration de type CMOS/CMOS apporte un bénéfice supérieur en terme de densité à l'intégration N/P. Cependant, les simulations effectuées sont basées sur des modèles compacts FDSOI avec du silicium (100). L'apport en performance permis par l'optimisation indépendante des deux transistors empilés (via le choix des matériaux de canal et l'orientation) n'a pas été pris en compte et pourrait éventuellement apporter une nuance à ces résultats.

CMOS/CMOS apporte un

Ce qu'il faut retenir
IV-A : Gain en densité

- Les résultats de la littérature sur le gain en densité obtenus grâce à l'intégration 3D séquentielle sont extrêmement variables puisque donnés pour des nœuds technologiques variables et/ ou des architectures différentes (présence de ligne de métaux entre les deux niveaux de transistors ou réalisation de cellules sur trois niveaux) et/ ou des cellules trop simples, sans rapport avec des circuits complexes.
- Pour obtenir une perspective du gain en densité fiable, adaptée à des nœuds technologiques avancés et à la technologie étudiée, nous avons suivi la démarche suivante :
 - Création d'un manuel de règles de dessin en prenant en compte des paramètres technologiques réalistes et basés sur des hypothèses connues.
 - Réalisation de bibliothèques de cellules standard avec définition d'un gabarit commun
 - Réalisation d'une synthèse logique sur une cellule plus complexe (~1000 transistors) permettant de prendre en compte la fréquence d'apparition de telle ou telle cellule de la bibliothèque.
- Le DRM réalisé est basé sur le nœud technologique 45nm et prend en compte entre autre la stratégie d'alignement différente en technologie planaire et en 3D, ainsi que la perte de cote dans les contacts.
- Deux bibliothèques de cellules (2D et 3D option nMOS empilés sur pMOS) ont été créées comportant les cellules standard indispensables à la réalisation d'un circuit complexe (NOT, NAND2, NOR2, BUFFER 3 ETATS, LATCH, FLIP-FLOP), elles ont été dessinées parfois jusqu'à la versions X16 qui délivre 16 fois plus de courant que la version de taille minimale X1.

- La première estimation du gain en densité obtenue en moyennant le gain sur toutes les cellules d'une bibliothèque donne un gain de 28.9% et de 36.7% pour les bibliothèques N/P et CMOS/ CMOS respectivement.
- Afin de tenir compte de la fréquence d'utilisation des cellules standard dans un circuit plus complexe, une synthèse logique d'une cellule 16 bit MutiPLY ACumulator (1000 transistors) a été réalisée. Le gain en densité est de 40 et 50% environ dans les cas N/P et CMOS/CMOS respectivement.
- L'architecture CMOS/CMOS apparaît plus intéressante en termes de densité. Elle permet un degré de liberté supplémentaire lors du dessin des cellules. Cependant cet avantage peut être nuancé par l'apport en performance apporté par l'optimisation indépendante des nMOS et de PMOS dans l'architecture N/P non prise en compte dans les simulations.
- Cette étude quantitative préliminaire laisse penser qu'un gain d'un nœud en densité au niveau du Front end sans modifier les règles technologiques (densité du 32 nm avec une technologie 45 nm) est possible. Cependant pour connecter toutes ces cellules avec la même technologie 45 nm, il serait nécessaire d'augmenter le nombre de niveaux de métaux. Pour que le gain en densité observé jusqu'au M1 puisse se retrouver au niveau du circuit, il semble nécessaire de réduire les dimensions du back end.

B - Gain en performances

B.1 -Gain en performances sur les transistors unitaires

L'architecture 3D séquentielle se prête facilement à l'optimisation indépendante des performances de n et pMOS grâce à la possibilité de réaliser des interconnexions à l'échelle du transistor, permettant de dessiner la cellule CMOS à cheval sur deux niveaux.

L'optimisation indépendante des deux types de transistors peut passer par le choix des empilements de grilles, des contraintes, des orientations et des matériaux de canal.

Comparée aux techniques de type « seed window », l'utilisation du collage moléculaire, donne un degré supplémentaire dans l'optimisation des performances des transistors. L'apport de contrainte peut être réalisé aussi via le substrat de départ et non seulement par le procédé de fabrication. De plus l'orientation des substrats peut être choisie indépendamment sur les deux niveaux empilés.

B.1.1 - Intégration SOI /SOI

Dans le cas d'une intégration SOI sur SOI, une option prometteuse est la co-intégration de l'orientation (100) pour le nMOS et (110) pour le pMOS.²

Des gains en mobilité de plus d'un facteur 2 ont été démontrés en passant d'une orientation (100) à l'orientation de surface (110) dans la direction (110) (Pic de mobilité à $225 \text{ cm}^2 \cdot (\text{V} \cdot \text{s})^{-1}$). [Yang 06a]. Pour réaliser cette co-intégration, plusieurs schémas d'intégration ont été proposés et réalisés et sont résumés dans la Figure 3.14.

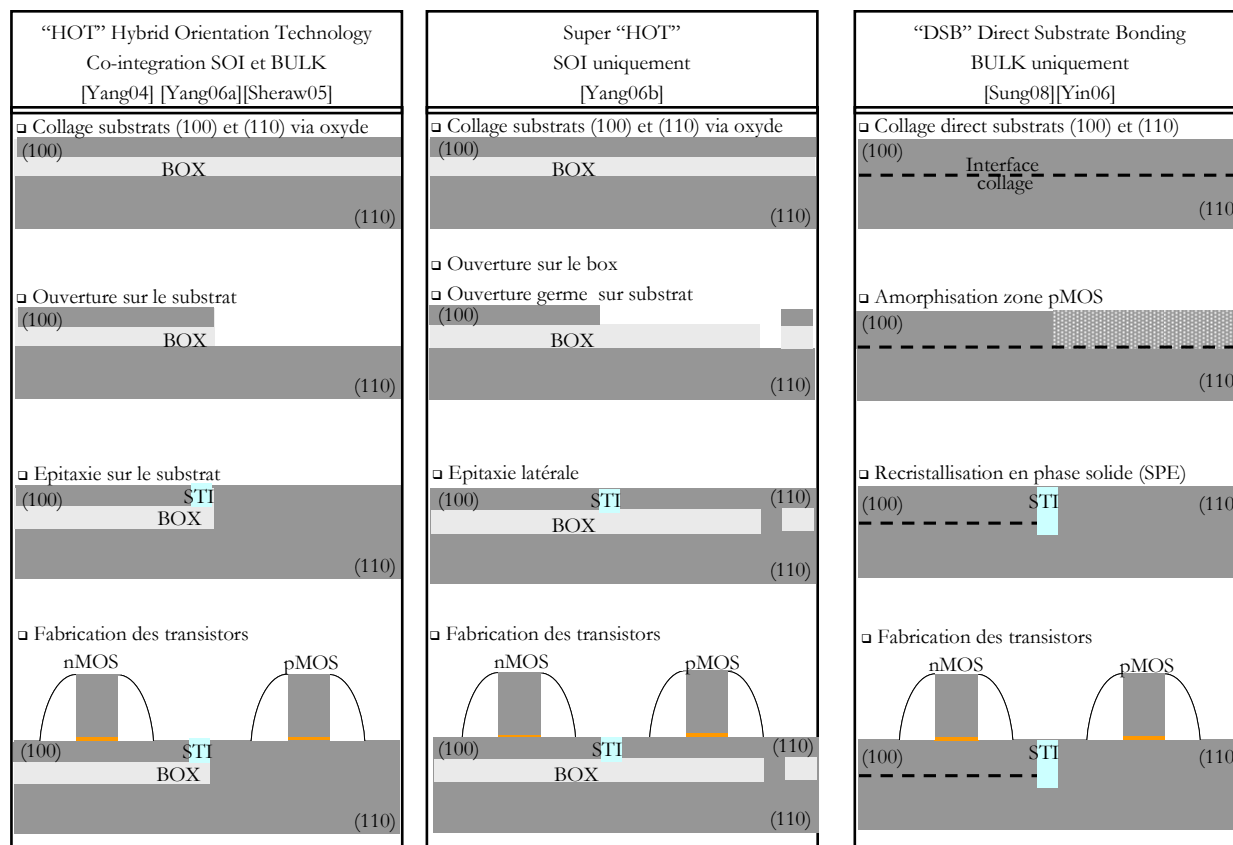


Figure 3.14 : Méthodes pour réaliser une co-intégration d'orientations avec un procédé planaire. Les procédés sont décrits de manière simplifiée, pour plus d'informations le lecteur pourra se reporter aux références données.

Dans [Yang06b], une évaluation des performances réalisée à l'aide du modèle compact SPICE montre que la co-intégration BULK-SOI (intégration « HOT » (Hybrid Orientation Technology)), apporte un gain en performance limité de 1 à 18% selon les portes étudiées, en passant à l'option dite « Super HOT », où les deux types de transistors sont sur SOI, le gain en performance se situe entre 16 et 21 % par rapport à une structure SOI planaire orienté classiquement 100.

² De manière standard, les transistors n et p sont réalisés sur des substrats de surface orientée (100). La mobilité des électrons dans un substrat massif est alors deux fois plus élevée que celle des trous. La technologie CMOS privilégie ainsi un nMOS plus performant que le pMOS à W égal.

L'intégration 3D séquentielle s'inscrit dans la lignée de l'intégration « super HOT » avec des transistors nMOS sur SOI (100) et des pMOS sur SOI (110). Dans la partie étude de coût, l'intégration 3D séquentielle avec double orientation sera comparée à l'intégration « super HOT ».

En plus de la co-intégration de ces deux orientations, des contraintes apportées par le procédé peuvent améliorer les performances des deux types de transistors. Ainsi, dans le cas du pMOS, des source/drain enterrés SiGe et/ou la réalisation de couches de nitrure compressives sur des substrats orientés (100) apportent un gain supplémentaire en termes de mobilité. [Sheraw05] [Wang06] [Tateshita06] [Yang 06a] [Yang07] [Packan08]. Un exemple de co-intégration d'orientation associé avec des contraintes optimisées pour les deux types de transistor est présenté dans la Figure 3.15.

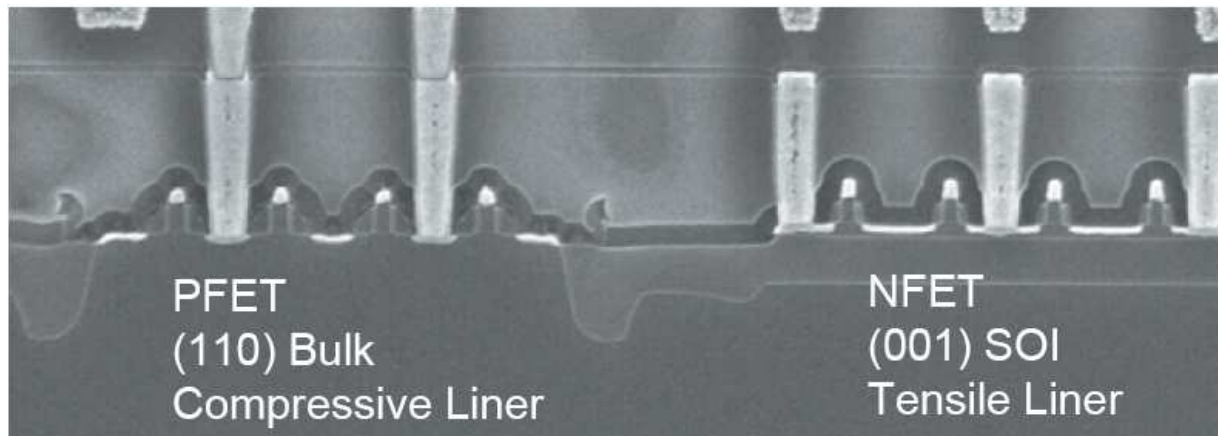


Figure 3.15 : Structure avec transistors n et pMOS à performances optimisées indépendamment par l'orientation du substrat et le type de contrainte (compression ou tension) appliquée [Sheraw05].

Une étude de simulation sur dispositifs longs, montre que le gain apporté par le changement d'orientation est réduit lorsque la contrainte compressive augmente, mettant en doute l'intérêt conséquent à changer d'orientation pour le pMOS. (Passage d'un gain d'un facteur 3.5 sans contrainte à un gain de 10% seulement pour une contrainte de 3GPa). [Packan08]. Cependant, une chute de mobilité plus faible avec la réduction des dimensions a été démontrée pour les substrats (110) par rapport aux substrats (100) [Yang07].

D'autre part, l'apport de contrainte à l'aide des couches de nitrures servant aussi de couche d'arrêt à la gravure des contacts (CESL pour Contact Etch Stop Layer) se complexifie pour des nœuds avancés. Le rapprochement des transistors nécessite d'utiliser des épaisseurs de nitrure plus faibles, qui contraignent moins. Dans ce contexte, l'intérêt pour les substrats double orientation est ravivé [Oishi05]. La difficulté observée à maintenir des contraintes élevées par l'utilisation des CESL pour des nœuds avancés peut aussi entraîner un regain d'intérêt pour les substrats contraints. L'intégration de contrainte via le substrat pour la zone active supérieure dans l'intégration ne requiert aucune étape supplémentaire, il suffit de remplacer le substrat SOI standard par un substrat SOI contraint.

Dans la partie II présentant les résultats électriques de l'intégration complète, nous avons démontré une co-intégration des orientations de surface (110) et (100): les transistors inférieurs sont réalisés sur une active dont la surface est orientée (100) et les transistors supérieurs sont réalisés sur une active dont la surface est orientée (110). Seuls des résultats dans la direction de conduction $\langle 100 \rangle$ moins intéressante en termes de mobilité que la direction $\langle 110 \rangle$ ont pu être

mesurés. Ceci est dû à un manque d'attention quant à la configuration des transistors sur le masque utilisé et non à une quelconque difficulté à orienter le substrat reporté dans telle ou telle direction de conduction.

B.1.2 - Co-intégration SOI/GeOI

Actuellement les transistors nMOS sur GeOI présentent des caractéristiques insuffisantes pour prétendre à une intégration dans un circuit. Le gain en mobilité des électrons dans le Ge massif par rapport à celle du Si ne se retrouve actuellement pas dans les dispositifs. Seul un gain de mobilité a été obtenu pour des pMOS sur GeOI par rapport au SOI. (Gain de l'ordre de 100% pour des matériaux non contraints [Chui03] [Yu04] à ~1000% pour du Ge contraint [Lee03][Weber05]). En conséquence, une co-intégration avec du silicium est nécessaire pour pouvoir profiter des meilleures performances du transistor pMOS Ge dans un circuit CMOS. Dans ce cadre, l'intégration 3D séquentielle permet de co-intégrer ces deux matériaux en empilant une zone active GeOI sur une zone active SOI tout en gagnant en densité.

En conclusion, grâce à la possibilité de réaliser des cellules CMOS à cheval sur les deux niveaux, l'association de l'intégration 3D séquentielle avec le collage moléculaire permet une optimisation des transistors n et pMOS en utilisant tous les bras de levier pour augmenter la mobilité des porteurs dans les transistors (contrainte, orientation et matériau du canal). Ce degré de liberté est particulièrement intéressant pour compenser la réduction de gain en courant de conduction des transistors d'une génération à l'autre dû au ralentissement de la réduction de la longueur de grille observée pour des nœuds avancés.

B. 2 -Gain en délai dans les interconnexions

Pour le nœud 45 nm, L'ITRS [ITRS Interconnect] annonce que les circuits intégrés devront comporter 12 niveaux de métaux et plusieurs km d'interconnexions par cm². Ceci donne une idée de la problématique complexe de la gestion du délai dans les interconnexions. La Figure 3.16 présente une observation en coupe des lignes de métal 1 à 8 de la technologie d'Intel 45nm.

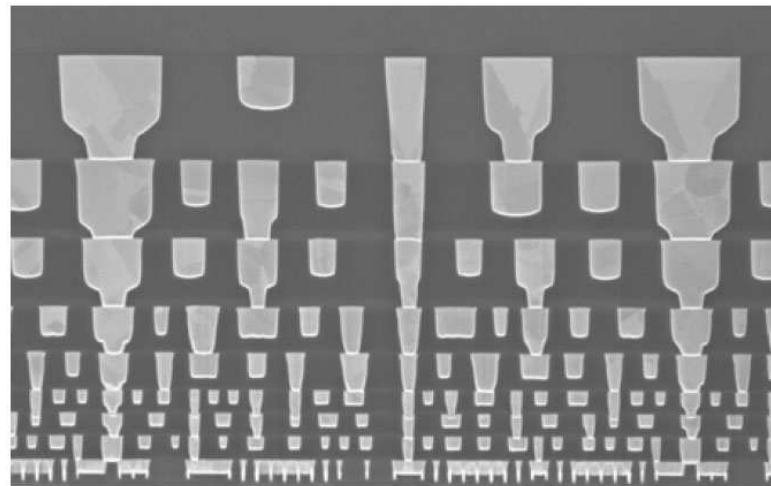


Figure 3.16 : Observation en coupe des niveaux de métaux 1 à 8 de la technologie d'Intel 45nm.

D'un coté, la longueur des interconnexions augmente car il y a toujours plus de transistors à interconnecter par unité de surface, et de l'autre, les sections de ces interconnexions ainsi que leur espacement sont réduits afin de limiter le nombre de niveaux de métaux nécessaires. Cette recherche de densification des réseaux d'interconnexions s'accompagne donc d'une augmentation du délai RC apparaissant dans celles-ci. Pour donner quelques ordres de grandeur, entre 1980 et 2000, le délai intrinsèque des transistors est passé de

~10 ps à ~1ps alors que celui apparaissant dans une interconnexion de 1mm de long est passé de ~1ps à ~100ps [Meindl01]. Le rapport relatif entre les deux délais s'inverse, et si aucune modification (technologique et/ou de design) n'avait été faite, le gain apporté par la réduction des dimensions du transistor n'aurait eu aucun impact sur la performance du circuit intégré.

Pour lutter contre l'augmentation du délai RC dans les interconnexions, des développements technologiques importants ont été effectués sur le Back-End : l'aluminium a été remplacé par le cuivre moins résistif et des diélectriques poreux à plus faible constante diélectrique ont été utilisés. Cependant, même avec l'introduction de ces nouveaux matériaux, le délai apparaissant dans les interconnexions longues reste problématique.

La Figure 3.17 présente l'évolution du délai dans une ligne métallique (niveau de métaux intermédiaires) avec les paramètres technologiques prévus par l'ITRS 2005 (prise en compte de l'introduction de diélectrique à faible permittivité et du cuivre, prise en compte de l'augmentation de la résistivité du aux joints de grains et aux collisions électroniques..)

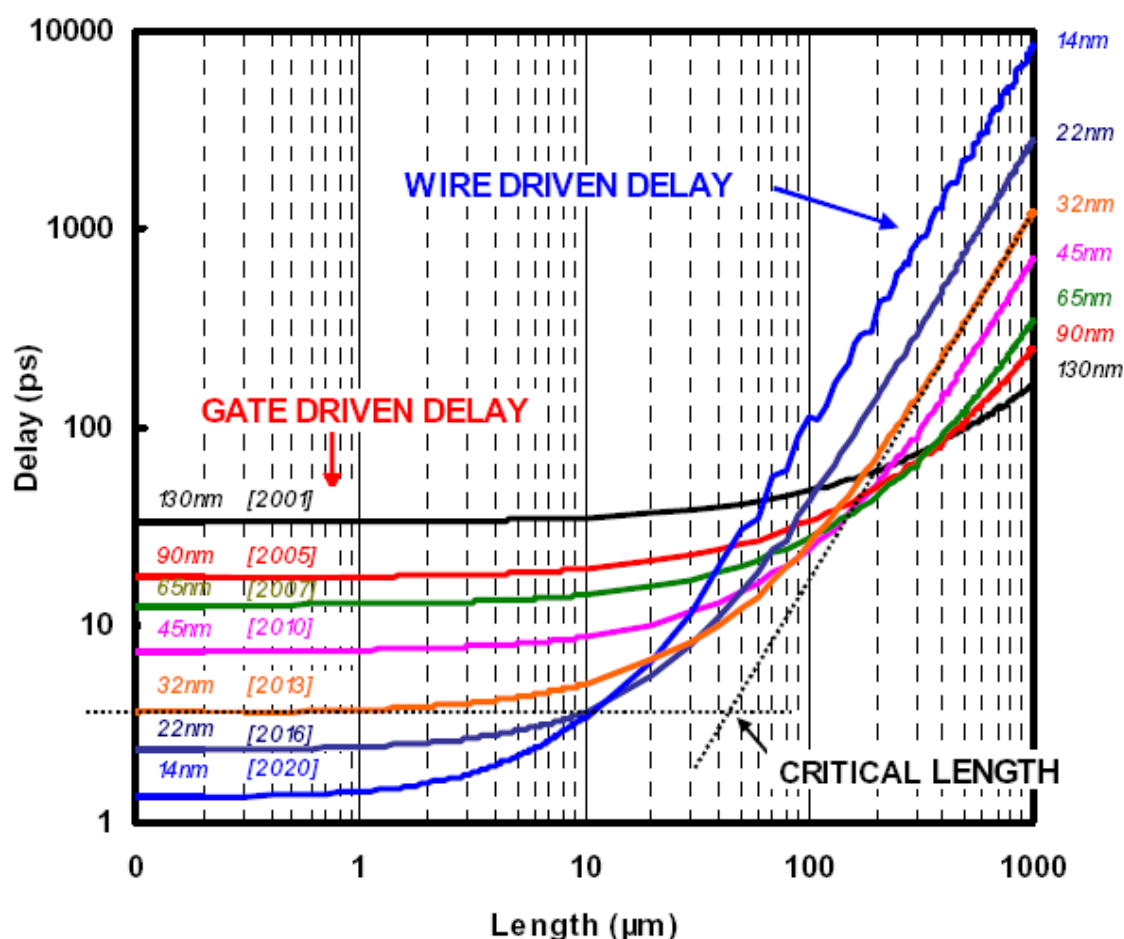


Figure 3.17 : Délai en fonction de la longueur de l'interconnexion pour différents nœuds technologiques [Sellier08]

La courbe fait apparaître deux régions, une aux faibles longueurs d'interconnexions, où le délai dans la ligne est limité par le délai intrinsèque du transistor et une aux grandes longueurs d'interconnexions où le délai est limité par la ligne elle-même. Pour bénéficier de la diminution du

délat intrinsèque du transistor avec la réduction des dimensions, les lignes doivent avoir des dimensions inférieures à la longueur critique définie comme l'intersection des asymptotes des deux régimes. Cette longueur critique est passée de 200 μm pour le nœud 90 nm à 20 μm pour le nœud 22nm. Une manière de limiter le délat dans les lignes longues est d'insérer des buffers³ le long de la ligne à des distances inférieures à la longueur critique. En conséquence, pour le nœud 45nm un buffer doit être inséré tout les 70 μm environ. A chaque insertion de buffer, la place utilisée pour le buffer ainsi que l'espace nécessaire pour faire redescendre le signal jusqu'aux transistors est perdu. Il en découle une double perte de place.

Ce contexte valorise nettement la technologie 3D parallèle car elle peut permettre de réduire la longueur des interconnexions. Comme illustré dans la Figure 3.18, l'objectif est de rapprocher des blocs de transistors lointains sur le même plan. De la même manière que dans l'étape de placement routage en technologie planaire, où l'outil optimise le placement des cellules pour réduire la longueur du réseau de routage les reliant, une étape de placement routage en 3D permet l'utilisation d'une dimension supplémentaire pour optimiser la longueur des interconnexions. Dans le cas de la technologie 3D parallèle, la densité maximale de vias atteignable est de l'ordre de 10⁸ vias.cm², dans ce cas les vias relient des groupes consécutifs de transistors (~100 transistors) dessinés en 2D.

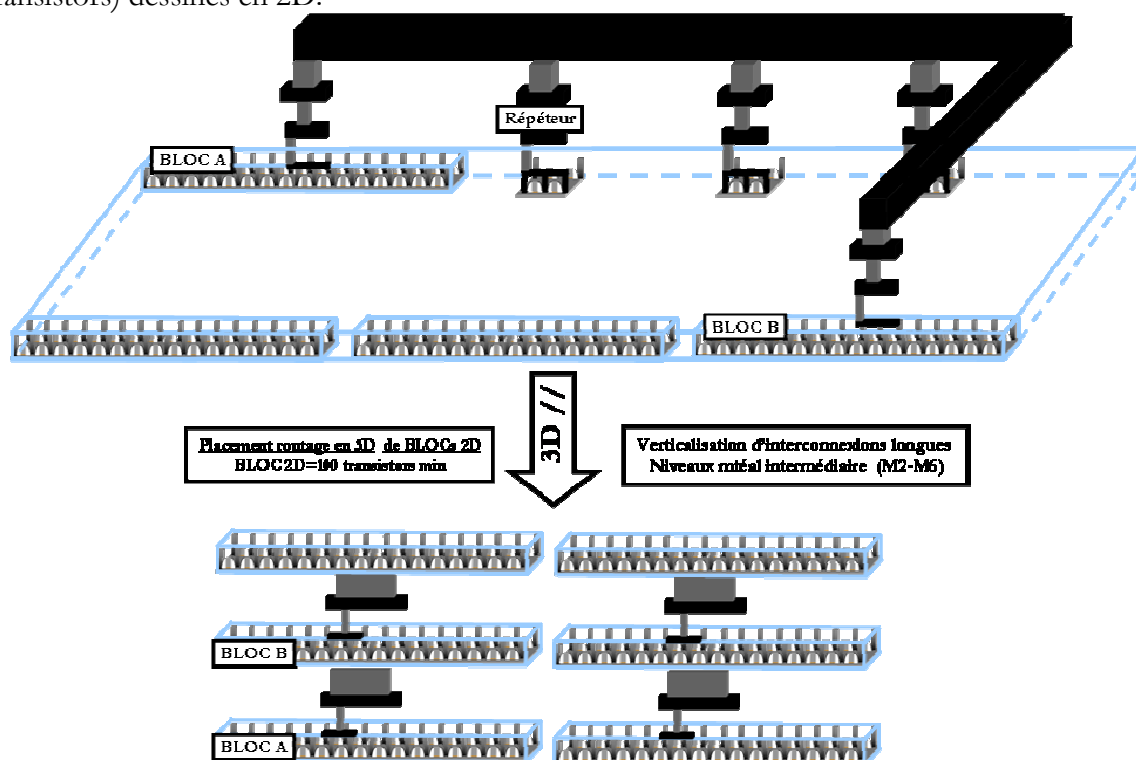


Figure 3.18 : Représentation schématique de la réduction des interconnexions avec l'intégration 3D parallèle reliant des blocs de transistors consécutifs. En général, à l'intérieur de chaque bloc, les transistors sont reliés par les niveaux M1 et M2.

³ Les buffers sont des cellules destinées à amplifier le courant et sont composés d'une succession d'inverseurs de taille croissante.

Dans la technologie 3D séquentielle, telle que nous l'avons pensée dans ce sujet de thèse, les cellules standard sont dessinées en 3D et le placement routage des cellules est réalisé en planaire. S'il y a réduction des longueurs des interconnexions, il s'agit d'interconnexions déjà relativement courtes en planaire. (cf Figure 3.19)

Pour pouvoir adresser la problématique des interconnexions longues avec une intégration 3D séquentielle, il faudrait que des niveaux de back end soient intercalés entre les niveaux de transistors. Pour cela, les transistors des niveaux supérieurs doivent être réalisés en dessous du budget thermique maximal supporté par le back end (~400°C) ou bien que les niveaux de métaux intermédiaires puissent supporter un budget thermique de 600°C.

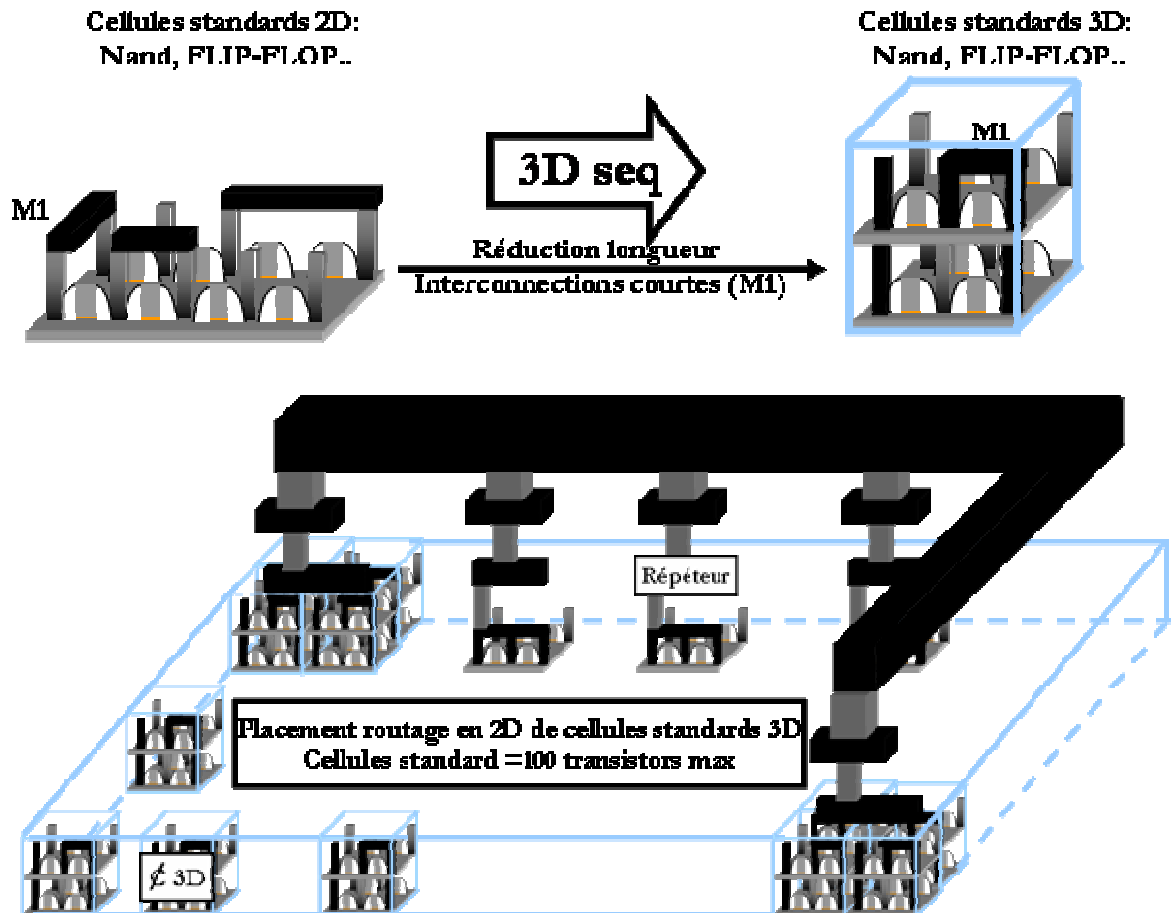


Figure 3.19 : Représentation schématique de la réduction des interconnexions avec l'intégration 3D séquentielle. Les interconnexions « verticalisées » sont des interconnexions courtes. Dans l'intégration séquentielle démontrée dans cette thèse, la réalisation d'un circuit complexe se fait par placement routage en 2D de cellules standard dessinées en 3D.

Dans la littérature, les études de gain en délai menées sur les cellules standard, donnent des gains allant de 14 à 30 % [Chan00] [Roos92] [Lim05] [Zhang99].

Cependant ce gain en délai dans les cellules standard sera certainement plus modéré au niveau d'un circuit complexe. En effet puisqu'il est possible de mettre plus de transistors sur la même surface, à surface constante, le réseau d'interconnexions pour les relier sera plus important. Quelle que soit l'option choisie (augmentation du nombre de niveaux de métaux, choix d'une technologie plus avancée pour le niveau back end), le délai des interconnexions va augmenter.

Pour conclure sur ce point, il serait nécessaire de réaliser l'étude sur des cellules beaucoup plus complexes routées au moins jusqu'aux niveaux de métaux M4 -M5.

Remarque sur la thermique :

Si l'intégration promet des bénéfices significatifs, le problème de l'échauffement thermique peut venir contrebalancer ces gains par une diminution de performance et de fiabilité [Banerjee01][Topol06][Rahman00][Loi06][Im00].

Il est évident que l'aspect thermique va être encore plus problématique dans un empilement 3D qu'il ne l'est déjà dans une intégration planaire. En effet, la puissance par unité de surface va être augmentée puisque l'on gagne en densité. De plus, à cause de la faible conductivité des diélectriques entre les niveaux empilés, la chaleur dégagée par les transistors et interconnexions va être plus difficile à évacuer qu'en planaire.

En conséquence, un gradient vertical de température pourrait entraîner des caractéristiques de fonctionnement différentes pour les transistors des niveaux empilés. La connaissance de ce gradient de température est essentielle pour réaliser le layout des circuits.

Actuellement les études sur la thermique sont surtout réalisées pour des architectures du type 3D parallèle où l'épaisseur du diélectrique inter-niveau et la densité de vias diffèrent largement avec l'empilement séquentiel. De plus, dans la 3D parallèle, la réflexion sur la thermique s'oriente beaucoup vers l'optimisation du placement routage en 3D pour limiter l'échauffement de température dans la structure. Dans la mesure où nous ne plaçons pas de niveaux de métal entre les différents niveaux de transistors, cette réflexion n'est pas applicable à l'intégration présentée.

Pendant cette thèse, la problématique de l'échauffement dans l'empilement 3D séquentiel n'a pas été étudiée. Pour la suite, cette étude apparait indispensable pour évaluer les pertes et les différentes problématiques liées à l'échauffement (en particulier la variabilité des caractéristiques qu'elle génère).

Ce qu'il faut retenir
IV-B : Gain en performance

- La possibilité de réaliser des cellules CMOS à cheval sur les deux niveaux offerte par l'intégration séquentielle, facilite la mise en oeuvre de tous les bras de leviers actuellement utilisés pour augmenter la mobilité des porteurs. Les deux types de transistors peuvent bénéficier des meilleures options.
- L'utilisation du collage moléculaire pour réaliser la zone active supérieure permet la co-intégration de différentes orientations et contraintes apportées directement par le substrat entre le niveau haut et bas par opposition aux techniques dites « seed window » décrites dans le chapitre I.
- Dans le cas d'une co-intégration de substrats d'orientations différentes sur isolant, l'architecture « super HOT » possède les mêmes limitations que les techniques de « seed

window » décrites dans le chapitre I (place de la fenêtre d'ouverture inutilisable pour réaliser des transistors, complexité de l'étape de cristallisation, qualité cristalline de la zone active et contrôle de son épaisseur).

- En terme de gain en délai, les intégrations 3D séquentielles et parallèles doivent être distinguées :

Dans l'intégration parallèle, les interconnexions « verticalisées » sont les interconnexions longues, dans lesquelles le délai est le plus grand. Une étape de placement routage en trois dimensions permet de trouver un optimum de positionnement pour réduire la longueur des interconnexions longues. Les blocs de cellules placés en 3D sont eux dessinés en planaire.

Dans le cas séquentiel, l'absence de niveaux métalliques entre les niveaux de transistors empêche la réalisation d'un placement routage en 3D. La troisième dimension apparaît seulement au moment du dessin des cellules standard. Les interconnexions « verticalisées » sont alors des connexions courtes (niveau M1). Le gain en délai est alors plus limité que dans une technologie 3D parallèle.

	3D parallèle	3D séquentiel
Dessin de cellules standard	en 2D	en 3D
Placement routage	en 3D	en 2D
Interconnexions verticalisées	interconnexions longues, aux niveaux routage intermédiaire et globale	interconnexions courtes, internes aux cellules standard (M1)

C - Etude de coût

C. 1 -Présentation du cadre de l'étude

L'objectif de cette partie est de faire une étude de coût comparative entre l'intégration 3D séquentielle et l'intégration planaire. Un procédé d'intégration étant composé d'étapes simples et maîtrisées telles que des étapes de gravure, lithographie, épitaxie. Ces étapes se retrouvent de la même manière dans le flot d'intégration planaire que de le flow d'intégration 3D. La seule étape que l'on pourrait qualifier de nouvelle (c'est-à-dire qu'on ne la retrouve pas dans une intégration planaire) est l'étape de collage. Cependant l'ensemble des étapes nécessaires pour réaliser le transfert de la zone active supérieure sont connues et se décrivent aussi par des étapes élémentaires (dépôt, CMP, préparation de surface. .etc). Le résultat de cette étude correspondra à un surcoût exprimé en pourcents de l'intégration 3D séquentielle par rapport à l'intégration planaire.

Nous avons choisi différentes stratégies d'intégration : une orientation que l'on qualifiera de standard plutôt orientée basse consommation et une option que l'on qualifiera de « haute performance ». La Figure 3.20 récapitule les différentes options étudiées.

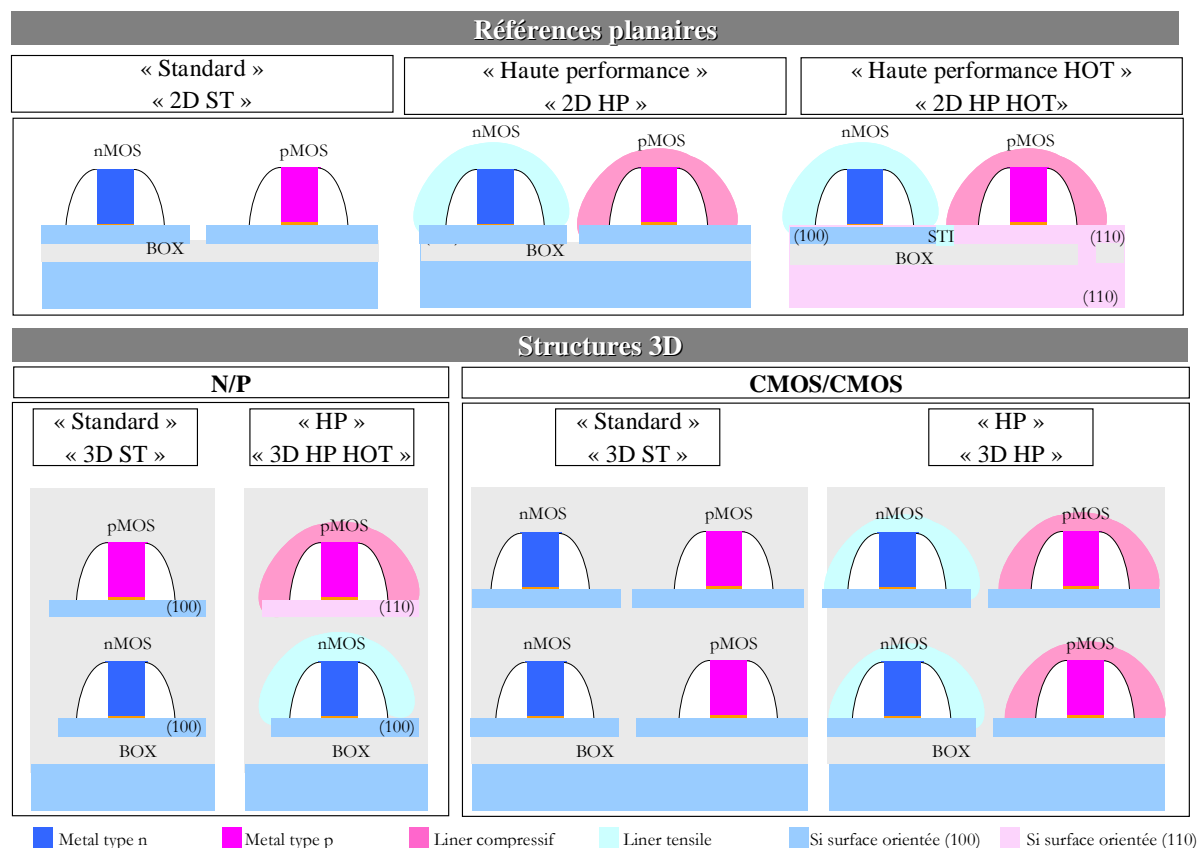


Figure 3.20 : Description des intégrations comparées dans l'étude de coût.

Le procédé standard correspond au procédé FDSOI tel qu'il est réalisé au Leti et présenté en figure 1.1 (ch1) (siliciuration comprise). De plus il intègre deux métaux de grilles adaptés pour le

nMOS et le pMOS. Pour passer à l'intégration de type haute performance, la réalisation de couches contraignantes adaptées pour les deux types de transistors sont prises en compte. Enfin une option supplémentaire dite « haute performance HOT » où un substrat SOI est remplacé par le substrat de type SUPER HOT proposé dans [Yang 06b]. Ce type de substrat permet d'optimiser l'orientation du substrat pour les deux types de transistors, i.e. (110) pour le pMOS et (100) pour le nMOS.

Nous comparerons ces deux options à leur équivalent en trois dimensions. Nous étudierons les deux architectures précédemment étudiées dans l'étude de densité ; intégration N/P avec un niveau de transistor N empilé au-dessus d'un niveau de transistor P et intégration CMOS/CMOS où les deux couches contiennent les deux types de transistors.

Lorsque nous ferons des comparaisons de coût, nous comparerons uniquement les options menant à des performances égales. Plus précisément, nous comparerons l'option 3DST à l'option ST planaire, l'option HP 3D à l'option HP 2D. Dans cette dernière, une distinction doit cependant être faite : nous pourrions comparer l'intégration planaire HP HOT uniquement avec l'option 3D HP N/P, où les deux types de transistors sont bien fabriqués sur des substrats d'orientations différentes. L'option 3D HP CMOS/CMOS ne permet pas d'optimiser l'orientation des zones actives pour les deux types de transistors sur les deux niveaux. Elle sera alors comparée à l'option planaire HP sur substrat SOI simple.

En plus des étapes du procédé planaire classique, l'option 3D va intégrer une étape de siliciuration optimisée (NiPtSi avec implantation W et F) pour les transistors inférieurs, le report de couche de la zone active supérieure par collage moléculaire et une lithographie contact supplémentaire (nécessaire pour réaliser des contacts de type traversant). La zone active supérieure peut être reportée en utilisant un substrat SOI ou un substrat Si massif avec un réseau de couches d'arrêt Si-SiGe comme décrit dans la Figure 3.21.

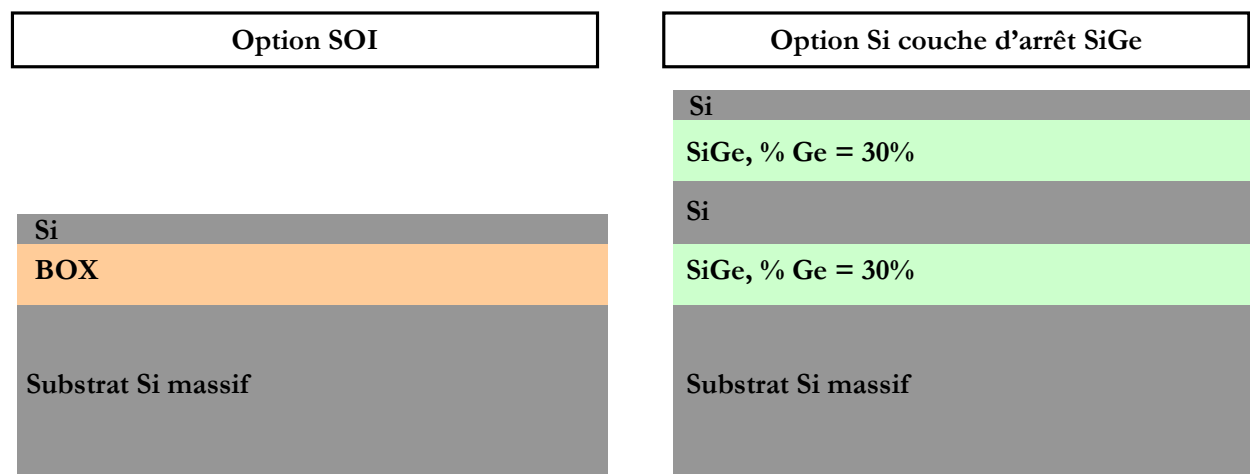


Figure 3.21 : Schéma des deux types de substrat considérés dans les études de coût pour réaliser la zone active supérieure

Le procédé front end est suivi d'un procédé back end avec 8 niveaux de métaux.

Le prix de chaque étape a été calculé en prenant en compte les frais d'achat et de connexion des équipements, l'amortissement sur 7 ans, la maintenance de l'équipement, le coût de la surface occupée en salle blanche ainsi que ses frais de fonctionnement, le temps de procédé par plaque et le salaire des opérateurs correspondants ainsi que les consommables. Cette étude a été menée

pour une hypothèse de 100'000 plaques fabriquées par an. Elle a été réalisée par Bernard Diem du CEA Leti.

C. 2 - Résultats

La Figure 3.22 présente le surcoût par plaque des différentes options d'intégration 3D par rapport aux références à performances équivalentes en planaire. Pour chaque option, i.e. « 3D N/P ST », « 3D CMOS/ CMOS ST », « 3D N /P HP HOT », « 3D CMOS/ CMOS HP », nous présenterons les résultats avec une active supérieure réalisée à partir d'un substrat SOI ou d'un substrat Si massif à couche d'arrêt SiGe. Nous considérons que le prix du substrat SOI vaut 9 fois le prix d'un substrat Si massif, i.e. 300€.

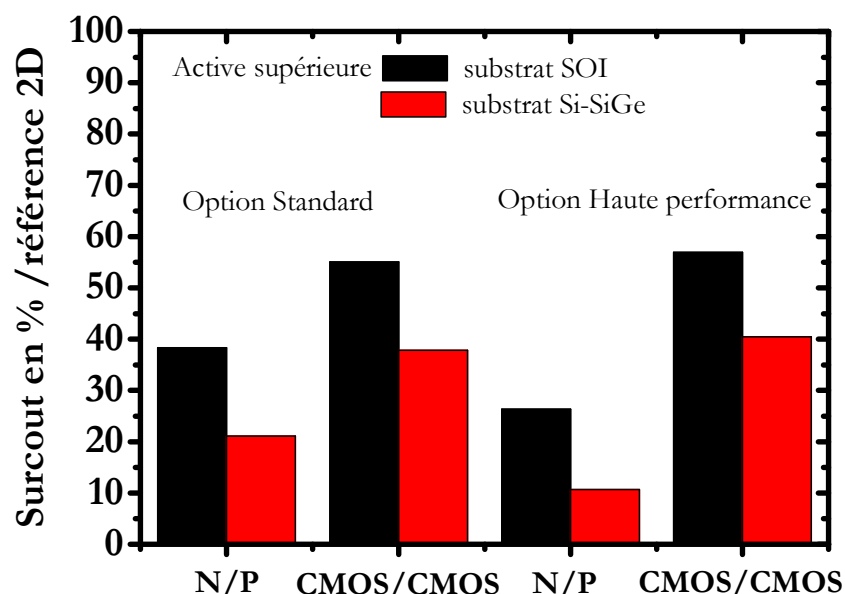


Figure 3.22 : Surcoût par plaque des différentes options d'intégrations 3D par rapport aux références à performances équivalentes en planaire.

indépendamment pour le nMOS et le pMOS en planaire, plus le gain sera important en passant à l'intégration 3D N/P.

Enfin, nous observons que l'utilisation d'un substrat SOI supplémentaire pour réaliser l'active supérieure contribue nettement à l'augmentation du surcoût. Lorsque l'on remplace le substrat SOI par un substrat Si massif avec couches d'arrêt en SiGe, le surcoût maximum reste inférieur à 40% quelles que soient les options. Un surcoût minimum de 11% est obtenu dans le cas de l'intégration de type haute performance N/P.

L'ensemble des intégrations 3D conduit à un surcoût se situant entre 10% et 55%.

L'option N/P est nettement moins chère que l'option CMOS/CMOS quelle que soit l'option de performance (ST ou HP). Dans l'option N/P nous observons une réduction du surcoût non négligeable lorsque l'on passe de l'option standard à l'option haute performance. Plus les transistors sont optimisés

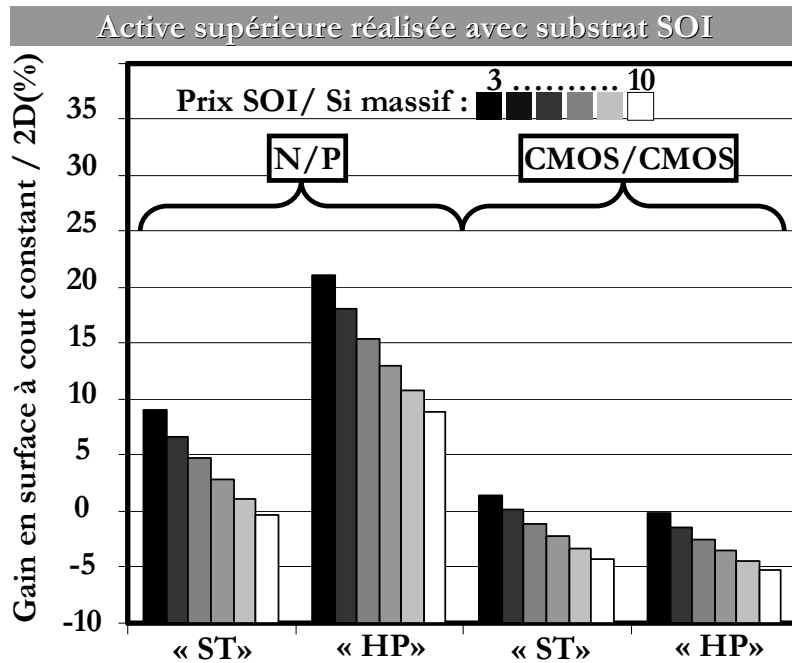


Figure 3.23 : Surcoût par plaque des différentes options d'intégrations 3D par rapport aux références à performances équivalentes en planaire. Cas avec la zone active supérieure réalisée à partir d'un substrat SOI

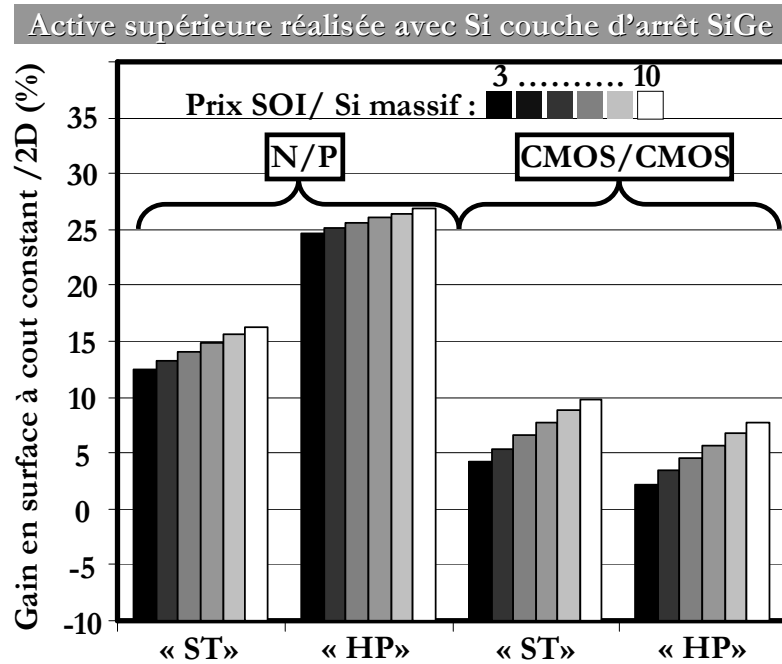


Figure 3.24 : Surcoût par plaque des différentes options d'intégrations 3D par rapport aux références à performances équivalentes en planaire. Cas avec la zone supérieure réalisée à partir d'un substrat Si avec couches d'arrêt SiGe

substrat Si massif avec des couches d'arrêts SiGe sont nettement plus encourageants avec un gain en densité d'environ 25% à coût constant pour la variante N/P haute performance.

Les valeurs des surcoûts par plaque sont à comparer aux gains en densité précédemment estimés.

Nous retiendrons un gain de 40% en surface pour l'option N/P et de 50% en surface pour l'option CMOS/CMOS. Grâce à l'étude conjointe du surcoût et du gain en densité, nous pouvons calculer le gain en densité possible à coût constant. Les résultats sont présentés dans les Figure 3.23 (cas zone active supérieure à partir d'un SOI) et Figure 3.24 (cas zone active supérieure à partir d'un substrat Si massif à couche d'arrêt SiGe).

Le rapport des prix du substrat SOI et Si massif est variable (d'un facteur 3 à un facteur 10.) Actuellement le rapport est estimé aux alentours de 9. Dans le cas d'une zone active supérieure réalisée avec un substrat SOI, nous observons que l'option CMOS/CMOS conduit finalement à une perte de densité pour un coût constant (jusqu' à 6%). Par contre l'option N/P permet bien de gagner en densité à coût constant.

Les résultats dans le cas où la zone active est réalisée à partir d'un

Nous observons que l'intégration 3D séquentielle est susceptible d'apporter un gain en densité à coût constant. Ce gain à coût constant est particulièrement remarquable car il est obtenu sans réduire les dimensions des transistors.

Cependant, cette étude a été réalisée en considérant un produit purement logique avec des transistors FDSOI. Tous les transistors d'un type (n ou p) ont la même tension de seuil. En conséquence, elle ne permet pas de conclure sur le gain en densité à coût constant pour un produit réel. En effet, dans un produit réel, différents types de transistors sont fabriqués dans le but de remplir des fonctions différentes. Par exemple, des transistors à faible tension de seuil sont utilisés pour les blocs logiques haute performance, alors que des transistors de type LSTP à plus forte tension de seuil sont utilisés dans les blocs de SRAM. Les transistors d'entrée/sortie possèdent des oxydes de grille plus épais. Enfin, l'option FDSOI n'est actuellement pas compatible avec la réalisation de prises ESD et le fait que les résistances d'accès soient plus importantes sur films minces que sur substrats massifs fait de la technologie FDSOI une option peut être plus adaptée pour des produits basse consommation. En partie pour cette raison, la technologie FDSOI est parfois pensée dans une logique de co-intégration avec des zones sur substrat massif comme présenté dans la Figure 3.25. Par exemple, les blocs de logique haute performance, les prises ESD, l'analogique, les composants RF pourraient être réalisés sur le substrat massif et la SRAM basse consommation sur un substrat FDSOI.

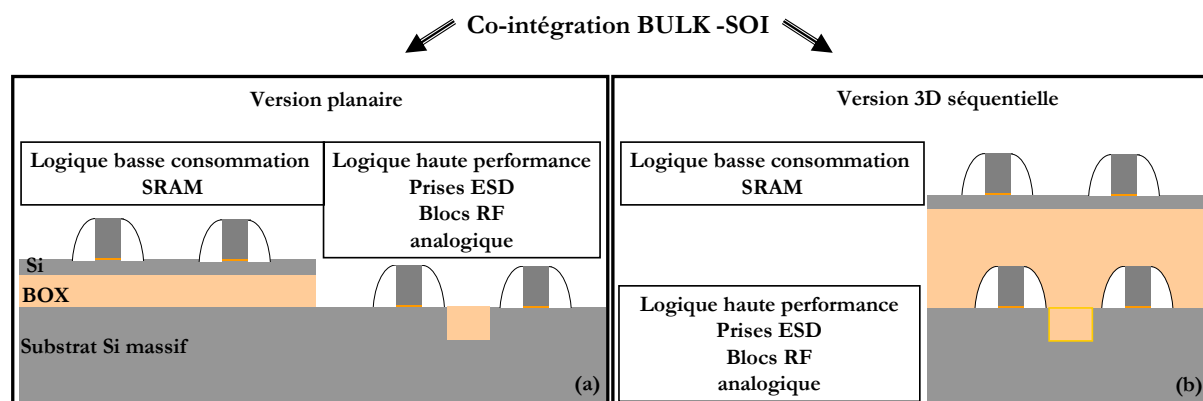


Figure 3.25 : Représentation schématique d'une co-intégration SOI/ BULK en planaire et en technologie 3D séquentielle.

Enfin, pour pouvoir conclure sur le gain en densité à coût constant sur un produit réel, il serait aussi nécessaire de mener une étude bien plus poussée que celle qui a été réalisée dans cette thèse. L'exemple de co-intégration SOI/ substrat massif avec la technologie 3D séquentielle serait un exemple d'intégration intéressant à étudier.

Néanmoins, dans l'hypothèse d'une intégration purement logique avec une seule valeur de tension de seuil pour les transistors n et p, il est tout de même intéressant de voir que l'intégration 3D séquentielle ne double pas le coût du procédé par deux. La part du prix du Back end dans le coût total du produit vient diminuer l'impact de l'augmentation du nombre d'étapes de front end dans l'intégration 3D séquentielle. Nous observons grâce à cette étude que la séparation des deux types de transistors sur les niveaux empilés peut réduire de manière non négligeable le surcoût du passage en 3D. Cette étude met donc en évidence l'intérêt du développement du report de la zone active à partir d'un substrat massif.

Ce qu'il faut retenir

IV-B : Gain en coût

- Une étude comparative de coût dans l'hypothèse d'un produit purement logique avec une tension de seuil unique pour les deux types de transistors a été menée. Différents schémas d'intégration en 3 dimensions ont été étudiés : Intégration de type N/P ou CMOS/ CMOS orientée haute performance ou basse consommation. Les intégrations 3D sont comparées à des références menant à des performances similaires. L'influence du prix relatif du substrat SOI par rapport à un substrat Si massif sur le surcoût de la technologie a été étudiée ainsi que le remplacement du substrat SOI par un substrat Si massif avec couche d'arrêt SiGe pour réaliser la zone active supérieure.
- Dans le cadre des hypothèses faites, le surcoût de la technologie 3D séquentielle par rapport à l'intégration planaire variant entre 10% et 55% selon les options d'intégration a été estimé.
- Nous observons que les options de type N/P conduisent à une diminution du surcoût conséquent par rapport à l'intégration de type CMOS/CMOS. Par exemple, l'estimation de surcoût passe de 41% à 11% pour l'option orientée haute performance par rapport à la référence planaire correspondante. Rappelons que la possibilité de réaliser des cellules CMOS à cheval sur les deux niveaux est une particularité de l'intégration 3D séquentielle.
- Nous observons également que plus les transistors n et p sont optimisés indépendamment plus le surcoût en 3D N/P est faible. Passage de 21% à 11% de l'estimation du surcoût par rapport aux intégrations planaires lorsque l'on passe de l'option standard à l'option haute performance. (Rappel : l'option haute performance permet de choisir une contrainte et une orientation optimisée pour les deux types de transistors)
- Enfin cette étude met en évidence l'intérêt financier du développement du report de la zone active supérieure à partir d'un substrat massif. L'estimation de la réduction de surcoût par rapport à la technologie planaire en utilisant un substrat Si massif avec couche d'arrêt en SiGe à la place d'un substrat SOI pour réaliser l'active supérieur est de 15% environ

D - Conclusion générale sur l'étude réalisée

Dans ce chapitre, nous avons mené des études visant à estimer les gains en densité et en surcoût de l'intégration 3D séquentielle par rapport à l'intégration planaire. Ces études ont été limitées au cas particulier d'intégration logique pure. Un des objectifs de ces études a été de comparer les deux schémas d'intégration permis par l'intégration 3D séquentielle ; à savoir : un schéma où les deux types de transistors sont réalisés sur des niveaux différents (3D N/P) et un schéma où les deux types de transistors peuvent être réalisés sur chaque niveau (3D CMOS/CMOS).

Les études réalisées sont à prendre avec beaucoup de précautions pour les raisons suivantes :

Etude de densité :

- Cette étude est d'abord basée sur un DRM qui est un DRM théorique, c'est-à-dire qu'il est extrapolé à partir de règles planaires et qu'il n'a pas été validé technologiquement par des lots fonctionnels et des études de rendement.

- Le gain en densité a été estimé par une étape de synthèse logique sur une cellule comportant environ 1000 transistors. En conséquence, cette étude ne nous permet de conclure sur le gain en densité après la synthèse physique (c'est-à-dire après optimisation du positionnement de toutes les cellules standard et réalisation des interconnexions (étape de routage)). En particulier, nous pouvons nous attendre à faire face à la problématique de congestion de routage au niveau de cette étape. D'autre part, l'étude de gain en densité a été réalisée sur une cellule de 1000 transistors seulement. Pour conclure sur cet aspect il est nécessaire d'aller jusqu'à la réalisation de circuits intégrés avec au moins 4 à 5 niveaux de métal.

Etude de coût

- De part sa réalisation au Leti, cette étude de coût peut être biaisée par rapport à une étude de coût dans un environnement industriel plus attaché à la notion de rentabilité.

- Le cadre de cette étude est relativement limité. Plus précisément il n'adresse que des empilements comportant de la logique. Le cas de transistors multi- V_T avec plusieurs valeurs de tension de seuil afin de réaliser sur le même substrat des transistors pour de la logique haute performance et des transistors pour des applications basse consommation (ex :SRAM), n'a pas été traité. De plus, cette étude a été menée dans le cadre restreint de la technologie FDSOI, dans laquelle nous ne connaissons pas pour l'instant de solution pour réaliser les prises de protection contre les décharges électriques (prises ESD) et suspectée d'être moins pertinente pour des applications haute performances, réalisation de fonctions radio-fréquence et analogiques que la technologie sur substrat silicium massif.

Cependant les études menées apportent des éléments de réponse supplémentaires par rapport aux études de l'état de l'art.

Etude de densité

- Réalisation d'une étude pour un nœud plus avancé (45nm) que les études de la littérature.
- Mise au point d'un DRM avec prise en compte des spécificités technologiques de l'intégration 3D séquentielle
- Comparaison de deux architectures (option N/P, CMOS/CMOS) sans ligne de métal entre les deux niveaux de transistors.

Etude de coût :

A notre connaissance, aucune étude de coût sur la technologie 3D séquentielle pour des intégrations logiques pures n'a été publiée. On notera cependant, les résultats présentés par Samsung sur l'intérêt financier de l'intégration 3D séquentielle pour des applications FLASH Nand. [Jung06]. Cette étude estime que l'intégration 3D séquentielle permet de continuer à réduire le coût par bit de 40% à chaque génération (cf Figure 3.26)

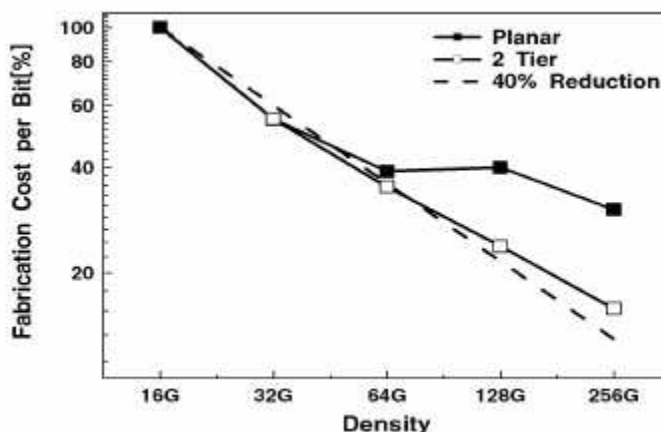


Figure 3.26 : Prédiction du coût de fabrication de mémoires FLASH au fur et à mesure que la densité des cellules augmente dans une intégration planaire et dans l'intégration 3D séquentielle selon le procédé présenté dans [Jung06]

L'étude de coût réalisée permet de pondérer les résultats de l'étude de densité. Ainsi, nous pouvons estimer le gain en densité à coût constant. Nous observons que le résultat de l'étude de densité et de l'étude de coût ne valorise pas la même option. Dans le cas de l'étude de densité, c'est l'option CMOS/CMOS qui apporte le plus de gain (~50% contre 40% dans le cas N/P). L'étude de coût met en valeur le surcoût plus important de cette option. La pondération de l'étude de coût par l'étude de densité permet de valoriser l'option N/P par rapport à l'option CMOS/CMOS.

Critère de comparaison	Option d'intégration	
	N/P	CMOS/CMOS
Densité	-	+
Performance	+	-
Coût	+	-

De plus l'option N/P peut conduire à des gains supplémentaires en termes de performances des transistors par le biais de l'optimisation indépendante des deux types de transistors, en intégrant par exemple du germanium (voir du germanium contraint) pour les transistors de type p.

Cette estimation au premier ordre est encourageante car la possibilité de réaliser les cellules CMOS à cheval sur les deux niveaux est une spécificité de l'intégration séquentielle, elle conforte donc l'intérêt généré par l'intégration 3D séquentielle et engage à réaliser des études plus complètes pour conclure de manière fiable sur les perspectives de gains apporté par cette technologie.

CHAPITRE IV

LES AUTRES APPLICATIONS DE L'INTEGRATION SEQUENTIELLE

En dehors des applications logiques hautes densités, l'architecture 3D séquentielle est prometteuse pour de nombreuses applications.

En premier lieu, nous citerons les applications mémoires SRAM et FLASH, domaine dans lequel le gain en densité est primordial.

De plus, la proximité des transistors empilés peut permettre une modification dynamique de la tension de seuil du transistor supérieur par couplage avec la grille du transistor inférieur à travers un diélectrique inter-niveau aminci. Cette particularité de l'intégration 3D séquentielle peut être mise à profit pour augmenter la stabilité des cellules SRAMs.

L'intégration 3D séquentielle semble aussi être intéressante pour la réalisation d'imageurs ultra denses. Enfin elle permet une simplification de co-intégration de technologies hétérogènes diverses.

Ce chapitre est dédié à l'étude de la pertinence de l'architecture 3D pour ces applications.

SOMMAIRE DU CHAPITRE IV :

LES AUTRES APPLICATIONS DE L'INTEGRATION 3D SEQUENTIELLE

A - LA STRUCTURE UTILD.....	119
A. 1 - Préambule sur l'intérêt de la modification dynamique de tension de seuil	119
A. 2 - Dimensionnements pour obtenir un couplage conséquent	120
A. 3 - Evolution en fonction du désalignement et cas avec grilles empilées de longueurs différentes.....	124
A. 4 - La réalité technologique.....	126
B - LES MEMOIRES SRAM	129
B. 1 - Les mémoires SRAM 3D.....	129
B. 2 - Les problématiques de conception de cellules SRAMs.....	130
B. 3 - La SRAM4T 3D UTILD	131
B.3.1 - <i>Présentation de la SRAM 4T et analyse de sa stabilité.....</i>	<i>131</i>
B.3.2 - <i>Positionnement des transistors en 3D.....</i>	<i>132</i>
B.3.3 - <i>Analyse des résultats</i>	<i>134</i>
B. 4 - La SRAM 6T 3D UTILD	137
B.4.1 - <i>Présentation du modèle compact.....</i>	<i>137</i>
B.4.2 - <i>Présentation de la SRAM 6T et analyse de sa stabilité.....</i>	<i>138</i>
B.4.3 - <i>Positionnement des transistors et analyse des résultats.....</i>	<i>139</i>
C - LES MEMOIRES FLASH	142
D - APPLICATION AUX IMAGEURS PETITS PIXELS.....	144

A - La structure UTILD

A. 1 - Préambule sur l'intérêt de la modification dynamique de tension de seuil

Les performances d'alignement de l'intégration séquentielle permettent de placer un transistor supérieur précisément au dessus d'un transistor inférieur (déviaton standard de l'alignement de l'ordre de 10nm). De plus, la distance entre deux transistors empilés peut atteindre la dizaine de nanomètres en optimisant l'épaisseur de diélectrique interniveau. (Pour rappel, l'épaisseur du diélectrique inter-niveau, T_{ILD} , est mesurée entre le haut de la grille du transistor inférieur et la zone active du transistor supérieur, comme décrit dans la Figure 4.1 (a)).

A de telles distances, un couplage électrostatique significatif entre les deux transistors empilés apparaît et va modifier leurs caractéristiques de fonctionnement. Ce couplage va induire, par exemple, une modification de la tension de seuil du transistor supérieur en fonction de la polarisation du transistor inférieur.

Ce couplage peut être considéré comme parasite en ce qu'il affecte le comportement des transistors. Cependant de nombreuses publications mentionnent l'intérêt de la modification dynamique de la tension de seuil afin d'augmenter les performances des mémoires SRAM [Hirano08],[Mukhopadhyay08] ou d'obtenir un MOSFET avec un fort courant à l'état passant et un faible courant à l'état bloqué [Assaderaghi97],[Ishigaki08].

La tension de seuil peut être modifiée en utilisant le couplage entre le canal et une électrode de contrôle arrière à travers un diélectrique, principe utilisé dans le transistor double grille à grilles indépendantes et le transistor SOI à BOX et film mince (UT2B : Ultra Thin Body and BOX), ou en polarisant la prise substrat comme dans le DTMOS (Dynamic Threshold voltage MOSFET). Ces trois autres exemples de dispositifs à tension de seuil dynamiquement modifiable sont présentés dans la Figure 4.1 (b), (c), (d).

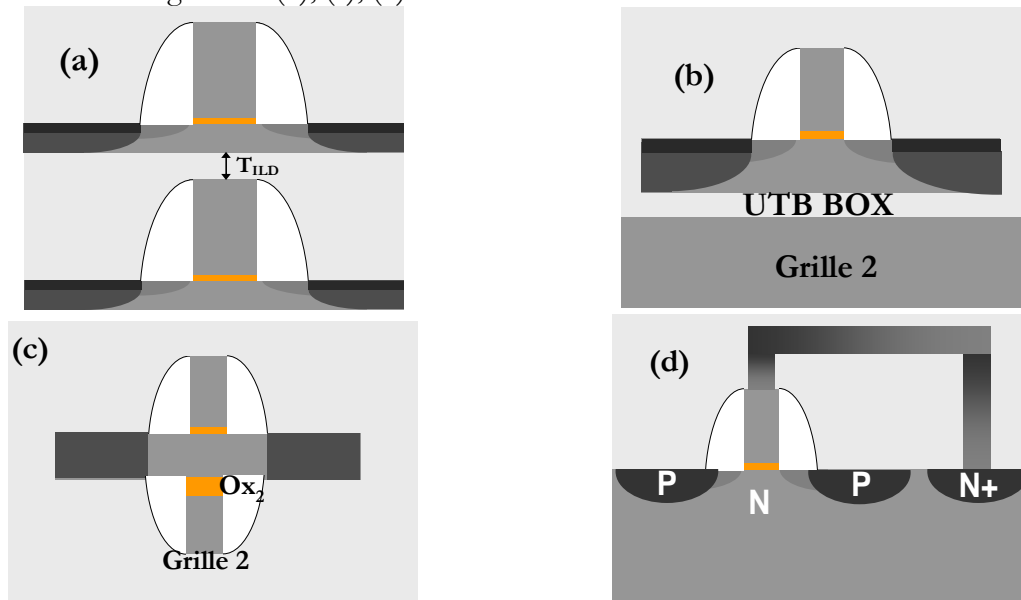


Figure 4.1 : Dispositifs permettant une modification dynamique de la tension de seuil (a) architecture 3D séquentielle avec diélectrique inter-niveau aminci (3DUTILD : Ultra Thin Inter Layer Dielectric) (b) Ultra Thin Body and BOX SOI (UT2B SOI) (c) Double gate MOSFET (DGMOS) (d) Dynamic Threshold voltage MOSFET (DTMOS)

Nous proposons d'utiliser la grille du transistor inférieur comme électrode de contrôle de la tension de seuil du transistor supérieur. Cette architecture se rapproche de celles du DGMOSFET et de l' UT2B SOI MOSFET car le contrôle de la tension de seuil est obtenu par couplage à travers un diélectrique. Elle diffère en ce que la grille inférieure n'a pas seulement pour rôle de modifier la tension de seuil du transistor supérieur mais aussi celui de faire commuter le transistor inférieur. Ainsi, elle fait intervenir deux transistors et permet donc de gagner en densité par rapport aux architectures planaires à modification de tension de seuil (cf Figure 4.2).

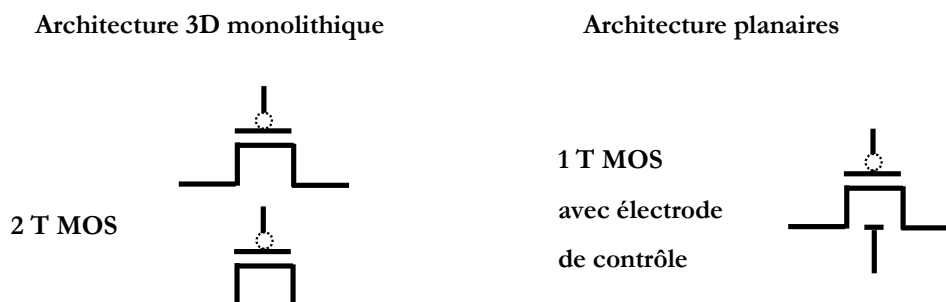


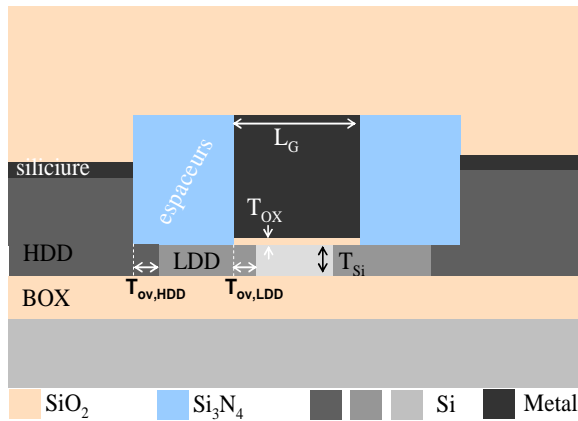
Figure 4.2 Mise en évidence de la potentialité de gain en densité avec l'architecture 3D séquentielle et diélectrique inter-niveau aminci. La grille contrôlant la tension de seuil du transistor supérieur permet aussi de faire commuter le transistor inférieur. Dans le cas du double grille, et de l'UT2B SOI, le contact face arrière ayant pour unique rôle de modifier la tension de seuil du transistor va augmenter la surface occupée par le transistor.

L'architecture 3D séquentielle avec un diélectrique inter-niveau suffisamment aminci pour modifier de manière significative la tension de seuil du transistor supérieur sera dénommée dans l'ensemble du texte par 3D UTILD pour Ultra Thin Inter Layer Dielectric.

L'objectif de cette partie est d'étudier la pertinence de l'architecture 3D UTILD pour prétendre à des applications avec modification de la tension de seuil dynamique. Pour cela nous avons étudié les paramètres influençant le couplage dans une architecture 3D séquentielle, à savoir : l'épaisseur du diélectrique inter-niveau, l'évolution du couplage en fonction de la taille des deux transistors superposés et l'influence du désalignement entre les deux transistors.

A. 2 - Dimensionnements pour obtenir un couplage conséquent

Cette étude a été réalisée avec le logiciel de simulation par éléments finis ATLAS de SILVACO. La structure du transistor simulée (cf Figure 4.3) correspond à une architecture de transistor FDSOI. L'objectif de cette étude n'est pas de simuler précisément un transistor FDSOI tel qu'il est réalisé au LETI mais d'étudier les interactions entre deux transistors empilés. Ainsi, un certain nombre d'hypothèses simplificatrices vis-à-vis de la technologie réelle peut être fait. Le dopage du transistor est réalisé à l'aide de 'boite de dopants', c'est-à-dire des régions avec une concentration uniforme de dopants : (une région canal non dopée, des régions faiblement dopées sous les espaceurs (LDD) et des régions fortement dopées (HDD)). Les accès sont surélevés au niveau des HDD (mais pas au niveau des LDD comme sur la technologie FDSOI) et les siliciures ont une résistivité nulle. Les espaceurs sont de forme rectangulaire. La grille est considérée comme entièrement métallique (au lieu d'une superposition de TiN, de polysilicium et de siliciure). L'oxyde de grille est en SiO₂ et dimensionné en épaisseur équivalente (au lieu d'une superposition SiO₂ et HfO₂).



T_{Si}	7 nm	Dopage canal	$1 \cdot 10^{15} \text{ cm}^{-3}$
T_{ox}	1 nm	Dopage LDD	$1 \cdot 10^{19} \text{ cm}^{-3}$
$T_{ov,LDD}$	6 nm	Dopage HDD	$1 \cdot 10^{20} \text{ cm}^{-3}$
$T_{ov,HDD}$	6 nm	V_{DD}	1 V

Figure 4.3 (a) : Description du MOSFET FDSOI simulé sous SILVACO (b) Paramètres constants dans l'ensemble des simulations

L'intégration 3D est étudiée pour répondre aux spécifications du nœud 45 nm. Dans cette étude, la longueur minimale sera de 35 nm. En choisissant une épaisseur de film de 7 nm, nous nous plaçons dans les conditions d'un transistor totalement dépleté et les effets quantiques apparaissant pour des épaisseurs de films inférieures à 7 nm peuvent être négligés [Widiez05]. Le modèle de 'CVT de Lombardi' est utilisé pour la mobilité et le modèle 'dérive-diffusion' pour le transport de charge.

Afin d'étudier la modification des propriétés électriques, due à la proximité des deux transistors empilés, les caractéristiques I_D-V_G du transistor supérieur lorsque la grille du transistor inférieur est polarisée de 0 à V_{DD} ont été simulées. Le type du transistor inférieur (p/n) n'influence pas ces caractéristiques quand l'électrode de grille est siliciurée, le travail de sortie de la grille inférieure perçu par le transistor supérieur est donc constant, que le transistor inférieur soit n ou p (cf Figure 4.4). Les deux transistors empilés ont la même longueur de grille dessinée ≈ 35 nm.

La Figure 4.5 présente les réseaux de courbes I_D-V_G du transistor pMOS supérieur, en fonction de la polarisation de la grille du transistor inférieur, pour des épaisseurs de diélectrique inter-niveau de 10, 50 et 150 nm.

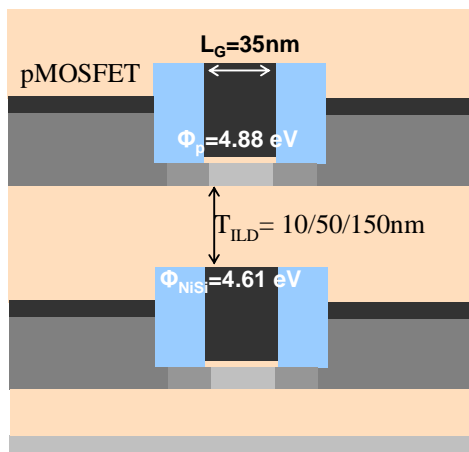


Figure 4.4 Description de l'empilement utilisé dans les simulations SILVACO menant aux simulations de la Figure 4.5 .

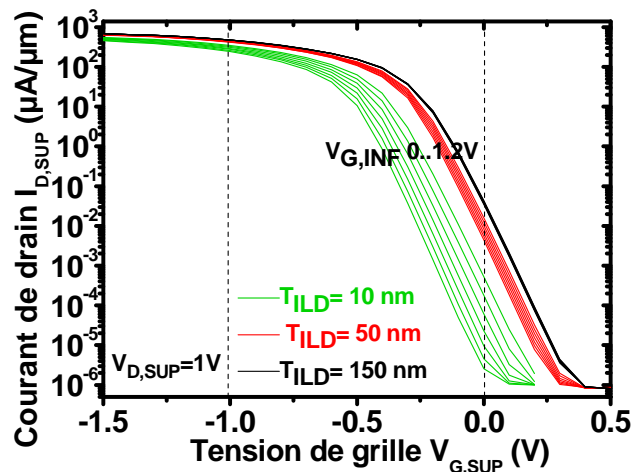


Figure 4.5 Evolution des caractéristiques I_D-V_G du transistor supérieur en régime saturé en fonction de la polarisation de grille du transistor inférieur pour des

épaisseurs de T_{ILD} de 10, 50 et 150 nm.

Nous observons une modification notable de l'aspect général des courbes pour $T_{ILD} \leq 50$ nm. La Figure 4.6, permet d'apprécier quantitativement les modifications des paramètres principaux du transistor supérieur (V_{TH} , I_{ON} , I_{OFF} , SS) en fonction de la polarisation de la grille inférieure.

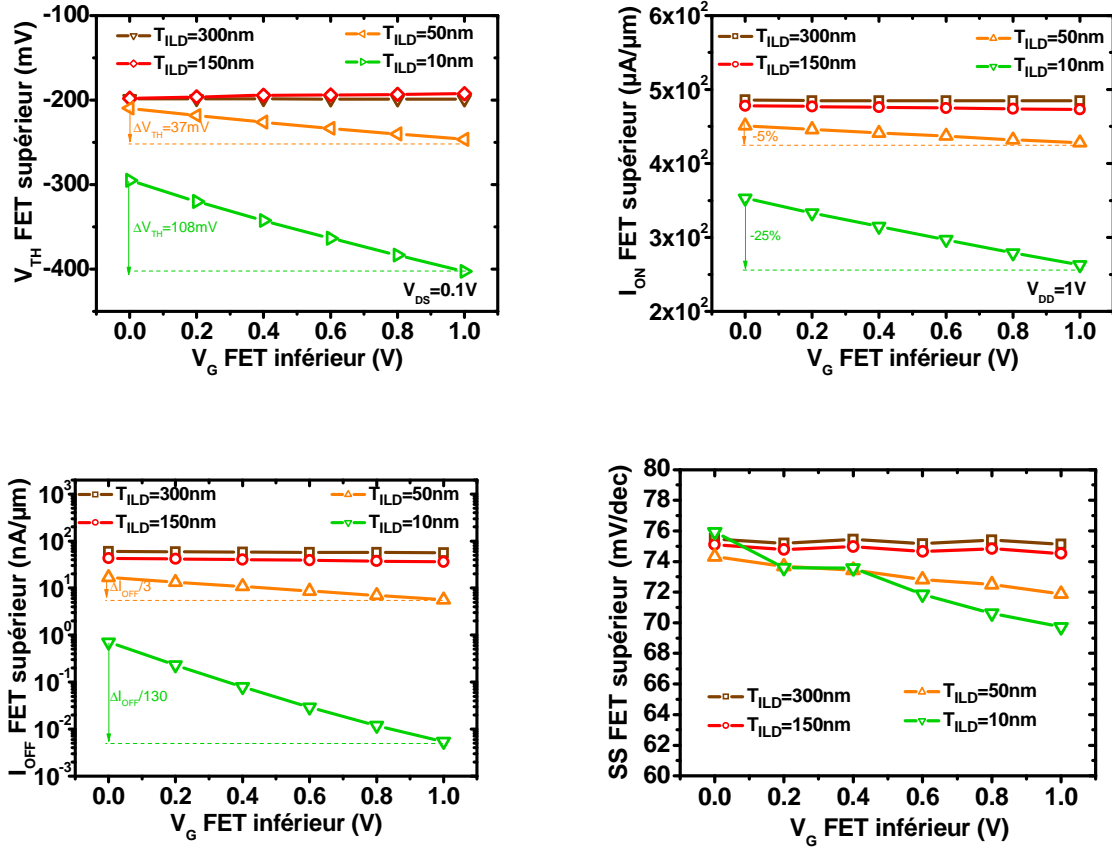


Figure 4.6 Evolution des paramètres électriques principaux du transistor supérieur (cas p-FET) avec la polarisation de la grille inférieure pour différents T_{ILD} . Evolution (a) de la tension de seuil V_{TH} , (b) du courant de conduction I_{ON} , (c) du courant de fuite I_{OFF} , (d) de la pente sous le seuil SS.

Nous observons que pour $T_{ILD} = 300$ et 150 nm, la modification des paramètres électriques est très faible, *a contrario*, la réduction de l'épaisseur du diélectrique inter-niveau $T_{ILD} = 50$ et 10 nm engendre une modification des paramètres électriques non négligeable.

Nous choisisons de définir deux épaisseurs critiques, la première T_{ISOLE} , pour laquelle le couplage peut être considéré comme négligeable et la deuxième $T_{COUPLAGE}$, pour laquelle le couplage est suffisant pour prétendre à des applications avec modification de tension de seuil.

L'épaisseur T_{ISOLE} est atteinte pour une modification de tension de seuil du transistor supérieur, inférieure à $1\%V_{DD}$, lorsque la grille basse passe de 0 à V_{DD} ($\Delta V_{TH, SUP(0/V_{DD})}$). Ce critère est arbitraire mais correspond à une valeur inférieure aux spécifications de la variabilité locale de la tension de seuil qui est de $3\%V_{DD}$. Dans le cas étudié, cette épaisseur se situe aux alentours de $T_{ILD} = 100$ nm.

Nous considérons que le couplage est suffisant pour prétendre à des applications avec modification de la tension de seuil lorsque $\Delta V_{TH, SUP(0/V_{DD})}$ est supérieur ou égal à 100 mV. Dans le cas de la structure particulière étudiée, $T_{COUPLAGE}$ est aux alentours de 10 nm- 20 nm.

Le MOSFET inférieur, lui, n'est pas influencé par les polarisations appliquées au transistor supérieur. En effet, ces caractéristiques I_D-V_G sont invariantes lorsque l'on polarise les électrodes du FET supérieur de 0 à V_{DD} et ce, jusqu'à une épaisseur de diélectrique inter-niveau de 10 nm.

La Figure 4.5 présente les réseaux de courbes I_D-V_G du transistor nMOS supérieur en fonction de la polarisation de la grille du transistor inférieur pour des épaisseurs de diélectrique inter-niveau de 10, 50 et 150 nm.

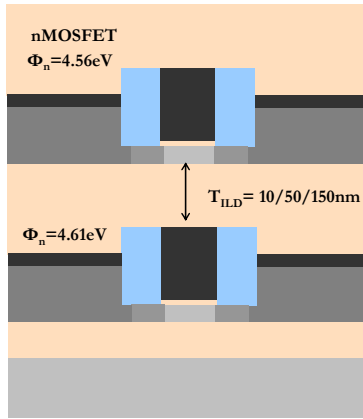


Figure 4.7 Description de l'empilement utilisé dans les simulations SILVACO menant aux simulations de la Figure 4.8

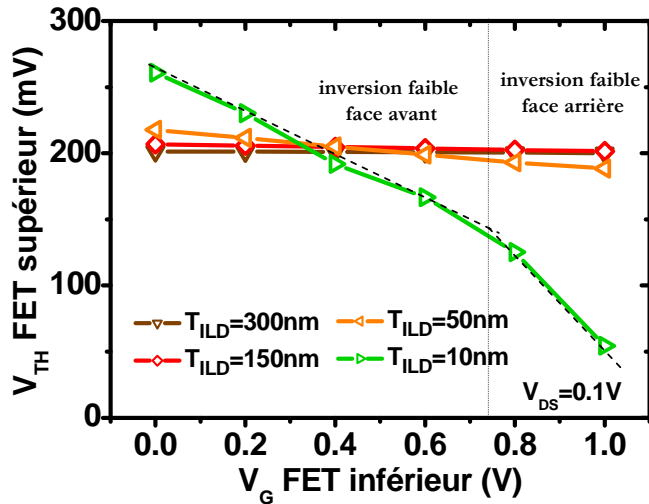


Figure 4.8 Evolution de la tension de seuil du nMOS supérieur avec la polarisation de la grille inférieure pour des épaisseurs de diélectrique inter-niveau de 10, 50, 150 et 300 nm.

La gamme d'épaisseur à partir de laquelle les paramètres électrique sont modifiés de manière significative reste la même que dans le cas du pMOS (~100nm). Cependant, nous observons une plus grande modification de la tension de seuil avec la polarisation de la grille inférieure, (par exemple dans le cas $T_{ILD}=10$ nm, $\Delta V_{TH, SUP(0/VDD)} = 206$ mV dans le cas du nMOS et 108 mV dans le cas du pMOS). Sur la courbe nous distinguons deux régimes : Un régime à basse tension de grille inférieure avec une pente faible et un régime à plus forte tension avec une pente plus importante. Ceci s'explique par le fait qu'à plus forte tension, c'est l'interface arrière qui contrôle la tension de seuil.

Afin d'obtenir l'amplitude de la modification de la tension de seuil voulue, il faudra donc dimensionner le diélectrique inter-niveau en fonction des positions relatives des travaux de sortie.

A. 3 -Evolution en fonction du désalignement et cas avec grilles empilées de longueurs différentes

Puisque les deux transistors ne sont pas auto-alignés, des désalignements entre grille haute et basse vont apparaître. L'évolution du contrôle du désalignement pour les parties logiques prévu par l'ITRS [ITRS lithography] est représentée dans la table suivante :

MPU/ ASICS Metal 1 (M1) 1/2 pitch (nm)	68	45	32
Overlay (3σ) (nm)	17	11	8

Table 4.1 : Extrait de l'ITRS 2007 des valeurs du désalignement entre deux couches à 3 σ.

Pour le nœud 45 nm, le désalignement à 3σ est d'environ 10 nm [ITRS section lithographie]. Un tel désalignement va modifier les caractéristiques électriques du transistor supérieur. L'objectif est de quantifier cette source de variabilité afin de savoir si elle est rédhibitoire pour l'utilisation du couplage dans les structures 3D séquentielles.

D'autre part, dans un circuit intégré, les longueurs des grilles des transistors ne sont pas systématiquement dessinées à la dimension minimale autorisée dans le DRM. Par exemple dans les zones de mémoire SRAM, les différents transistors sont dessinés avec des longueurs de grille différentes afin d'optimiser les marges de stabilité¹. Par exemple, pour le nœud 45 nm, les transistors de la zone logique sont dessinés à la taille minimale ($L_G \sim 35\text{nm}$) alors que les transistors des cellules SRAM auront des longueurs de grilles différentes au sein de la même cellule (typiquement des longueurs de variant entre 45 à 65 nm). En conséquence, le scénario où les transistors empilés ont des tailles différentes est à étudier.

Les simulations suivantes permettent de quantifier l'influence de la modification de la taille du transistor inférieur ainsi que de l'alignement entre les deux grilles empilées. La longueur de la grille du transistor supérieur reste constante et égale à 45 nm, alors que la longueur de la grille inférieure prend les valeurs 45, 55 et 65 nm. L'épaisseur du diélectrique inter-niveau est de 10nm.

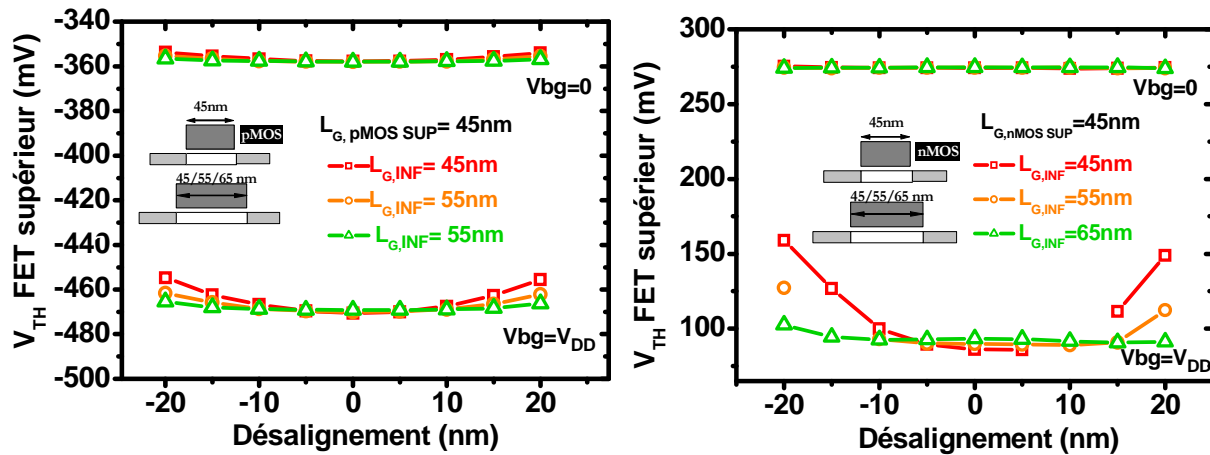


Figure 4.9 : Evolution de la tension de seuil d'un transistor supérieur ($L_G=45\text{ nm}$) en fonction du désalignement entre les deux grilles des transistors empilés pour deux conditions de polarisation de la grille inférieure de dimension variable (45/55/65nm) (a) le transistor supérieur est de type p (b) le transistor supérieur est de type n.

¹ Ce point sera détaillé dans la partie IV.B consacrée aux mémoires SRAM.

Nous observons que le désalignement a une influence sur les performances du transistor supérieur. Dans le cas où les deux transistors empilés ont tous deux la même longueur de grille (45 nm), la modification maximale de tension de seuil est de 8mV dans le cas d'un pMOS et de 40 mV dans le cas d'un nMOS pour un désalignement de 15 nm. Dans le cas du nMOS, cette modification dépasse la variabilité locale de la tension de seuil tolérée ($3\%V_{DD}$). Il est donc nécessaire de chercher à limiter cette contribution supplémentaire à la variabilité totale.

Le désalignement affecte plus le transistor nMOS où l'influence du couplage dans la position alignée est originellement plus important ($\Delta V_{TH, SUP(0/V_{DD})}$ vaut 185mV dans le cas du nMOS et 112 mV dans le cas du pMOS).

D'autre part, on observe aussi qu'en agrandissant la grille inférieure de 20 nm par rapport à la grille supérieure, la modification de tension de seuil entre la position alignée et désalignée de 15 nm passe de 8 mV à 2.6mV pour le pMOS et de 40 à 1mV pour le nMOS.

En conséquence, pour mieux contrôler la variabilité apportée par ce désalignement, nous pouvons :

- Optimiser les valeurs de travaux de sortie des deux grilles afin d'atteindre la modification de tension de seuil souhaitée et ne pas dépasser cette valeur. Plus le couplage est fort, plus l'influence du désalignement se fait sentir.
- Aggrandir la longueur de grille du transistor inférieur.

Les simulations suivantes permettent de quantifier l'influence de la modification de la taille du transistor supérieur ainsi que de l'alignement entre les deux grilles empilées. La longueur de la grille du transistor inférieur reste constante et égale à 45 nm, alors que la longueur de la grille du transistor supérieur (type n) prend les valeurs 45,55 et 65 nm. L'épaisseur du diélectrique inter-niveau est de 10nm.

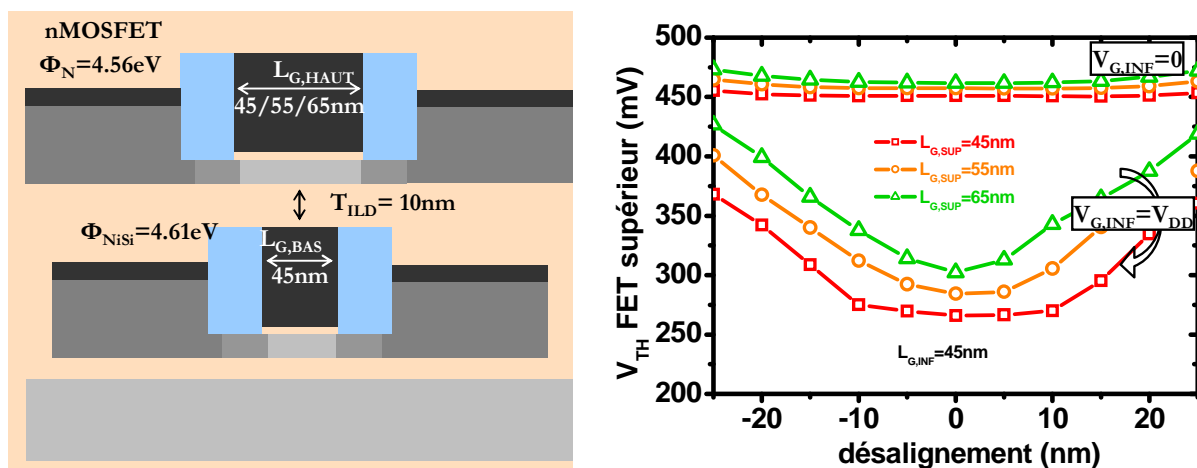


Figure 4.10 Evolution de la tension de seuil pour un nMOSFET de longueur 45/55/65 nm empilé sur un transistor de longueur constante (45nm).

Nous observons que cette configuration est plus défavorable en termes de variabilité car elle est très sensible au désalignement (modification d'environ 60mV pour un désalignement de 15 nm dans la configuration n-FET supérieur de 65nm empilé sur FET de 45nm de longueur de grille). Il sera donc indispensable de choisir avec attention les emplacements des transistors au sein d'une cellule SRAM en tenant compte de ce point particulier.

A. 4 -La réalité technologique

A.4.1.a - Influence du désalignement dans le sens de la largeur de grille

Dans la partie précédente, nous avons étudié l'influence du désalignement dans le sens de la longueur de grille (axe Y représenté sur la Figure 4.11), mais un désalignement peut aussi apparaître dans le sens de la largeur de grille (axe X). Dans ce cas, l'utilisation des dépassements de grilles par rapport à la zone active peuvent rendre la structure insensible au désalignement selon cet axe. En effet, en planarisant le niveau poly inférieur comme décrit dans la Figure 4.11 (b), la distance entre la grille inférieure et la zone active supérieure reste constante, que la zone active soit bien alignée au dessus de la zone active supérieure ou qu'elle soit désalignée et soit placée au dessus de la zone d'isolation.

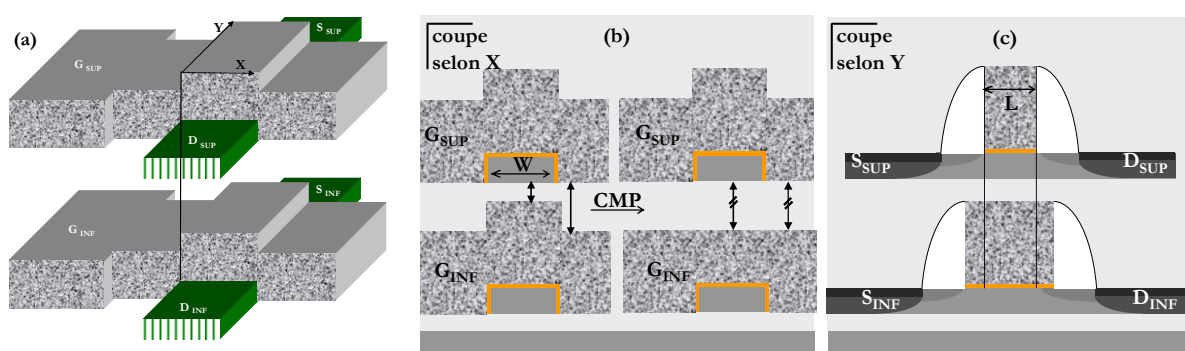


Figure 4.11 (a) Définition des axes pour décrire les désalignements potentiels dans l'architecture 3D monolithique. (b) Description d'une solution pour supprimer l'influence du désalignement dans le sens perpendiculaire au canal. (i.e. Planarisation de la topographie de la grille) (c) Description d'une solution permettant de limiter l'impact du désalignement selon l'axe parallèle au canal (i.e. élargissement de la grille inférieure)

A.4.1.b - Impact du contrôle de l'épaisseur du diélectrique inter-niveau

Nous avons vu au cours de cette étude que pour obtenir un couplage conséquent entre transistor supérieur et inférieur, le diélectrique inter-niveau doit être relativement fin. Dans le cas de la stabilisation de la SRAM à quatre transistors et à six transistors qui sera décrit par la suite, le diélectrique a été dimensionné à 10 nm et 20 nm respectivement. Mais ces épaisseurs faibles sont-elles contrôlables avec une précision suffisante ? Quelle variabilité supplémentaire le contrôle de l'épaisseur de l'ILD va-t-elle apporter ? Il y a-t-il des solutions technologiques pour limiter les variations d'épaisseur ?

Nous nous reportons au Chapitre I sur le développement des briques technologiques. L'épaisseur finale du diélectrique inter-niveau se décompose en deux contributions, $T_{ILD, BAS}$ et $T_{ILD, HAUT}$ (Figure 4.12). Dans l'intégration réalisée dans cette thèse, ces deux épaisseurs sont contrôlées par polissage mécano-chimique (CMP : Chemical mechanical Polishing). Or, le contrôle d'épaisseur par CMP conduit à une variation importante de celle-ci sur l'ensemble de la

plaque. Par exemple, lors de l'étape de planarisation de la partie basse du diélectrique inter-niveau précédant l'étape de collage, un amincissement d'environ 250nm est nécessaire. Suite à cette étape une variation d'épaisseur de $T_{ILD, BAS}$ d'environ 30nm sur l'ensemble de la plaque est observée. Une telle variation rend ces procédés inutilisables pour la réalisation d'ILD ultrafine dans l'intégration 3D monolithique.

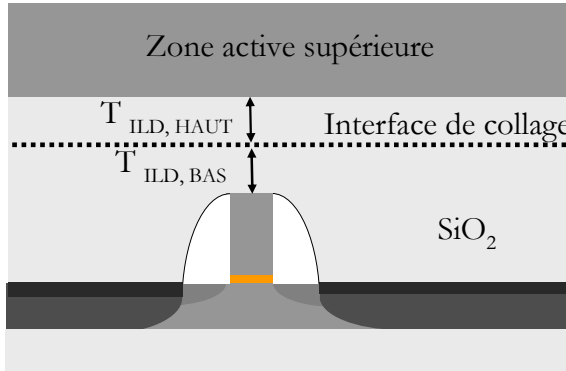


Figure 4.12 Décomposition de l'épaisseur du diélectrique inter-niveau en deux contributions.

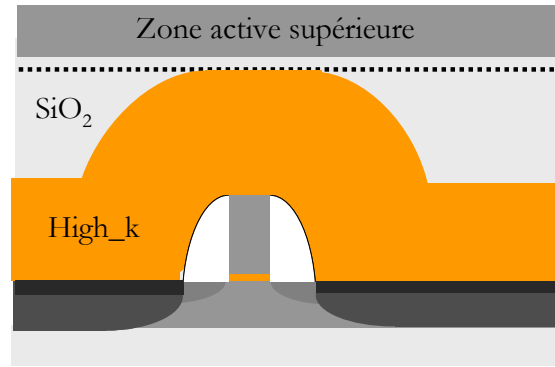


Figure 4.13 Utilisation de high-k pour satisfaire à un couplage suffisant entre les deux niveaux de transistors avec une variabilité améliorée par rapport à l'architecture avec ILD en SiO_2

Afin de répondre aux exigences en termes de contrôle de l'épaisseur de diélectrique inter-niveau, nous proposons différentes options d'intégration.

Pour le contrôle de $T_{ILD, BAS}$, on pourrait envisager de remplacer l'étape de planarisation et d'amincissement par CMP par une planarisation à l'aide d'une résine HSQ suivi d'un amincissement chimique. Les résines HSQ sont des solutions liquides qui, étalées par tournette permettent de supprimer la topographie, après recuit, elles conduisent à la formation d'un diélectrique (classiquement le SiO_2). Cette étape permet de supprimer la topographie du niveau transistor inférieur.

Une autre option pour relâcher les contraintes sur le contrôle de l'épaisseur du diélectrique inter-niveau, peut être d'utiliser un high-k dans la zone où le couplage est requis comme présenté dans la Figure 4.13. Par exemple, un empilement high-k/ SiO_2 peut être réalisé puis planarisé en utilisant le high-k comme couche d'arrêt. Une épaisseur équivalente de SiO_2 de 10nm conduit à une épaisseur de high-k d'environ 50 nm avec une constante de diélectrique de 20. Le contrôle de son épaisseur est alors moins critique. En effet, on observe dans les résultats de simulation présentés dans la Figure 4.14 que pour une EOT de 10 nm visée en utilisant du SiO_2 , une variation d'épaisseur physique de 1 nm va engendrer une modification de tension de seuil d'environ 20 mV. En utilisant un high-k de constante diélectrique 20, le contrôle de la dimension du diélectrique est plus relâché et supporte une différence d'épaisseur physique de 5 nm par rapport à la valeur physique visée afin de contrôler la tension de seuil à 20mV près.

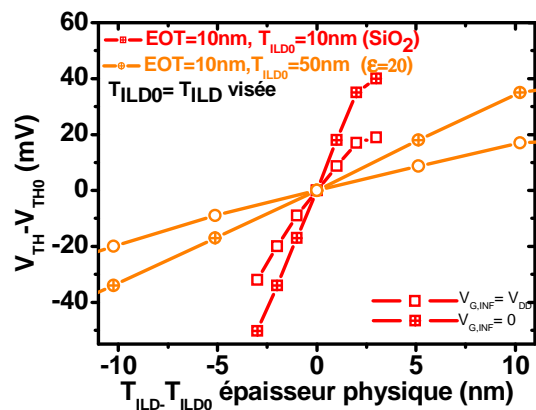


Figure 4.14 Impact de la variation d'épaisseur sur le contrôle de la tension de seuil, pour $T_{ILD} = 10/20$ nm. Evolution pour le contrôle d'épaisseur pour un high-k.

Enfin, si T_{ILD} est augmenté, les variations d'épaisseur auront une influence plus faible sur le contrôle de la tension de seuil.

Limiter la contribution $T_{ILD, haut}$ est moins complexe, car cela n'implique pas de planarisation d'une topographie. Ce diélectrique peut être réalisé par oxydation thermique, permettant de contrôler les épaisseurs formées au nm près, ou encore complètement supprimé en développant des techniques de collage direct SiO_2/Si ou High-k/Si.

Ce qu'il faut retenir

IV.A : La structure 3D UTILD

- Il existe une gamme d'épaisseur de diélectrique inter-niveau pour laquelle les caractéristiques électriques statiques du transistor supérieur ne sont pas modifiées. Pour les nœuds sub-45 nm, le couplage peut être négligé pour des épaisseurs de diélectrique inter niveau supérieur à 100nm.
- Pour satisfaire aux applications avec modification de tension de seuil, la modification de tension de seuil doit être de l'ordre 100mV. Dans le cas de l'architecture 3D monolithique et les règles de dessin du nœud 45 nm, cela conduit à une épaisseur de diélectrique inter-niveau de l'ordre de 10-20 nm (en EOT).
- L'évolution des paramètres électriques du FET supérieur dépend de sa propre longueur de grille, de la longueur de grille du transistor inférieur, des travaux de sortie relatifs des grilles haute et basse, du désalignement. Tous ces paramètres doivent être pris en compte pour modéliser correctement le comportement du transistor lors de l'établissement d'un modèle compact.
- Les caractéristiques électriques du transistor du niveau inférieur ne sont pas modifiées par la proximité du transistor supérieur
- Le désalignement et le contrôle de l'épaisseur du diélectrique inter-niveau sont des sources de variabilité mais ces contributions peuvent être limitées par des options technologiques ou de conception et ne semblent pas rédhibitoires pour l'utilisation de l'architecture 3D UTILD.
 - Pour limiter l'influence d'un désalignement dans le sens de la longueur de grille sur les valeurs de tension de seuil, nous pouvons augmenter la longueur de grille inférieure.
 - Pour supprimer l'influence du désalignement dans le sens de la largeur de grille, nous pouvons supprimer la topographie de la grille au dessus de la zone active à l'aide d'une étape de CMP.
- La variabilité liée au contrôle de l'épaisseur de diélectrique inter-niveau peut avoir une influence réduite en le réalisant par exemple avec un high-k.

B - Les mémoires SRAM

B.1 - Les mémoires SRAM 3D

Une forte demande existe pour obtenir des SRAM haute densité dans tous les domaines d'applications de la SRAM. Dans les circuits intégrés pour des processeurs, la place occupée par les SRAMs avoisine souvent les soixante pourcents de la taille de la puce pour des microprocesseurs et les quatre-vingt pourcents de la taille de la puce pour des SoC. La cellule SRAM la plus classique comporte six transistors, ce qui fait d'elle une mémoire peu dense comparée à une DRAM par exemple. Afin de surpasser ce handicap, les 6 transistors de la cellule ont été répartis sur plusieurs étages dans des intégrations 3D séquentielles [Takao92] [Liu02] [Jang04] [Jung04] [Jung05] [Jung07] [Son07]. Dans les études les plus abouties, la cellule 6T est intégrée sur 3 niveaux comme représentée dans la Figure 4.15. Les 3 paires de transistors décrits dans la Figure 4.15 : transistor d'accès (TA pour Access), de conduction (TD pour drive) et de charge (TL pour Load) sont placés sur des niveaux distincts.



Figure 4.15 Observation TEM d'une coupe de SRAM empilée sur 3 niveaux [Jung07]

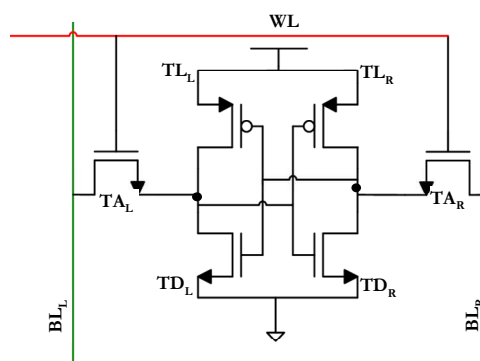


Figure 4.16 Représentation schématique d'une SRAM à 6 transistors.

La surface de la cellule SRAM 6T ainsi obtenue est de $0.16\mu\text{m}^2$ avec une longueur de grille minimum de 65 nm [Jung 05]. La surface occupée par cette SRAM 3D est équivalente à celles obtenues avec des technologies de type 32 nm en planaire comme le démontre la table récapitulative des démonstrations de SRAM de la littérature (cf table 4.1). Pourtant la longueur de grille utilisée place aussi cet exemple dans la catégorie nœud 90 nm.

Référence	Nœud (nm)	$L_{G,MIN}$ (nm)	Surface (μm^2)	Surface/ (noeud ²)
Intel [Ghanie03]	90	45	1	123
ST [Devoivre02]	90	70	1.36	168
Samsung [Jung 05]	80	65	0.16	25
Intel [Tyagi05]	65	35	0.57	135
IBM [Luo04]	65	43	0.625	148
ST [Arnaud04]	65	45	0.5	118
Intel [Auth08]	45	35	0.346	171

Toshiba[Nii06]	45	34	0.248	122
TSMC[Cheng07]	45	30	0.242	120
ST[Josse06]	45	35	0.25	123
Intel [Natarajan 08]	32	30	0.171	167
Toshiba[Hasegawa 08]	32	30	0.124	121
TSMC [Wu07]	32	30	0.15	146
IBM/ ST[Arnaud 08]	32	28	0.157	153
Samsung [Jung 05]	80	65	0.16	25
IBM [Haran08]	22	25	0.1	207

Table 4.1 : Benchmark SRAM pour les nœuds 90-32 nm.

Ainsi, empiler la SRAM sur trois niveaux permet de gagner environ 3 nœuds technologiques en terme de densité (passage d'une technologie 80 nm à 32 nm) et ce sans changer de technologie. Garder des dimensions de transistors relâchées permet de réduire les coûts de développement et de fabrication via l'utilisation d'équipements des générations précédentes et de limiter la variabilité (effets canaux courts plus faibles).

Les gains obtenus en utilisant l'intégration 3D monolithique font de la SRAM une application phare de cette technologie.

B. 2 -Les problématiques de conception de cellules SRAMs

L'augmentation des fluctuations des paramètres du procédé de fabrication (L , W , V_{TH} , μ_0 , N_a , T_{OX} , ...) et la réduction de la tension d'alimentation avec la réduction des dimensions augmentent de plus en plus la sensibilité des cellules mémoires SRAMs aux différentes sources de bruit.

Pour réduire la surface des cellules, de nombreuses règles de dessin sont violées. Malgré la pression pesant sur la réduction de la surface de la cellule, les dimensions des transistors (L et W) ne correspondent pas généralement aux règles minimales afin de limiter l'impact de la variabilité sur la stabilité des cellules.

Lorsqu'une cellule SRAM est dimensionnée, les caractéristiques recherchées sont :

- Une stabilité adéquate au bon fonctionnement des diverses opérations (stabilité en lecture, caractérisée par la SNM (Static Noise Margin), stabilité en rétenion caractérisée par la RNM (Retention Noise Margin))
- Une marge en écriture suffisante pour des opérations d'écriture efficaces WM (Write Margin)
- Un courant de conduction (I_{CELL}) maximal pour atteindre de bonnes performances,
- Un courant en rétenion (I_{OFF}) minimal pour minimiser la puissance consommée statique,
- Une taille minimale de cellule

Globalement, l'ensemble de ces critères sollicite des dimensionnements (choix de L et W) de la cellule antagonistes. Sa conception relève alors d'un ensemble de compromis qui est arbitré par l'application visée.

Actuellement la stabilité des cellules correspond au critère le plus problématique dans la conception des cellules SRAMs. La sensibilité des marges de stabilité en lecture, rétenion et marge en écriture aux variations des paramètres technologiques est bien mise en évidence par ces mesures issues de [Bhavnagarwala05] sur les 512 cellules appartenant au même plan de mémoire.

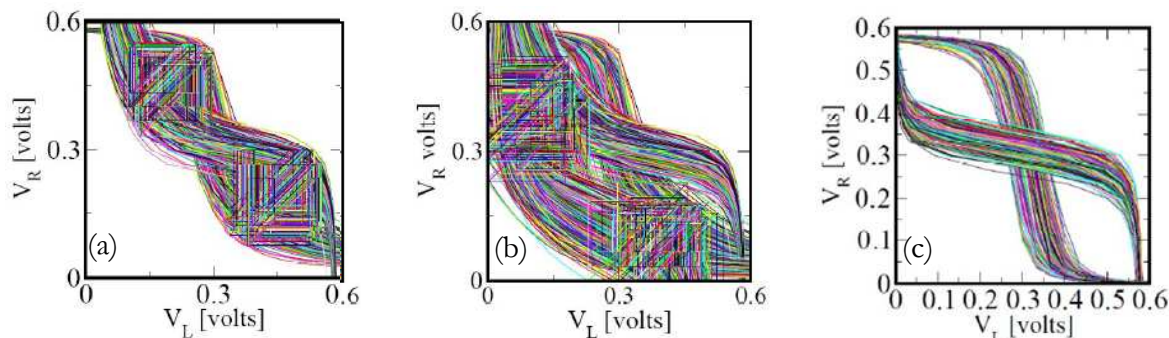


Figure 4.17 Caractéristiques statiques en tension pour une matrice de 512 cellules SRAM6T dans le mode (a) lecture (b) écriture (c) rétention illustrant l'impact de la variabilité sur la stabilité des SRAMs. [Bhavnagarwala05]

Pour augmenter les marges de stabilité, nous proposons d'utiliser le couplage entre les deux transistors empilés dont nous présenterons deux modes de réalisation de SRAMs. Pour chacune d'entre elles, nous présenterons d'abord les problématiques de stabilité de la cellule mémoire, puis le positionnement des transistors, enfin, les performances des cellules ainsi que le gain en densité correspondants.

B. 3 -La SRAM4T 3D UTILD

B.3.1 - Présentation de la SRAM 4T et analyse de sa stabilité

Possédant seulement quatre transistors, la SRAM 4T est très attractive en termes de densité (typiquement 30 à 40% de moins que la classique 6T). Malheureusement, les cellules 4T présentent des problèmes d'instabilité [Takeda00] qui impliquent l'utilisation de techniques de rafraîchissement ou d'un bloc de contrôle des données en rétention. La Figure 4.18 représente la vue schématique d'une cellule 4T dite load-less (les transistors de charge sont absents par rapport à la SRAM 6T). Cette cellule est composée de deux types de transistors, les transistors d'accès de type p, et les transistors de conduction de type n. La cellule est symétrique, on désignera les transistors de droite par l'indice R (pour Right) et de gauche par indice L (pour Left). $BL_{L/R}$ désigne la ligne de Bit (BitLine Left/Right) et WL , la ligne de mot (Word Line).

Le dimensionnement et la tension de seuil des deux transistors au sein d'une paire (transistors d'accès ou transistors de drive) sont strictement égaux, cela permet que le comportement de la cellule (stabilité) soit symétrique lorsqu'un « 1 » ou un « 0 » est mémorisé dans la cellule.

Pour décrire les problématiques de stabilité, nous nous plaçons par exemple dans le cas où un « 1 » est mémorisé (i.e. $L= '1'$ and $R= '0'$).

Selon le régime dans lequel la stabilité est analysée, (rétention ou lecture), les transistors critiques pour le maintien de l'information diffèrent. La Figure 4.18 met en évidence les transistors intervenant dans la stabilité pour ces deux modes.

Pendant la période de rétention, les lignes de bit sont pré chargées à V_{DD} et les lignes de mot maintenues à V_{DD} pour bloquer les transistors d'accès. Dans ces conditions de polarisation, le nœud le plus instable est le nœud $L='1'$ (puisque le nœud de stockage R est maintenu à 0 grâce au TD_R qui est passant). Pour que la donnée soit maintenue, le transistor d'accès TA_L doit fuir

suffisamment pour maintenir le nœud L à V_{DD} . En conséquence TA_L doit avoir un courant de fuite I_{OFF} supérieur à celui du transistor TD_L et du courant de grille de TD_R , ce qui donne

$$I_{TA_L-OFF} > I_{TD_L-OFF} + I_{TD_R-GRILLE} \quad (EQ 4.1)$$

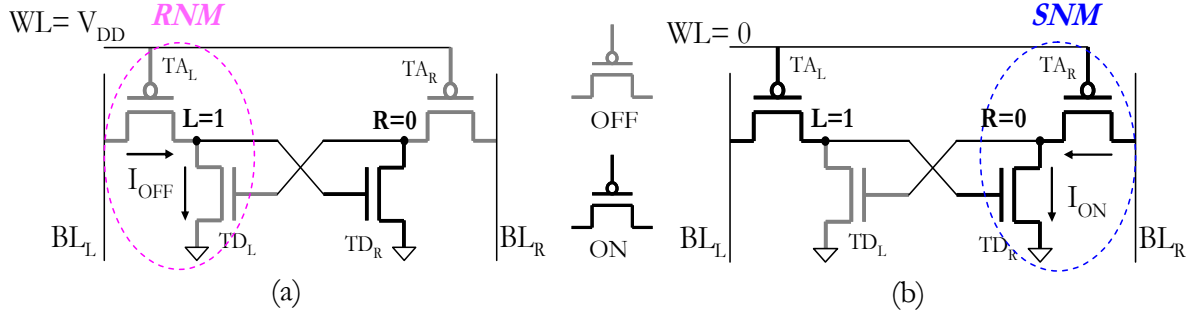


Figure 4.18 Représentation schématique des problématiques de stabilité d’une SRAM 4T (a) en mode rétention. La stabilité en rétention dépend du rapport des courants I_{OFF} des transistors TA_L et TD_L (b) en mode lecture, la stabilité en lecture dépend du rapport des courant I_{ON} des transistors TA_R et TD_R

Pendant la période de lecture, les lignes de bit sont pré chargées à V_{DD} et la ligne de mot est maintenue à 0 de manière à lire l’information contenue par la cellule. Dans ces conditions, le nœud le plus instable est le nœud $R=0$ (il est relié à la BL pré chargée à ‘1’ par TA_R qui est passant). Dans le but de maintenir le potentiel du nœud R proche de la masse, le transistor TD_R doit avoir un courant I_{ON} plus fort que celui de TA_R , ce qui donne:

$$I_{TA_R-ON} \leq I_{TD_R-ON} \quad (EQ 4.2)$$

Ces deux conditions sont *de facto* en conflit ; pour améliorer la marge en rétention, la tension de seuil des transistors d’accès doit être diminuée afin d’obtenir un fort courant de fuite, alors que, pour améliorer la marge en lecture la tension de seuil des transistors d’accès doit être augmentée afin d’obtenir un faible courant de conduction. En conséquence, seul un contrôle dynamique de la tension de seuil offre un moyen d’augmenter à la fois la marge de stabilité en lecture et en rétention.

B.3.2 - Positionnement des transistors en 3D

La configuration décrite dans la Figure 4.19 , utilisant le couplage entre niveau haut et bas pour modifier dynamiquement la tension de seuil permet l’amélioration des marges de stabilité en rétention et en lecture. En effet, placer TA_L au dessus de TD_L (dont la grille est polarisée à 0) va abaisser sa tension de seuil, donc augmenter le I_{OFF} de TA_L (Figure 4.19 courbe 1) conduisant à une augmentation de la stabilité en rétention. (cf EQ 4.1), et en positionnant TA_R au dessus de TD_R (dont la grille est polarisée à V_{DD}) va augmenter sa tension de seuil et en conséquence diminuer son I_{ON} (insert de la Figure 4.10 courbe 2) conduisant à une stabilité en lecture améliorée. (cf EQ 4.2). Remarquons qu’ainsi la stabilité en lecture et en rétention sera améliorée quelque soit la donnée stockée dans la cellule mémoire (‘0’ ou ‘1’)

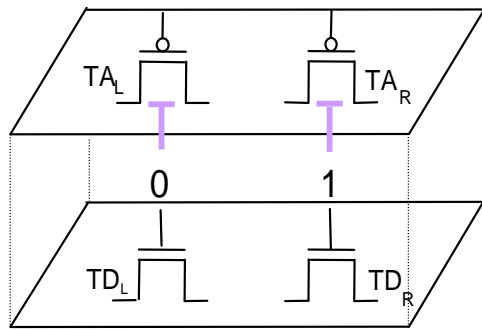


Figure 4.19 Positionnement des transistors dans la cellule SRAM4T à stabilité améliorée grâce à la technologie 3D séquentielle avec diélectrique inter niveau ultrafin.

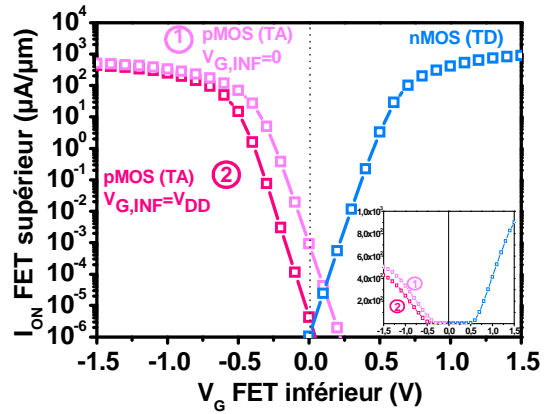


Figure 4.20 Courbes I_D - V_G des transistors de drive et des transistors d'accès en fonction de la polarisation de la grille inférieure obtenues par simulation. ($T_{ILD}=10\text{nm}$)

Comme le couplage modifie le courant de fuite exponentiellement, alors que le courant de conduction suit une dépendance quadratique, l'avantage de cette architecture se situe principalement dans le régime de rétention. Les tensions de seuil des transistors seront donc adaptées de sorte à augmenter majoritairement la RNM.

B.3.3 - Analyse des résultats

Gain en stabilité

La marge de stabilité correspond à la tension minimale à appliquer aux points de stockage de la cellule (i.e 'L' et 'R') nécessaire pour faire basculer la cellule dans l'état de mémorisation opposé (perte de l'information contenue initialement dans la cellule). L'extraction d'une marge de stabilité se fait en traçant l'évolution du potentiel du nœud L lorsque le potentiel du nœud R varie de 0 à V_{DD} et vice versa. Puis en symétrisant l'une des deux courbes par rapport à la première bissectrice, une courbe de type 'papillon' est alors obtenue. La marge de stabilité correspond à la mesure du côté du plus petit carré s'insérant entre les deux courbes comme représenté sur la Figure 4.21 .

La Figure 4.22 schématise les interconnexions et les polarisations utilisées pour extraire les marges de stabilité en lecture et en rétention grâce à la simulation.

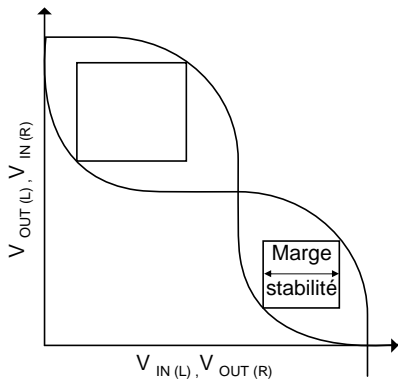


Figure 4.21 Représentation graphique des marges de stabilité, la marge de stabilité correspond à la SNM lorsque la cellule est polarisée en lecture (transistors d'accès passant) et RNM, lorsqu'elle est en rétention (transistors d'accès bloqués)

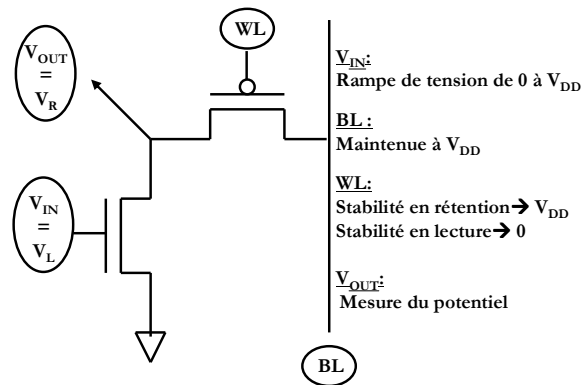


Figure 4.22 Schématisation des interconnexions et polarisations pour l'extraction des marges de stabilité en rétention et en lecture.

Les caractéristiques des transistors ont été choisies pour correspondre aux critères de l'ITRS pour le nœud 45nm et les applications LSTP. Plus précisément la tension de seuil des FETs a été ajustée pour obtenir un courant I_{OFF} autour de $10\text{pA}/\mu\text{m}$ pour une tension d'alimentation de 1.1V.

L'épaisseur du diélectrique inter-niveau est de 10 nm et permet une augmentation du courant de fuite des transistors d'accès d'un facteur 180 lorsque la grille du niveau inférieur passe de 0 à V_{DD} . Ce facteur est supérieur à celui nécessaire (valant typiquement 100) pour assurer une bonne stabilité en rétention [Thomas03].

La Figure 4.23 présente les courbes papillons correspondant aux simulations avec les paramètres récapitulés dans la Figure 4.24 . Une importante marge de stabilité en rétention (RNM=320mV) et une marge de stabilité en lecture plus faible (SNM=150 mV) sont obtenues. Ces valeurs sont atteintes pour une cellule dite « non dimensionnée » car les longueurs et largeurs des transistors des n et pMOSFETs sont égales ($W_n=W_p$ and $L_n=L_p$).

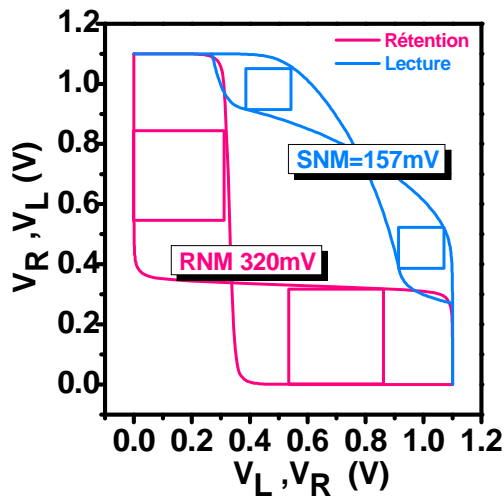


Figure 4.23 Courbes papillons obtenues par simulation en rétention et en lecture.

	V_{TH} (mV)	I_{ON} ($\mu A/\mu m$)	I_{OFF} ($\mu A/\mu m$)
pMOS au dessus GRILLE BAS à V_{DD}	-0.36	$2.6 \cdot 10^{-4}$	$1.5 \cdot 10^{-11}$
pMOS GRILLE BAS à 0	-0.23	$3.5 \cdot 10^{-4}$	$2.7 \cdot 10^{-9}$
nMOS	0.44	$5 \cdot 10^{-4}$	$1.3 \cdot 10^{-11}$

Figure 4.24 Récapitulation des caractéristiques des MOSFETs utilisés dans les simulations ($L_G=32nm$, $T_{ILD}=10nm$)

Afin d'équilibrer les marges de rétention et de lecture, la cellule doit être dimensionnée. Pour ce faire, il est nécessaire de réaliser des simulations en 3 dimensions afin d'évaluer l'influence de la variation des largeurs de grille W sur les stabilités. La Figure 4.25 représente la structure simulée en 3D, avec $W_p=80nm$ et $W_n=160nm$.

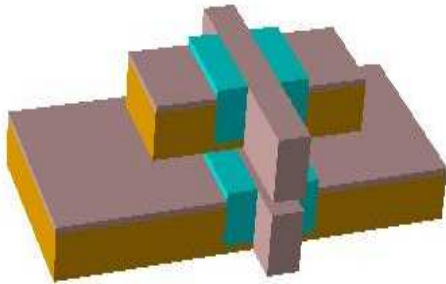


Figure 4.25 Structure en trois dimensions permettant de dimensionner la SRAM4T 3D UTILD avec ATLAS

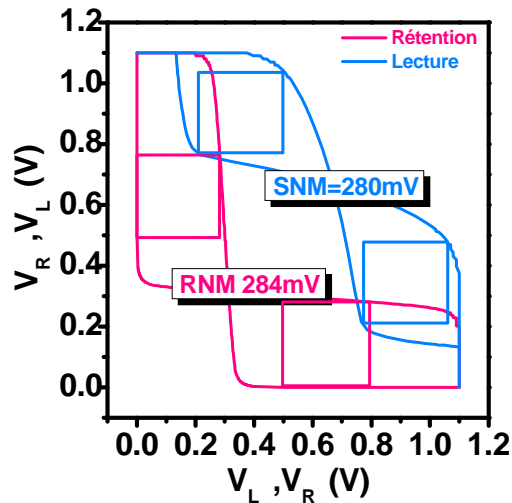


Figure 4.26 Courbes papillons obtenues par simulation ATLAS en trois dimensions en rétention et en lecture.

Le rapport de 2 entre W_n et W_p apporté par le dimensionnement permet d'augmenter le rapport entre le courant I_{ON} de drive (nMOS) et celui du transistor d'accès, ce qui renforce la stabilité en lecture selon l'inéquation 4.2. La Figure 4.26 confirme en effet que la stabilité en lecture a été renforcée, car la valeur de la marge de stabilité en lecture passe de 157mV à 280mV. Les valeurs des marges de stabilité ont donc été équilibrées grâce au dimensionnement. Ces marges dépassent les spécifications usuelles en termes de stabilité (20 % de V_{DD}) pour les cellules mémoires SRAMs.

Ces résultats montrent que l'exploitation de la modification de la tension de seuil dynamique permet de concevoir une SRAM 4T fonctionnelle avec des marges de stabilité importante en rétention et en lecture et ce, sans besoin de technique de rafraichissement.

Gain en densité

Les layout de la SRAM4T en 2D et en 3D ont été comparés. Ces layout ont été réalisés avec des règles de dessin 65 nm pour la partie back-end et 45 nm pour la partie front-end (Figure 4.27 Figure 4.27). Un gain en densité de 16.4% a été obtenu en empilant la cellule sur 2 niveaux.

Notons que pour obtenir de telles marges de stabilité avec les technologies DG et UTB FDSOI, un contact face arrière est nécessaire. En conséquence, à stabilité comparable, un gain supplémentaire est attendu grâce à la suppression du contact face arrière par rapport aux technologies planaires.

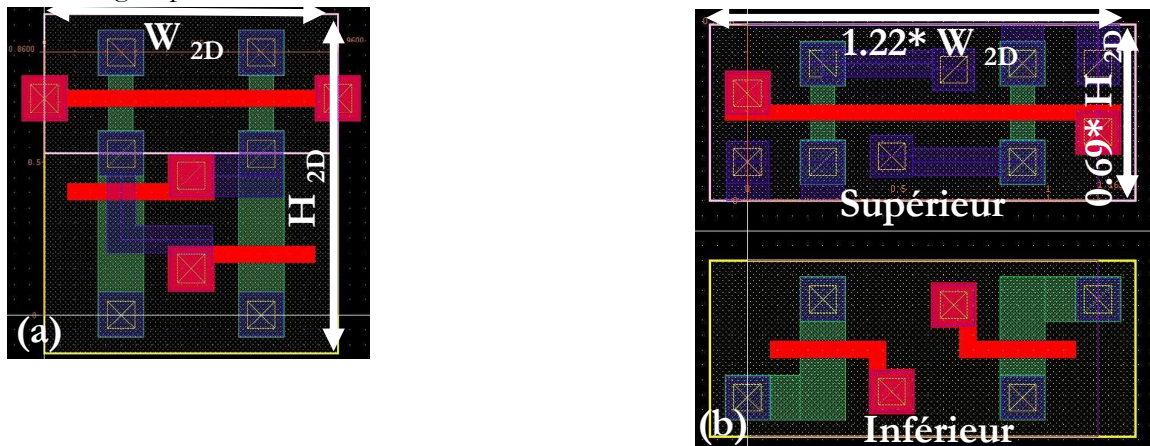


Figure 4.27 (a) Layout planaire et (b) layout avec la technologie 3D séquentielle UTILD de la SRAM4T. Le positionnement des transistors dans (b) permet d'améliorer la stabilité de la cellule via la modification dynamique de tension de seuil. (En vert, les zones actives, en rouge les zones de grille, en jaune, les contacts, en bleu les lignes de métal)

B. 4 -La SRAM 6T 3D UTILD

B.4.1 - Présentation du modèle compact

Dans le cas de la cellule SRAM 6T, l'étude de stabilité à l'aide de la simulation se complexifie, rendant le temps de codage de la structure et les durées de simulations trop importants. Pour palier à cette difficulté, il est nécessaire de développer des modèles compacts. De plus, grâce à la courte durée des simulations, son utilisation permettra de tenir compte d'autres figures de mérite importantes lors de la conception de cellules SRAM, à savoir la marge en écriture WM (Write Margin) pour une écriture efficace, le courant de conduction (I_{CELL}) maximal pour atteindre de bonnes performances, un courant en rétention (I_{OFF}) minimal pour minimiser la puissance consommée statique et le nombre maximal de bits par ligne (NBL). La conception de la réalisation particulière de la SRAM 6T a été réalisée par des concepteurs. Des compléments d'information sur la conception de cette cellule peuvent être trouvés dans [Thomas 09].

Dans ce paragraphe, nous nous limiterons donc à décrire brièvement les facteurs permettant l'amélioration des performances de la cellule (c'est-à-dire le positionnement des différents transistors) et présenterons les caractéristiques (stabilité, performance, densité) résultantes.

Nous avons utilisé la simulation TCAD pour mettre au point le modèle compact nécessaire à la conception optimisée de la cellule SRAM. Ces simulations ont permis de fournir les données d'entrée décrivant les modifications des caractéristiques électriques avec le couplage pour calibrer le modèle.

Les comportements électriques des transistors supérieurs et inférieurs sont décrits par deux modèles différents. La Figure 4.28 permet de mettre en évidence les paramètres à prendre en compte lors de l'établissement des modèles compacts.

Dans le cas du transistor inférieur, le modèle compact d'un transistor FDSOI peut être directement utilisé puisque les paramètres électriques de celui-ci ne sont pas modifiés par la présence du transistor supérieur.

En ce qui concerne le transistor supérieur, son comportement s'apparente plus à celui d'un MOS Double Grille avec un oxyde de grille arrière plus épais. En effet les paramètres du FET supérieur vont être modifiés uniquement en fonction de la polarisation de la grille du transistor inférieur et non des polarisations des accès. Cependant, son comportement électrique est foncièrement différent de celui d'un transistor double grille classique puisqu'au vu des épaisseurs de diélectrique en jeu, l'interface arrière ne se situe jamais dans le régime d'inversion forte.

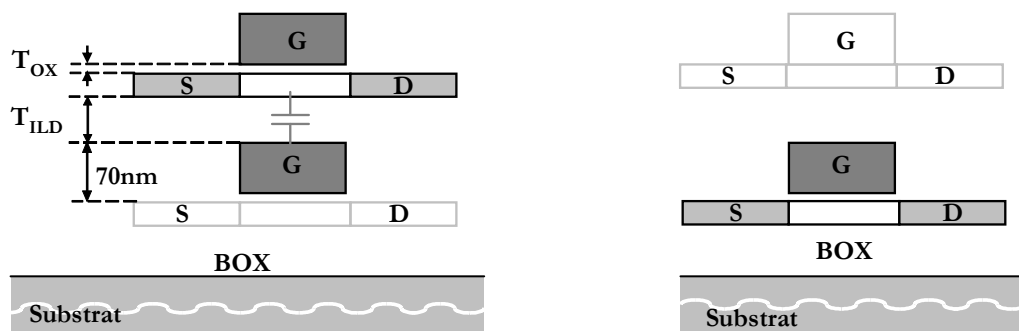


Figure 4.28 Représentations schématiques des environnements influençant le comportement statique du transistor (a) supérieur et (b) inférieur.

B.4.2 - Présentation de la SRAM 6T et analyse de sa stabilité

La Figure 4.28 représente la SRAM 6T ainsi que l'identification des transistors critiques pour la stabilité en régime de lecture et d'écriture.

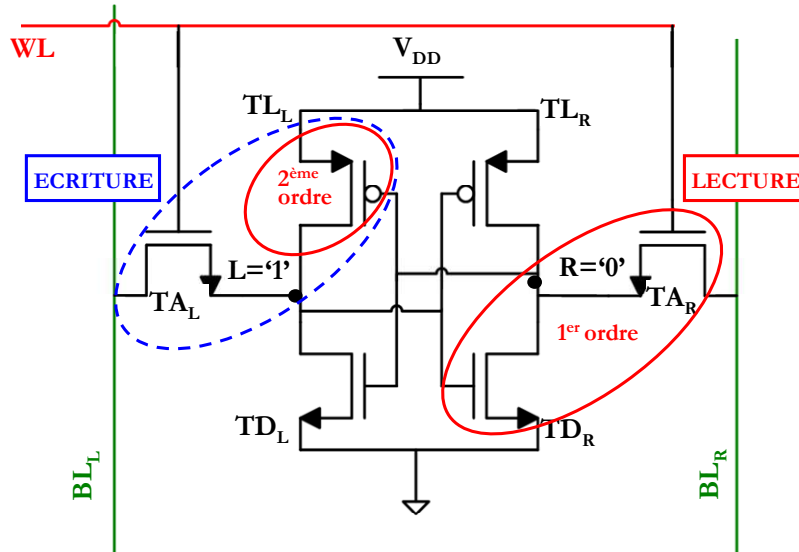


Figure 4.29 Vue schématique d'une SRAM 6T mettant en valeur les transistors dont le dimensionnement est critique en termes de stabilité dans les modes d'écriture et de lecture. (Pour la phase de lecture les BL sont préchargées à 1)

Contrairement à la SRAM 4T, la cellule 6T va être étudiée dans le mode de lecture uniquement car elle possède intrinsèquement de bonnes marges en rétention grâce à ces deux inverseurs rétro-couplés.

Comme précédemment, il s'agit d'identifier le nœud qui présente la plus haute probabilité de basculer en premier. En observant les courants entrant en jeu dans la configuration de polarisation de lecture, on observe que le nœud 'R'=0 est le plus critique. L'équation de stabilité pour ce nœud est donc la suivante:

$$I_{G-TLL} + I_{OFF-TLR} + I_{ON-AR} \leq I_{ON-DR} \quad (\text{EQ 4.3})$$

Selon cette équation et en négligeant les courants I_{G-TLL} et $I_{OFF-TLR}$, on observe qu'au premier ordre, la stabilité en lecture est améliorée par une augmentation du courant traversant le transistor de conduction et une diminution du courant traversant le transistor d'accès. De plus, une grande tension doit aussi être maintenue au nœud L, la stabilité en lecture dépend donc aussi, mais au second ordre, du courant traversant le transistor de charge.

D'autre part, pour pouvoir écrire la cellule, il faut que la tension du nœud 'L' puisse descendre jusqu'à la valeur de la tension de seuil de l'inverseur (TL_R, TD_R) connecté au nœud R. Donc pour permettre l'écriture de la cellule il faut que le courant traversant TL_L soit inférieur à celui traversant TA_L .

$$I_{ON-AL} \geq I_{ON-LL} \quad (\text{EQ 4.4})$$

Comme dans le cas de la SRAM 4T, nous observons que ces deux marges sont comme des vases communicant et qu'en améliorant un critère on dégrade nécessairement l'autre.

B.4.3 - Positionnement des transistors et analyse des résultats

Gain en stabilité

Le positionnement des transistors décrit dans la Figure 4.30 permet d'améliorer certains critères sans pour autant dégrader les autres.

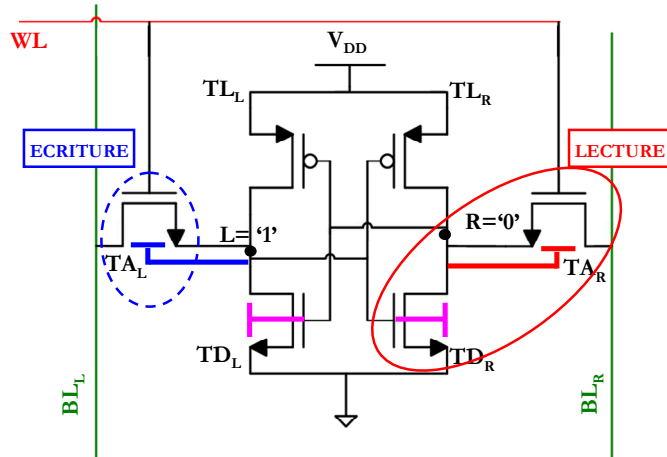


Figure 4.30 Positionnement des transistors dans la cellule SRAM6T à stabilité améliorée grâce à la technologie 3D monolithique avec diélectrique inter niveau ultrafin.

Régime d'écriture

Placer T_{AL} au dessus du nœud L va abaisser sa tension de seuil et ainsi renforcer son courant de conduction par rapport à celui de T_{LL} . La marge en écriture résultante est augmentée (cf EQ 4.4)

Régime de lecture :

Placer T_{AR} au dessus du nœud R va élever sa tension de seuil et ainsi affaiblir son courant de conduction par rapport à celui de T_{DR} . La marge en lecture résultante est augmentée (cf EQ 4.3)

Une réduction de la fuite des transistors de drive est obtenue en plaçant ceux-ci au dessus d'une électrode maintenue à la

même valeur que celle de leurs grilles.

Les caractéristiques résultantes de cette réalisation particulière sont récapitulées dans la Figure 4.31 L'utilisation du modèle compact a permis d'optimiser finement le dimensionnement de cette cellule.

Cellule	SRAM6T 2D	SRAM6T 3D UTILD	Comparaison
SNM (mV)	229	252	+ 10,4 %
RNM (mV)	430	435	~
WM (mV)	381	386	~
I_{CELL} (μA)	33	33	=
I_{OFF} (pA)	14.5	12.7	- 12.3 %
I_{OFF-PG} (pA)	0.74	0.65	- 11.7 %
NBL (10^6)	45.2	50.6	+ 12 %

Figure 4.31 Extraction des caractéristiques de la cellule

Gain en densité :

En termes de gain en densité, la cellule en 3D occupe 20% de surface en moins par rapport à la 6T en planaire avec le même I_{CELL} pour occuper au final une surface de $0.290\mu m^2$ (cf. Figure 4.32). Le DRM présenté dans le chapitre III, n'est pas adapté au cas des SRAMs où de nombreuses règles sont violées. Les dérogations sur les règles du DRM3D monolithique ont été inspirées de [Boeuf05]. Les règles essentielles sont récapitulées dans la Figure 4.32

	Taille/Espacement (nm)
Active	55 / 80
Poly	45 / 80
Contact interne	50 / 90
Contact selectif bas	60 / 90
Contact selectif haut	50 / 90
M1	70 / 60

Figure 4.32 Récapitulation des règles critiques utilisées pour le design de la SRAM 6T 3D UTILD. Les dérogations de règles par rapport au DRM 3D monolithique 45 nm sont inspirées de [Boeuf05]

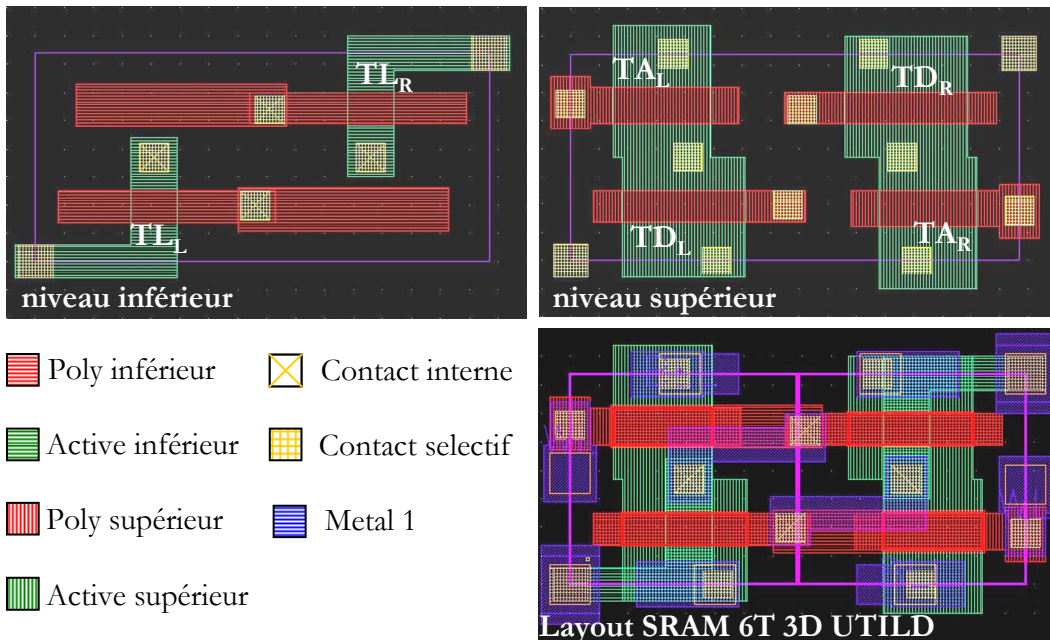


Figure 4.33 : Layout de la cellule SRAM 6T 3D UTILD vues par niveau et layout complet jusqu'au métal 1.

Ce qu'il faut retenir
IV.B : Application aux SRAMs

Contexte

- Les SRAMs 6T sont encombrantes et occupent actuellement une surface considérable de la surface des circuits intégrés (classiquement la surface occupée par la SRAM peut atteindre les 60% de la surface de la puce pour un microprocesseur et les 80% de la surface de la puce pour un SoC).
- La réduction des dimensions engendre une réduction des marges de stabilité via l'augmentation de la variabilité et la diminution des tensions d'alimentation.
- La modification dynamique de la tension de seuil des transistors est une solution pour augmenter ces marges de stabilité.
- Grâce à la technologie 3D séquentielle, il est possible de gagner en densité en ne réduisant pas les dimensions des transistors. En effet, [Jung 05] démontre une cellule de $0.16\mu\text{m}^2$ avec une longueur de grille minimale de 65 nm. En conséquence, les SRAMs apparaissent comme une application phare pour l'intégration 3D séquentielle.
- Garder des dimensions de transistors relâchées permet de réduire les coûts de développements via l'utilisation d'équipements des générations précédentes et de limiter la variabilité (effets canaux courts plus faibles).
- De plus en utilisant le couplage à travers un diélectrique inter-niveau ultra fin, l'architecture 3D séquentielle offre la possibilité de modifier dynamiquement la tension de seuil de façon similaire aux technologies DGMOS et UT2B-SOI. Cette particularité peut être mise à profit pour augmenter les marges de stabilité des SRAMs.
- La réalisation de SRAM avec modification de tension de seuil ne peut pas être obtenue avec une architecture 3D parallèle à cause des performances d'alignement insuffisantes pour placer les transistors supérieurs et inférieurs précisément l'un au dessus de l'autre.

La SRAM 4T :

- La SRAM 4T présente une instabilité en rétention, sans dispositif de rafraîchissement, les marges de stabilité requises (20% VDD) ne peuvent être atteintes à la fois en rétention et en lecture. Les paramètres de la cellule sont alors généralement ajustés pour obtenir une marge en lecture suffisante et l'instabilité en rétention est corrigée à l'aide de systèmes de rafraîchissement.
- L'utilisation du couplage dynamique grâce à l'architecture 3D séquentielle permet de rendre stable en rétention la SRAM4T sans dispositif de rafraîchissement. La réalisation particulière présentée permet d'atteindre des marges de stabilité en rétention et de lecture

de plus de 25% V_{DD} .

- Le gain en densité par rapport au layout planaire est de 16% environ.
- Les autres technologies (DGMOS, UTB FDSOI, DTMOS) permettant de modifier dynamiquement la tension de seuil utilisent toutes un contact supplémentaire (contact face arrière ou prise substrat). En conséquence, à stabilité comparable, le gain en densité obtenue avec l'architecture 3D monolithique sera accru grâce à la suppression de ce contact.
- La cellule 4T ainsi obtenue est extrêmement intéressante en termes de densité par rapport à la cellule 6T car la cellule SRAM 4T est déjà 30 à 40% plus petite que la cellule SRAM 6T.

La SRAM 6T :

- La réalisation d'une cellule plus fiable et à plus basse consommation est obtenue grâce à la modification de tension de seuil obtenue par couplage entre l'étage supérieur et inférieur : La réalisation particulière présentée permet d'améliorer la marge en lecture (+10%), le nombre maximal de bits par ligne (+12%), et de réduire la consommation statique (-12%), sans dégrader les autres paramètres importants (WM, RNM, I_{CELL})
- La cellule 6T présentée occupe une surface réduite de 20% par rapport à une architecture planaire classique.

C - Les mémoires FLASH

La demande pour des mémoires FLASH NAND à bas coût et à haute densité a augmenté considérablement avec la montée en puissance des applications vidéo et audio portables.

La réduction des dimensions entraîne une dégradation des caractéristiques électriques et de la fiabilité des cellules. Pour le nœud 30 nm, il semble que de nombreuses limites soient atteintes, forçant à innover par rapport à l'architecture quasi inchangée depuis des générations : avec une grille de contrôle et une grille flottante en poly et un interpoly en ONO, cf Figure 4.34 .

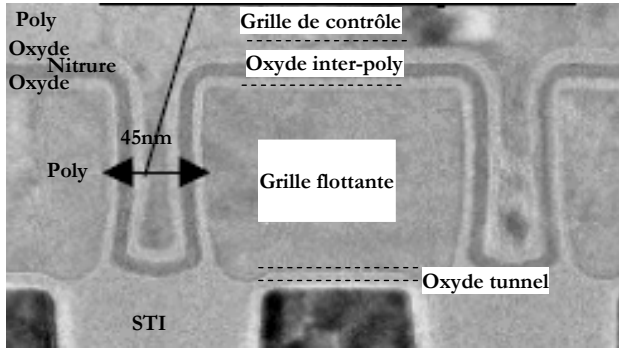


Figure 4.34 Architecture classique pour une cellule FLASH nand. Grille de contrôle et flottante en poly et oxyde interpoly de type ONO (Oxyde-Nitrure-Oxyde)

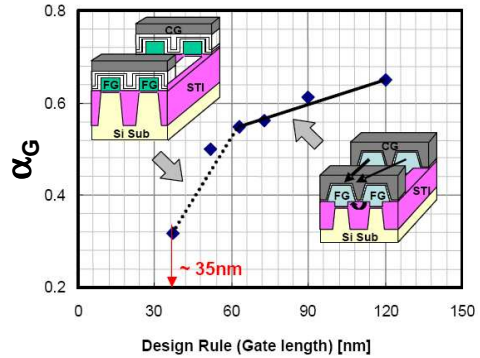


Figure 4.35 Evolution du coefficient de couplage entre la grille de contrôle et la grille flottante en fonction du nœud technologique.

En effet, l'espacement entre les cellules se réduisant, le recouvrement de la grille de contrôle sur le flanc de la grille flottante devient impossible et rend le couplage entre la grille de contrôle et la grille flottante insuffisant pour répondre aux critères prévu par l'ITRS (cf Figure 4.35).

D'autre part, le rapprochement des cellules entraîne l'apparition d'un couplage parasite entre cellules adjacentes. Les électrons stockés dans la grille flottante d'une cellule peuvent décaler la tension de seuil de la cellule voisine.

De plus, avec la réduction des dimensions, on observe une augmentation des effets canaux courts (réduction de la longueur de grille) et une diminution des courants due à la réduction des largeurs des transistors.

Enfin due à la constante diminution de la capacité du diélectrique interpoly, le nombre d'électrons nécessaires sur la grille flottante pour coder un bit est réduit à chaque génération [Molas04]. Quand le nombre d'électrons atteint les limites statistiques, la distribution de tension de seuil dans les cellules du plan de mémoire va devenir incontrôlée et les niveaux logiques imprévisibles [ITRS PIDS 07].

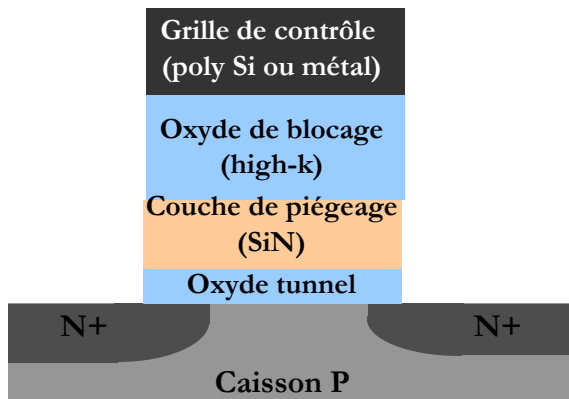


Figure 4.36 Architecture innovante avec remplacement de la grille flottante par une couche à piégeage de charge ainsi que réalisation de l'oxyde de blocage en high-k.

Une intégration 3D séquentielle de cellule FLASH a été démontrée avec succès par Samsung avec deux niveaux de cellules FLASH empilés (cf Figure 4.37)

Les cellules empilées sont de type TANOS (TaN/Al₂O₃ /Nitrure/Oxyde/Si) et la zone active supérieure semble être en poly recristallisé.

Pour répondre à ces problématiques, des architectures innovantes sont proposées. Une option prometteuse semble être une cellule intégrant un diélectrique permettant de piéger des charges à la place de la grille flottante en poly associée avec un high-k pour un meilleur couplage de la grille de contrôle avec le diélectrique à piégeage de charge (cf Figure 4.36).

Une autre option pour continuer à gagner en densité sans avoir à changer d'architecture est d'empiler les cellules en 3D.

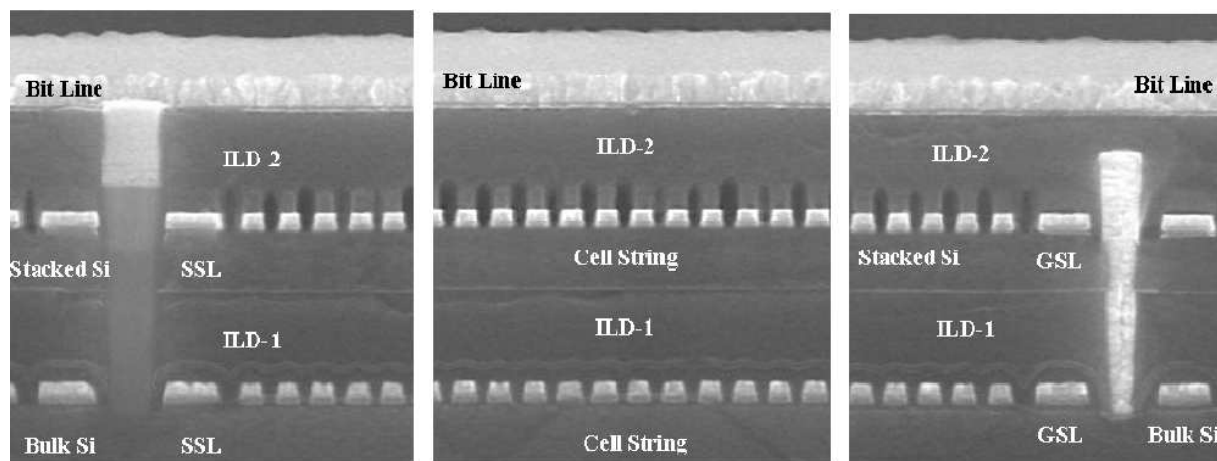


Figure 4.37 Vue en coupe des cellules FLASHs empilées démontrées dans [Jung06]

Samsung évoque l'intérêt de l'intégration séquentielle par rapport à l'intégration parallèle en termes de coût puisque les étapes de Back-end peuvent être réalisées en commun dans le cas séquentiel. Cependant l'éventuel gain en coût par rapport à une intégration 3D parallèle apparaît limité car le back end est relativement simple pour une intégration FLASH (seulement 2 niveaux de métaux).

D'autre part, au vue de la faible densité de connexions nécessaires dans une structure de type NAND, l'architecture séquentielle ne s'impose pas comme unique solution 3D. Seul un faible gain en coût pourrait faire préférer l'intégration séquentielle à l'intégration parallèle. De plus, la maturité de l'intégration parallèle par rapport à l'intégration séquentielle joue aussi en défaveur de l'intégration séquentielle, pour laquelle il faudrait développer une technologie faible budget thermique pour le niveau supérieur. A court terme, l'option parallèle semble donc être la plus intéressante pour cette application.

Le domaine des FLASH NOR, par contre impose l'utilisation de l'intégration séquentielle de par la grande densité d'interconnexions entre les deux niveaux requise. Actuellement, le marché de la NOR ne cesse de diminuer, complètement supplanté par l'architecture NAND. La possibilité de l'intégrer en 3D la FLASH NOR peut éventuellement la rendre plus compétitive par rapport à la NAND.

Afin d'évaluer les performances de cellules FLASH 3D, des matrices de cellules FLASH NAND et NOR ont été intégrées dans le nouveau masque.

D - Application aux imageurs fortement miniturisés

L'introduction de fonctionnalités d'imagerie dans les dispositifs portables a poussé la conception d'imageurs CMOS associant des résolutions élevées à un encombrement, une consommation et un coût réduit. Cet objectif a été atteint en diminuant la taille des pixels jusqu'au noeud 1.4 μ m. Cependant cette miniaturisation n'a pas eu que de des avantages, elle s'est accompagnée d'une dégradation des performances électro-optiques. Même si certaines optimisations permettent de diminuer ce constat, les architectures conventionnelles vont se montrer insuffisantes dans la perspective de pixels submicroniques : seule une conception radicalement différente des pixels pourra répondre aux problématiques liées à la réduction des dimensions.

Les pixels faible bruit modernes sont basés sur une architecture à photodiode clampée associée à quatre transistors. La présence des transistors à l'intérieur du pixel limite le facteur de remplissage et la capacité de stockage de la photodiode, dégradant le rapport signal sur bruit et la dynamique. Pour répondre à ces problématiques, une architecture originale a été proposée [Coudrain08][Coudrain09]. Elle associe une illumination face arrière des photodiodes à la réalisation des transistors sur un niveau dédié comme présenté dans la Figure 4.38 . Dans le cas d'un pixel 3D, la photodiode ne souffre plus de la présence des transistors (excepté la grille de transfert); la surface photosensible peut ainsi être augmentée de 35% pour un pixel 1.4 μm , et la diode éloignée des sources de courant d'obscurité.

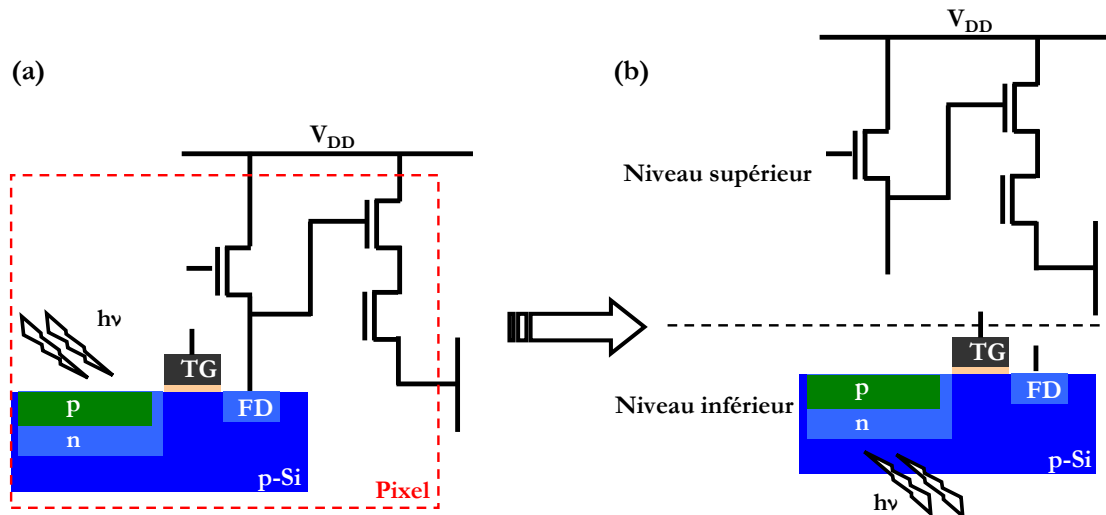


Figure 4.38 : (a) Schéma électrique d'un pixel conventionnel planaire à 4 transistors et (b) Pixel 3D avec illumination face arrière.

Bien que déjà démontrée sur des capteurs complexes [Suntharalingam05] [Koyanagi01] avec des largeurs de pixels de plus de 8 μm , la possibilité d'intégrer un pixel CMOS en trois dimensions, n'a jamais été appliquée au contexte particulier des imageurs fortement miniaturisés. Dans la technologie 1.4 μm standard, la largeur des contacts est de 90 nm. En passant en 3D la taille de ces vias doit rester constante afin de laisser le maximum de surface pour les photodiodes (cf Figure 4.39). L'approche 3D parallèle ne peut pas être appliquée ici car les contacts ayant des dimensions de l'ordre de l'ordre de 90 nm pour un pixel 1.4 μm , ne sont pas compatibles avec un alignement au micron. Dès lors, l'approche séquentielle s'impose.

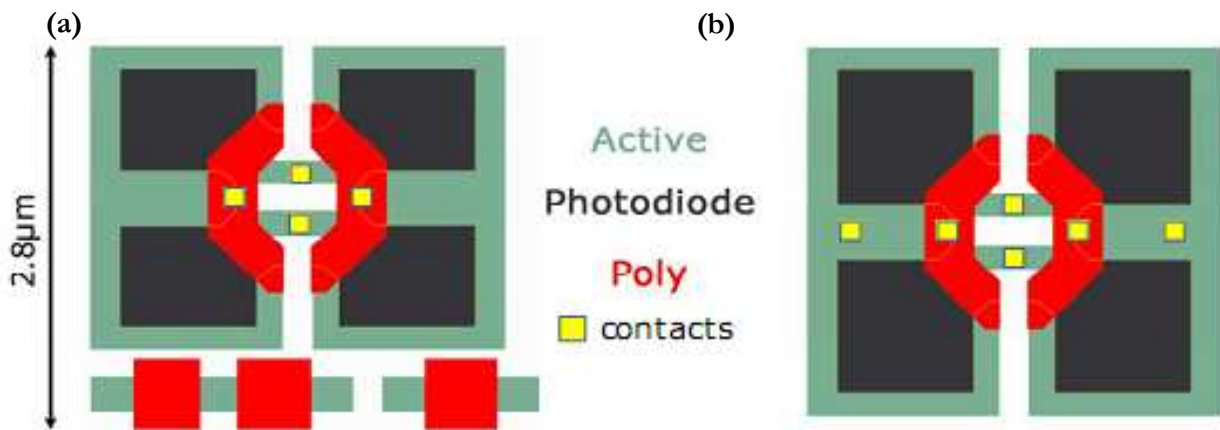


Figure 4.39 : (a) Layout d'un pixel à 4 transistors conventionnel planaire et (b) Layout d'un pixel 3D avec illumination face arrière (architecture proposée par Coudrain et al [Coudrain08])

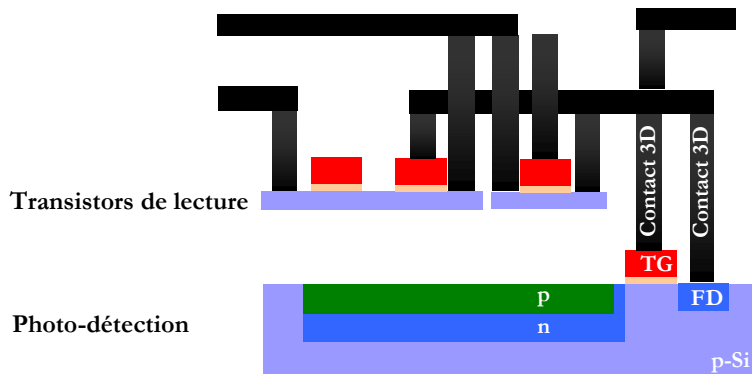


Figure 4.40 Vue schématique en coupe du pixel 3D avec illumination face arrière.

Pour réaliser ce type d'empilement il est nécessaire de limiter le budget thermique du transistor supérieur pour ne pas dégrader les performances de la photodiode et du transistor de transfert via la diffusion des dopants et le démouillage en phase solide du siliciure. Le transistor inférieur est siliciuré avec du Cobalt. Les transistors du niveau supérieur seront réalisés avec le procédé décrit dans cette thèse.

Pour satisfaire à une application imageur, les transistors à bas budget thermique doivent présenter de faibles niveaux de bruit basse fréquence au risque de dégrader les performances de l'imageur dans des conditions de faible éclaircissement [Leyris06].

La Figure 4.41 présente la densité de bruit en $1/f$ en fonction du courant de drain mesuré sur un transistor nMOS faible budget thermique réalisé au Leti et une référence haute température avec le même empilement de grille : $(\text{HfO}_2/\text{TiN}/\text{Poly})$. Nous observons que le niveau de bruit n'est pas dégradé par rapport au procédé haute température. Nous observons de plus une bonne corrélation de la mesure avec le modèle $\Delta N/\Delta \mu$ pour les deux technologies. Le procédé basse température conduit donc à un comportement standard en termes de bruit. Cette figure de mérite montre que l'intégration 3D séquentielle est compatible avec une technologie imageur.

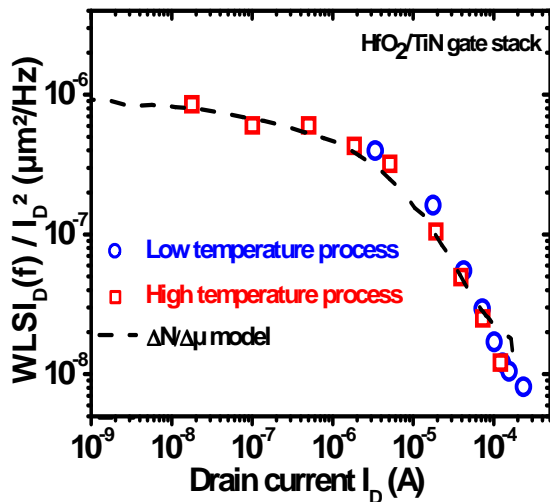


Figure 4.41 : Densité de bruit en 1/f en fonction du courant de drain à 1Hz pour des transistors nMOS faible BT et comparaison à une référence haute température.

Par ailleurs, l'intégration 3D séquentielle peut permettre la réalisation de pixels de petite taille avec un nombre plus élevé de transistors ainsi que l'ajout de mémoires sur le niveau supérieur. Ceci pourrait ouvrir la voie à des capteurs à la fois miniaturisés et capables de fonctionnalités étendues : imagerie instantanée faible bruit, niveau de charge de la diode, capteurs à grande vitesse ou à grande dynamique...) limités pour l'instant à des pixels larges (10μm minimum). De plus le marché des imageurs ne cesse de croître et la part de marché pour les imageurs fortement miniaturisés destinés aux applications portable est d'environ 75%. En conséquence la réalisation de pixels fortement miniaturisés en trois dimensions apparaît comme une application prometteuse de l'intégration 3D séquentielle.

Des pixels avec cette technologie ont été dessinés dans le futur masque 3D réalisé au Leti² et permettront de réaliser un démonstrateur complet avec le niveau photodiode à l'étage inférieur et un étage avec les transistors de lecture.

² Les layouts ont été dessinés par P. Coudrain de ST microelectronics

CONCLUSION GENERALE

Conclusions

L'intégration 3D fait actuellement figure d'alternative potentielle à la simple réduction des dimensions pour continuer à gagner en densité d'une génération à la suivante. Elle est perçue comme un moyen de limiter les effets canaux courts en permettant de relaxer les contraintes dimensionnelles pesant sur le transistor et comme un moyen de réduire le délai dans les interconnexions. Elle permet aussi de réaliser des co-intégrations de technologies hétérogènes.

Cette thèse porte sur l'intégration à trois dimensions séquentielle, le terme séquentiel traduisant le fait que les transistors des différents niveaux sont fabriqués les uns après les autres sur un même substrat. La difficulté majeure de ce type d'intégration est la réalisation des niveaux de transistors supérieurs (comprenant les étapes de création de la zone active ainsi que celles du procédé de fabrication du transistor) à bas budget thermique afin de préserver le niveau de transistor inférieur de toute dégradation. Les transistors à bas budget thermique doivent cependant présenter des performances similaires à celles des transistors réalisés avec un budget thermique classique.

Ce type d'intégration se différencie de l'intégration dite de type « parallèle », où les deux plaques sont fabriquées en parallèle puis sont associées pour réaliser l'empilement tridimensionnel. Dans ce cas, les budgets thermiques nécessaires à la réalisation des différents niveaux de l'empilement ne sont pas limités. L'inconvénient majeur de cette intégration se situe dans les faibles densités d'interconnexions entre les différentes strates atteignable. Les performances d'alignement entre les deux substrats étant actuellement de l'ordre du micromètre, la densité de vias 3D est limitée à 10^8 vias/cm², alors que pour le nœud technologique 45 nm, la densité de contact à interconnecter est de 10^{10} contacts/cm². Ainsi le transfert d'information entre le niveau inférieur et le niveau supérieur est limité à un contact 3D tous les 100 contacts.

Dans le cas de l'intégration 3D séquentielle, les performances d'alignement des niveaux supérieurs sur les niveaux inférieurs ($\sigma \sim 10$ nm) permettent de réaliser des interconnexions 3D à l'échelle du transistor. Il est ainsi possible de venir réaliser les cellules CMOS à cheval sur les deux niveaux (par exemple un inverseur avec le nMOS empilé sur le pMOS).

Malgré cet avantage considérable par rapport à l'option 3D parallèle, l'intégration 3D séquentielle est moins étudiée actuellement. C'est en partie la complexité technologique de l'intégration séquentielle qui la place comme une intégration à plus long terme (~10ans). C'est dans ce contexte que nous avons entrepris la démonstration et l'étude de cette technologie. Les résultats obtenus au cours cette thèse sont présentés dans ce manuscrit articulé en quatre chapitres.

Dans le premier chapitre, nous avons identifié les défis technologiques de l'intégration 3D séquentielle, proposé des solutions techniques permettant de les relever, justifié ensuite de leur pertinence pour des nœuds technologiques avancés (sub-45nm) puis comparé nos résultats à ceux présentés dans la littérature. Les développements ont principalement porté sur quatre briques technologiques :

- 1- Les travaux réalisés ont permis de développer un siliciure à base de nickel stable en température pour le FET inférieur en prévision des budgets thermiques nécessaires à la réalisation du FET supérieur. L'intégration d'un siliciure est indispensable pour adresser des nœuds avancés et permet de démarquer ces travaux de l'état de l'art de l'intégration séquentielle, où l'intégration d'un siliciure avec le procédé SALICIDE n'avait pas été démontrée.

2- La réalisation de la zone active supérieure de qualité cristalline équivalente à un substrat massif est la problématique essentielle de l'intégration 3D séquentielle. Nous proposons d'utiliser le collage moléculaire afin de reporter la zone active supérieure au dessus du niveau de transistor inférieur préalablement fabriqué.

Les reports de films sont complets (sur plaque de 200mm de diamètre) et les films transférés sont fins (jusqu'à 10 et 50 nm dans le cas Si et Ge respectivement). Le collage est réalisé à bas budget thermique (200°C). Les caractérisations acoustiques et infrarouges ne révèlent aucun défaut de collage majeur. La qualité du collage est confirmée pendant les étapes de retrait substrat et de gravure des contacts qui n'entraînent pas de décollement (même local) des deux plaques.

Cette méthode se démarque des autres techniques de réalisation de l'active supérieure présentées dans la littérature par : a) la qualité cristalline de la zone active qu'elle permet d'obtenir, b) le contrôle de l'épaisseur du film atteint, c) la réalisation d'une zone active « pleine plaque » (Dans le cas des techniques de recristallisation permettant d'obtenir une zone active monocristalline, on observe une perte de densité due aux fenêtres d'ouverture, nécessaire au transfert de l'information cristalline de la zone active inférieure à la zone active supérieure, et des défauts cristallins aux fronts de cristallisation, qui rendent ces zones impropres pour la réalisation de transistors), d) le choix possible de l'orientation cristalline des deux zones actives empilées. Elle est pour toutes ces raisons, la technique la plus à même de répondre aux exigences requises pour une application industrielle. Elle est d'ailleurs utilisée quotidiennement par SOITEC pour réaliser des substrats SOI.

3- Pour la réalisation du FET supérieur nous avons proposé de limiter le budget thermique global à 650°C. Deux options ont été envisagées, une option germanium bénéficiant du faible budget thermique nécessaire pour l'activation des dopants, et une option silicium où l'activation thermique des dopants à haute température est remplacée par une activation par recristallisation en phase solide à 600°C (SPER, Solid Phase Epitaxy Regrowth).

Nous avons montré que la SPER permet d'obtenir des compromis de résistance de couches en fonction de la profondeur de jonction équivalentes à celles atteintes avec les méthodes d'activations les plus avancées (FLASH, laser). Pour réaliser les surélévations des sources et drains à 650°C, les épitaxies Si ont été remplacées par des épitaxies SiGe avec une proportion de Ge de 30%. La surélévation des sources et drain est une étape indispensable pour réaliser des transistors sur film minces (<10nm).

4- Enfin, les épaisseurs de films de zones actives et de diélectrique inter niveau ont été optimisées afin de permettre de connecter les niveaux supérieurs et inférieurs avec un seul niveau lithographique via une gravure sélective avec arrêt sur les accès. Le développement d'un diélectrique inter-niveau fin (100nm), nous permet de conserver des techniques de gravure des contacts semblables à celles utilisées en planaire.

L'intérêt de cette intégration, à savoir la grande précision d'alignement permettant la réalisation d'interconnexions denses entre les deux niveaux empilés, est bien vérifié. L'écart type du désalignement entre active supérieur et active inférieur est de 7 nm, cette valeur est comparable par exemple à celle obtenue lors de l'alignement du niveau grille sur le niveau active (stepper lithographie DUV 248nm).

Dans le chapitre 2, nous avons présenté et caractérisé l'intégration complète réalisée. Nous avons analysé les caractéristiques des transistors inférieurs et supérieurs, celles des inverseurs et cellules SRAM fabriquées à cheval sur les deux niveaux. Nous avons intégré les briques technologiques présentées dans le chapitre précédent.

La démonstration du transistor inférieur est remarquable par rapport à l'état de l'art en ce que ces accès sont siliciurés. Après réalisation du FET supérieur, nous avons montré que le siliciure à base de nickel n'a pas subi d'agglomération en phase solide.

Les transistors supérieurs sont réalisés sur des zones actives reportées par collage moléculaire suivi du retrait du substrat initial. Deux types de transistors ont été fabriqués, des transistors sur SOI (n et pMOSFETs) afin d'adresser les applications de type basse consommation (LSTP, Low Standby Power) et GeOI (pMOSFETs) pour adresser des applications du type haute performance (HP, High Performance).

Les cellules CMOS sont alors réalisées verticalement avec par exemple un inverseur avec le pMOSFET sur substrat germanium empilé au dessus d'un nMOSFET sur substrat silicium.

Dans le cas du silicium, l'activation des dopants par SPER à 600°C a été réalisée avec succès et des transistors fonctionnels jusqu'à des longueurs de grille de 50nm ont été démontrés. Même sans travaux d'ingénierie de jonction, les transistors présentent un bon contrôle des effets canaux courts jusqu'à des longueurs de grille de 50nm, attribué à la réduction de la diffusion des dopants permise par les faibles budgets thermiques utilisés. Un avantage supplémentaire de la réduction du budget thermique est la réduction de la recroissance de l'oxyde de silicium formé au pied du high-k. Nous avons constaté que les fuites de jonctions sont légèrement plus importantes (~1 décade) dans le cas SPER que dans la référence haute température. Cette dégradation peut s'expliquer par la présence de défauts résiduels (End Of Range : EOR) situés au delà de l'interface Si amorphe/Si cristallin et dont la dissolution est incomplète avec un budget thermique de (600°C) et au caractère plus abrupt des jonctions conduisant à une augmentation du courant tunnel bande à bande.

Dans le cas du transistor supérieur sur active germanium, les caractéristiques des transistors correspondent aux caractéristiques classiquement obtenues dans la filière GeOI du Leti.

Les structures CMOS 3D (inverseurs et SRAMs) sont fonctionnelles. Les cellules 3D co-intégrées pMOS-GeOI sur nMOS-SOI et mono-matériau SOI sur SOI ont été démontrées. Il s'agit de la première démonstration de cellules 3D interconnectées avec une zone active supérieure réalisée par collage moléculaire. Les SRAMs sont stables en lecture et en rétention.

Dans le chapitre 3, nous nous sommes intéressés aux perspectives de gains de l'intégration 3D séquentielle. Les réflexions ont été menées dans le cadre d'applications logiques haute densité avec la technologie FDSOI (nœud 45nm).

Pour obtenir une perspective du gain en densité fiable adapté à des nœuds technologiques avancés et à la technologie étudiée, nous avons suivi la démarche suivante :

- Création d'un manuel de règles de dessin basé sur des hypothèses connues et prenant en compte des paramètres technologiques réalistes.
- Réalisation de bibliothèques de cellules standard (2 options comparées nMOS empilés sur nMOS et CMOS empilés sur CMOS)
- Réalisation de la synthèse logique sur une cellule Multiply ACcumulator (~1000 transistors).

Les gains pour les deux types d'intégration ont été calculés en comparant la taille de la cellule en 3D avec un layout panaire. Nous avons estimé un gain de 40% dans le cas N/P et de 50% dans le cas CMOS/CMOS. Cette étude a montré que l'option CMOS/CMOS, plus flexible dans le positionnement des transistors, conduit à un gain plus important que l'option N/P. Cependant cet avantage peut être nuancé par l'apport en performance, apporté par l'optimisation

indépendante des nMOS et de pMOS dans l'architecture N/P non prise en compte dans les simulations.

La perspective de gain en densité d'un nœud technologique sans réduction des dimensions des transistors est particulièrement attrayante. Garder des dimensions de transistors relâchées permet de réduire les coûts de développements et de fabrication via l'utilisation d'équipements des générations précédentes et de limiter la variabilité.

Concernant les perspectives de gain des transistors individuels, l'architecture 3D séquentielle se prête facilement à l'optimisation indépendante des performances de n et pMOS grâce à la possibilité de réaliser des interconnexions 3D à l'échelle du transistor, permettant de dessiner la cellule CMOS à cheval sur deux niveaux.

L'optimisation indépendante des deux types de transistors peut passer par le choix des empilements de grilles, des contraintes, des orientations et des matériaux de canal. L'utilisation du collage moléculaire, comparée aux techniques dites de « seed window », donne un degré de liberté supplémentaire dans l'optimisation des performances des transistors. L'apport de contrainte peut être réalisé via le substrat de départ et non seulement par le procédé de fabrication. De plus l'orientation des substrats peut être choisie indépendamment sur les deux niveaux empilés. Dans le cas de l'intégration seed window, le transfert de l'information cristalline via le germe implique la même orientation et contrainte pour le niveau inférieur et le niveau supérieur.

Quant au gain en délai possible dans les interconnexions, une distinction entre la technologie 3D séquentielle sans ligne de métal entre les niveaux de transistors empilés et la technologie parallèle est nécessaire.

Dans l'intégration parallèle, les interconnexions « verticalisées » sont les interconnexions longues, dans lesquelles le délai est le plus grand. Une étape de placement routage en trois dimensions permet de trouver un optimum de positionnement pour réduire la longueur des interconnexions longues. Les blocs de cellules placés en 3D sont dessinés en planaire.

Dans le cas séquentiel, l'absence de niveaux métalliques entre les niveaux de transistors empêche la réalisation d'un placement routage en 3D. La troisième dimension apparaît alors seulement au moment du dessin des cellules standard. Les interconnexions « verticalisées » sont alors des connections courtes (niveau M1). En conséquence, même si un gain en délai est probable au sein des cellules standard, l'intégration 3D séquentielle sans ligne de métal entre les niveaux de transistors est moins pertinente que l'intégration parallèle pour répondre à la problématique du délai dans les interconnexions.

Enfin, une estimation du surcoût par plaque du procédé 3D séquentiel par rapport au procédé planaire a été réalisée. Elle a été menée dans le cadre d'un produit purement logique avec une tension de seuil unique pour les deux types de transistors. Différents schémas d'intégration en 3 dimensions ont été étudiés : Intégration de type N/P ou CMOS/ CMOS orientée haute performance ou basse consommation (option standard). Pour les deux options, l'intégration de deux métaux de grille, une étape de siliciuration et la réalisation de huit niveaux de métal ont été prises en compte. Dans l'option haute performance, les deux types de transistors sont optimisés indépendamment en co-intégrant deux types d'orientation de surface des zones actives et deux types de couches de nitrure. L'influence du prix relatif du substrat SOI par rapport à un substrat Si massif sur le surcoût de la technologie a été étudié ainsi que le remplacement du substrat SOI par un substrat Si massif avec couche d'arrêt SiGe pour réaliser la zone active supérieure.

Dans le cadre des hypothèses faites, nous avons estimé que le surcoût de la technologie 3D séquentielle par rapport à l'intégration planaire varie entre 10% et 55% selon les options d'intégrations. Une intégration 3D séquentielle avec deux niveaux de transistors empilés ne double donc pas le coût du procédé. La part du prix du back end et du substrat dans le coût total

vient diminuer l'influence de l'augmentation du nombre d'étapes de front end dans l'intégration 3D séquentielle.

La séparation des deux types de transistor sur les niveaux empilés peut réduire de manière non négligeable le surcoût du passage en 3D.

Plus les nFETs et pFETs sont optimisés indépendamment, plus le surcoût en 3D N/P est faible. Par exemple, nous estimons à 21% et 11% les surcoûts par rapport aux intégrations planaires, pour l'option standard et haute performance respectivement, avec l'utilisation d'un substrat Si massif avec couches d'arrêt SiGe pour réaliser l'active supérieure.

Enfin, cette étude met en évidence l'intérêt financier du développement du report de la zone active supérieure à partir d'un substrat massif. L'estimation de la réduction de surcoût par rapport à la technologie planaire en utilisant un substrat Si massif avec couche d'arrêt en SiGe à la place d'un substrat SOI pour réaliser l'active supérieur est de 15% environ.

L'étude de coût réalisée permet de pondérer les résultats de l'étude de densité. Ainsi, nous pouvons estimer le gain en densité à coût constant. Nous observons que le résultat de l'étude de densité et de l'étude de coût ne valorise pas la même option. Dans le cas de l'étude de densité, c'est l'option CMOS/CMOS qui apporte le plus de gain (~50% contre 40% dans le cas N/P). L'étude de coût met en valeur le surcoût plus important de cette option. La pondération de l'étude de coût par l'étude de densité permet de valoriser l'option N/P par rapport à l'option CMOS/CMOS. De plus l'option N/P peut conduire à des gains supplémentaires en termes de performances des transistors par le biais de l'optimisation indépendante des deux types de transistors, en intégrant par exemple du germanium (voir du germanium contraint) pour les transistors de type p.

Cette estimation au premier ordre est encourageante car la possibilité de réaliser les cellules CMOS à cheval sur les deux niveaux est une spécificité de l'intégration séquentielle, elle conforte donc l'intérêt généré par l'intégration 3D séquentielle et engage à réaliser des études plus complètes pour conclure de manière fiable sur les perspectives de gains apporté par cette technologie.

Le chapitre 4 est consacré à l'étude des applications potentielles de l'intégration 3D autre que la logique haute densité.

Nous montrons dans une première partie que la proximité verticale des deux transistors empilés permet une modification significative de la tension de seuil du transistor supérieur lorsque la grille du transistor inférieur est polarisée de 0 à V_{DD} . Cette caractéristique peut être mise à profit dans des architectures SRAM pour améliorer les marges de stabilité (selon les cas, rétention, lecture, écriture).

Deux configurations particulières de SRAM, une version à quatre transistors et une version à 6 transistors menant à des gains conjoints en densité et stabilité par rapport à l'intégration planaire ont été proposées.

D'autre part, l'intégration 3D séquentielle a été identifiée comme une solution privilégiée pour répondre aux problématiques des imageurs CMOS pour des pixels fortement miniaturisés (génération < 1.4 μm). En effet, la réalisation des transistors de lecture sur un niveau supérieur dédié, associé avec une illumination face arrière permet d'augmenter la surface photosensible de 35% pour un pixel 1.4 μm et d'éloigner la diode des sources de courant d'obscurité. La taille des contacts nécessaire pour relier les deux niveaux empilés (< 100nm) exclut une intégration de type 3D parallèle. Nous avons montré que les transistors à bas budget thermique ont des niveaux de bruit compatible avec une application imageur. Les imageurs CMOS fortement miniaturisés semblent être une niche pour l'intégration 3D séquentielle.

Ces deux applications valorisent nettement l'intégration séquentielle par rapport à l'intégration parallèle car elles requièrent des interconnexions à l'échelle du transistor pour être réalisées.

D'autres applications peuvent potentiellement constituer un marché pour l'intégration 3D séquentielle : (i) les mémoires FLASH, domaine d'application pour lequel la densité est un argument majeur, (ii) les co-intégrations de technologies hétérogènes diverses, e.g. co-intégration Si-Ge pour des applications photoniques ou encore des co-intégrations SOI/Si massif.

Même si nous avons privilégié l'étude de transistors FDSOI sur les deux niveaux, l'intégration séquentielle n'est pas limitée à cette application. L'intégration 3D séquentielle peut être élargie à des transistors partiellement déplétés (PDSOI) et des transistors sur substrat massif.

Perspectives :

Nous avons identifié des développements et études supplémentaires qui présentent un intérêt pour certaines applications :

- Optimisation des modules de jonction :

- Une étude de la désactivation des dopants du niveau inférieur lié au budget thermique du niveau supérieur, ainsi qu'une étude sur la désactivation des jonctions SPER du niveau supérieur avec les étapes de back-end.

- Un travail d'ingénierie de jonctions avec le procédé SPER, ainsi qu'une étude sur les courants de jonction permettant de statuer sur la pertinence de ce procédé pour des applications à très basse consommation.

- Développements de contacts innovants :

- Pour des applications hautes densités, il serait souhaitable de développer un contact de type « traversant », plus compact que le contact dit « chevauchant ».

- Le développement d'un contact « interne », confiné entre les niveaux de transistor inférieur et supérieur ne remontant pas jusqu'au niveau métal 1 peut permettre de libérer un espace pour le routage. Cette interconnexion locale entre deux niveaux facilite la connexion entre le niveau n et le niveau n+1 dans une intégration à plus de deux niveaux.

- Développement de lignes métalliques entre deux niveaux de transistor.

La réalisation de ligne de métal entre les niveaux de transistors empilés, permettra de réaliser un placement routage en 3D. Dans ce cas, l'architecture 3D séquentielle pourrait répondre à la problématique de gain en délai dans les interconnexions.

- Pour pouvoir conclure sur le gain en densité dans des circuits intégrés, il serait nécessaire de réaliser une étude de densité sur des cellules beaucoup plus complexes routées au moins jusqu'aux niveaux de métal M4 -M5.

- Pour pouvoir conclure sur le surcoût de l'intégration 3D séquentielle, il est nécessaire de mener ces études sur des applications précises.

Au terme de ces études préliminaires, l'intégration 3D séquentielle semble *a priori* prometteuse : En effet, nous avons mis en évidence des perspectives de gain (coût, surface, performance) pour des applications logiques haute densité, et identifié d'autres applications justifiant sa réalisation. De plus nous avons démontré la faisabilité de cette intégration. Ce travail engage donc à poursuivre son étude.

Pour statuer sur l'intérêt de l'intégration séquentielle et quantifier précisément les gains apportés par cette technologie, les études doivent être dorénavant menées dans une optique plus industrielle, en intégrant des paramètres non étudiés pendant cette thèse, comme l'aspect de rendement, de variabilité et de fiabilité.

De plus, il serait souhaitable de considérer des applications précises. En travaillant dans un cadre moins général, la limitation des hypothèses et des paramètres des études permettra d'apporter des conclusions quantitatives fiables.

PROCEEDINGS DE CONFERENCES ET JOURNAUX :**PAPIERS EN AUTEUR PRINCIPAL :****1/ In-depth investigation of the mechanisms impacting C-V/G-V characteristics of Ge/GeON/HfO₂/TiN stacks by electrical modeling**

P. Batude, X. Garros, L. Clavelier, C. Le Royer, J. M. Hartmann, V. Loup, P. Besson, L. Vandroux, Y. Campidelli, S. Deleonibus, F. Boulanger, *Microelectronic engineering* 84 (2007)

2/ Insights on fundamental mechanisms impacting Ge metal oxide semiconductor capacitors with high-k/metal gate stacks

P. Batude, X. Garros, L. Clavelier, C. Le Royer, J. M. Hartmann, V. Loup, P. Besson, L. Vandroux, Y. Campidelli, S. Deleonibus, F. Boulanger, *J. Appl. Phys.* 102, 034514 (2007)

3/Enabling 3D monolithic Integration

P. Batude, M. Vinet, A. Pouydebasque, L. Clavelier, C. LeRoyer, C.Tabone, B. Previtali, L.Sanchez , L. Baud, A. Roman, V. Carron, F. Nemouchi, S. Pocas, C. Comboroure, V. Mazzocchi, H. Grampeix, F.Aussenac and S. Deleonibus, *ECS journal VO16*, pp47 (2008)

4/ 3D CMOS integration: Introduction of dynamic coupling and application to compact and robust 4T SRAM

P. Batude, M.-A. Jaud, O. Thomas, L. Clavelier, A. Pouydebasque, M. Vinet, S. Deleonibus, A. Amara, *IEEE International Conference on Integrated Circuit Design*, pp81 (2008)

5/GeOI and SOI 3D Monolithic Cell integrations for High Density Applications,

P. Batude, M. Vinet, A. Pouydebasque, C. Le Royer, B. Previtali, C. Tabone, L. Clavelier, S. Michaud, A. Valentian, O. Thomas, O. Rozeau, P. Coudrain*, C. Leyris*, K. Romanjek, X. Garros, L. Sanchez, L. Baud, A. Roman, V. Carron, H. Grampeix, E. Augendre, A. Toffoli, F. Allain, P. Grosgeorges, V. Mazzocchi, L. Tosti, F. Andrieu, J.-M. Hartmann, D. Lafond, S. Deleonibus and O. Faynot. *Digest of technical papers symposium on VLSI technology* (2009)

6/ Advances in 3D CMOS Sequential Integration

P. Batude, M. Vinet, A. Pouydebasque, C. Le Royer, B. Previtali, C. Tabone, J.-M. Hartmann, L. Sanchez, L. Baud, V. Carron, A. Toffoli, F. Allain, V. Mazzocchi, D. Lafond, N. Bouzaida, O. Thomas, O. Cueto, A. Amara¹, S. Deleonibus and O. Faynot; *IEEE International device Meeting (IEDM)* (2009), papier accepté

AUTRES PAPIERS**7/Hard X-ray photoemission experiments on novel Ge-based metal gate/high-k stacks**

Rubio-Zuazo, E. Martinez, **P. Batude**, L. Clavelier, F. Soria, A. Chabli, G. R. Castro, *International Conference on Frontiers of Characterization and Metrology. AIP Conference Proceedings, Volume 931*, pp. 329-333 (2007).

8/ Setting up 3D Sequential Integration for Back-Illuminated CMOS Image Sensors with Highly Miniaturized Pixels with low temperature Fully Depleted SOI transistors

P. Coudrain, **P. Batude**, X. Gagnard, C. Leyris, S. Ricq, M. Vinet, A. Pouydebasque, N. Moussy, Y. Cazaux, B. Giffard, P. Magnan, P. Ancey, P. Magnan, *IEEE International Electron Device Meeting IEDM 2008*

9/High-k and Metal-Gate pMOSFETs on GeOI Obtained by Ge Enrichment: Analysis of ON and OFF Performances

Le Royer, C. Vincent, B. Clavelier, L. Damlencourt, J.-F. Tabone, C. **Batude, P.** Blachier, D. Truche, R. Campidelli, Y. Nguyen, Q.T. Cristoloveanu, S. Soliveres, S. Le Carval, G. Boulanger, F. Billon, T. Bensahel, D. Deleonibus, *IEEE Electron Device Letters*, vol. 29, issue 6, pp. 635-637(2008)

10/System on Wafer : a new silicon concept in SiP

G. Poupon, N. Sillon, D. Henry, C. Gillot, A Mathewson, L. Di Ciccio, B. Charlet, P. Leduc, M. Vinet, **P. Batude**, *Proceedings of the IEEE, special issue 3-D integration technologies*, Volume: 97, Issue: 1 (2009)

11/Enabling 3D-Integrated Back-Illuminated CMOS Image Sensors for Highly Miniaturized Pixels,

P. Coudrain, **P. Batude**, X. Gagnard, C. Leyris, M. Vinet, A. Castex, C. Lagahe-Blanchard, A. Pouydebasque, Y. Cazaux, B. Giffard, P. Ancey, P. Magnan, *IEEE Transactions on Electronic devices 2009, papier accepté.*

12/Compact 6T SRAM cell with robust Read/Write stabilizing design in 45nm Monolithic 3D IC technology,

O. Thomas, M. Vinet, O. Rozeau, **P. Batude**, A. Valentian, *IEEE International Conference on Integrated Circuit Design (2009)*

13/Towards 100 nm Contact Technology for Three-Dimensional Back-Illuminated Miniaturized Pixels,

Perceval Coudrain, P. Magnan, **P. Batude**, X. Gagnard, C. Leyris, L. Depoyan, Y. Cazaux, B. Giffard, P. Ancey, *IEEE International Image Sensor Workshop (2009)*

PAPIERS INVITES:**14/ 6th International Conference on Silicon Epitaxy and Heterostructures (ICSI-6):**

“Co-integration (including 3D /vertical) of SOI/GOI”, M. Vinet, C. Le Royer, **P. Batude**, J.-M. Hartmann, C. Tabone, J.-F. Damlencourt, B. Vincent, A. Pouydebasque, B. Previtali, L. Sanchez, L. Baud, A. Roman, V. Carron, F. Nemouchi, P. Scheiblin, Y. LeCunff, O. Thomas, E. Augendre, V. Mazzocchi, H. Grampeix, O. Faynot.

15/ Electro-Chemical Society fall meeting 09:

“3D monolithic integration of Si nMOSFET with Ge pMOSFET” M. Vinet, **P. Batude**, C.Tabone, B. Previtali, C. LeRoyer, A. Pouydebasque, L. Clavelier,A. Valentian, O. Thomas, S. Michaud, L. Sanchez, L. Baud, A. Roman, V. Carron, F. Nemouchi, S. Pocas, C. Comboroure, V. Mazzocchi, H. Grampeix, F.Aussenac, A. Amara and S. Deleonibus.

16/ System design for 3D silicon intergration Workshop (D43D 09): Titre, auteur, présentateur non défini**17/ European Materials Research Society (E-MRS 2010):** Titre, auteur, présentateur non défini**CONFÉRENCES:**

- 1/ The 15th biennial Conference on Insulating Films on Semiconductors (INFOS) (Athens 2007)
- 2/ The 2008 International Conference on IC Design & Technology (ICICDT) (Grenoble 2008)
- 3/ The Tenth International Symposium on Semiconductor Wafer Bonding joint to the 214th International Meeting of ECS (Honolulu 2008)
- 4/ The 2009 Symposium on VLSI Technology (Kyoto 2009)
- 5/ The 2009 International Electron Device Meeting (Baltimore 2009)

BREVETS :

- 1/ DD 10575 VR : Circuit à transistors intégrés dans trois dimensions et ayant une tension de seuil ajustable dynamiquement, **P.Batude**, M-A.Jaud, L.Clavelier.
- 2/ DD 10575 VR : Cellules SRAMs comportant des transistors intégrés dans trois dimensions et ayant une tension de seuil ajustable dynamiquement, O.Thomas, **P.Batude**, M. Vinet, A. Pouydebasque.

A

- [Abou-Samra]: S-J. Abou-Samra, J.Arweiler, A.Guyot, « Low power SOI CMOS Multipliers 2D vs 3D », IEEE European Solid-State CIRcuits Conference (ESSIRC), p196 (1998)
- [Akasaka86]: Y. Akasaka, T. Nishimura, "Concept and basics technologies for 3-D IC Structure", Technical Digest of International Electron Devices Meeting, pp 488 (1986)
- [Akiyama83]: S. Akiyama, S. Ogawa, M. Yoneda, N. Yoshii, Y. Terui; "Multilayer CMOS device fabricated on laser recrystallised silicon islands"; IEEE Internatonal Electron Device Meeting (IEDM); pp352 (1983)
- [Akiyama84]: S. Akiyama, M. Yoneda, S. Ogawa, N. Yoshii, Y. Terui, "Fabrication technologies for multilayer CMOS device," Proceedings of the Symposium on Very Large Scale Integrations (VLSI) technology, pp. 28 (2004).
- [Ando88]: M Ando, T Okazawa, H Furuta, M. Ohkawa, J. Monden, N. Kodama, K. Abe, H. Ishihara, IS Sasaki; "A 0.1 uA Standby Current, Bouncing-Noise-Immune 1 Mb SRAM"; technical digest of Symposium. on VLSI Technology, pp49 (1988)
- [Arnaud04] : F. Arnaud, B. Duriez, B. Tavel, L. Pain, J. Todeschini, M. Jurdit, Y. Laplanche, F. Boeuf, F. Salvetti, D. Lenoble, J.P. Reynard, F. Wacquant, P. Morin, N. Emonet, D. Barge, M. Bidaud, D. Ceccarelli, P. Vannier, Y. Loquet, H. Leninger, F. Judong, C. Perrot, I. Guilmeau, R. Palla, A. Beverina, V. DeJonghe, M. Broekaart, V. Vachellerie, R.A. Bianchi, B. Borot, T. Devoivre, N. Bicais, D. Roy, M. Denais, K. Rochereau, R. Difrenza, N. Planes, H. Brut, L. Vishnobulta, D. Reber, P. Stolk, M. Woo; " Low cost 65nm CMOS platform for Low Power & General Purpose applications"; Symposium on VLSI Technology Digest of Technical Papers, pp10 (2004)
- [Arnaud 08] : F.Arnaud, J.Liu, Y.M.Lee, K.Y.Lim, S.Kohler, J.Chen, B.K.Moon, C.W. Lai, M.Lipinski, L.Sang, F.Guarin, C.Hobbs, P.Ferreira, K.Ohuchi, J.Li, H.Zhuang, P.Mora, Q. Zhang, D.R.Nair, D.H. Lee, K.K.Chan, S.Satadru, S.Yang, J.Koshy, W.Hayter, M.Zaleski, D.V.Coolbaugh, H.W.Kim, Y.C. Ee, J.Sudijono, A.Thean, M.Sherony, S.Samavedam, M.Khare, C.Goldberg, A.Steege; "32nm General Purpose Bulk CMOS Technology for High Performance Applications at Low Voltage"; IEEE International Electron Device Meeting (IEDM), pp 1 (2008)
- [Assaderaghi97]. Assaderaghi , D. Sinitsky , S. A. Parke , J. Bokor , P. K. Ko and C. Hu "Dynamic threshold-voltage MOSFET (DTMOS) for ultra-low voltage VLSI," IEEE Transactions on Electron Devices, vol. 44, pp. 414, Mar. 1997.
- [Auth08]: C.Auth, A. Cappellani, J-S. Chun, A. Dalis, A. Davis, T. Ghani, G. Glass, T. Glassman, M. Harper, M. Hattendorf, P. Hentges, S. Jaloviar, S. Joshi, J. Klaus, K. Kuhn, D. Lavric, M. Lu, H. Mariappan, K. Mistry, B. Norris, N. Rahhal-orabi, P. Ranade, J. Sandford, L. Shifren, V. Souw, K. Tone, F. Tambwe, A. Thompson, D. Towner, T. Troeger, P. Vandervoorn, C. Wallace, J. Wiedemer, C. Wiegand; "45nm High-k + metal gate strain-enhanced transistors"; Symposium on VLSI Technology Digest of Technical Papers, pp128 (2008)

B

- [Bai06] W. P. Bai, N. Lu, A. Ritenour, M. L. Lee, D. A. Antoniadis, and D. L. Kwong, "Ge n-MOSFETs on lightly doped substrates with high-kappa dielectric and TaN gate," *Ieee Electron Device Letters*, vol. 27, pp. 175-178, 2006.
- [Banerjee01]: K. Banerjee, S- J. Souri, P. Kapur, K-C. Saraswat; "3-D ICs: A Novel Chip Design for Improving Deep-Submicrometer Interconnect Performance and Systems-on-Chip Integration"; *Proceedings of the IEEE*, vol 89, pp602 (2001)
- [Batude 08b]: P. Batude, M. Vinet, A. Pouydebasque, L. Clavelier, C. LeRoyet, C.Tabone, B. Previtali, L. Sanchez , L. Baud, A. Roman, V. Carron, F. Nemouchi, S. Pocas, C. Comboroure, V. Mazzocchi, H. Grampeix, F.Aussenac, S. Deleonibus; "Enabling 3D monolithic integration"; *Pocceedings of the Electro-Chemical Society (ECS) spring meeting VOL 16 pp47 (2008)*.
- [Batude09a]: P. Batude, M. Vinet, A. Pouydebasque, C. Le Royer, B. Previtali, C. Tabone, L. Clavelier, S. Michaud, A. Valentian, O. Thomas, O. Rozeau, P. Coudrain*, C. Leyris*, K. Romanjek, X. Garros, L. Sanchez, L. Baud, A. Roman, V. Carron, H. Grampeix, E. Augendre, A. Toffoli, F. Allain, P. Grosgeorges, V. Mazzochi, L. Tosti, F. Andrieu, J.-M. Hartmann, D. Lafond, S. Deleonibus O. Faynot; "GeOI and SOI 3D Monolithic Cell integrations for High Density Applications"; *Technical Digest of the Symposium on Very Large Scale Integration (VLSI)*, pp166 (2009)
- [Batude09b]: P. Batude, M. Vinet, A. Pouydebasque, C. Le Royer, B. Previtali, C. Tabone, J.-M. Hartmann, L. Sanchez, L. Baud, V. Carron, A. Toffoli, F. Allain, V. Mazzocchi, D. Lafond, N. Bouzaida, O. Thomas, O. Cueto, A. Amara, S. Deleonibus, O. Faynot, "Advances in 3D CMOS Sequential Integration"; *accepté à IEEE International Electron Device Meeting (IEDM) (2009)*
- [Bhavnagarwala05]: A. Bhavnagarwala, S. Kosonocky, C. Radens, K. Stawiasz, R. Mann, Q.Ye, "Fluctuation limits & scaling opportunities for CMOS SRAM cells", *IEEE International Electron Devices Meeting Technical Digest (IEDM)* , pp 659(2005)
- [Boeuf05] : F.Boeuf, F. Arnaud, C. Boccaccio, F. Salvetti, J. Todeschini, L. Pain, M. Jurdit, S. Manakli, B. Icard, N. Planes, N. Gierczynski, S.Denorme, B. Borot, C.Ortolland, B. Duriez, B. Tavel, P. Gouraud, M. Broekart, V. Dejonghe, P. Brun, F. Guyader, P.Morin, C. Reddy, M. Aminpur, C. Laviro, S. Smith, J.P. Jacquemin, M. Mellier, F. André, N. Bicais-Lepinay, S. Jullian, J. Bustos and T.Skotnicki. "0.248 μm^2 and 0.334 μm^2 Conventional Bulk 6T-SRAM bit -cells for 45nm node Low Cost - General Purpose Applications", *Symposium on VLSI Technology Digest of Technical Papers*, pp130 (2005)
- [Bonnetier07]: S. Bonnetier, B. Imbert, M Hopstaken, D. Galpin, R. Gwoziecki, D. Barge, S. Zoll, O. Anilturk, E. Sicurani, C. Caillat, A. Barr, R. Gonell, Y. Espinoux, Pi. Mur, Nicolas Mayet, Andrea Gotti, M-T Basso; 'Silicide pre-clean effects on NiPtSi thermal stability for 65 nm technologies and beyond'; *Microelectronic Engineering*, Volume 84, Issue 11, pp2528, *European Workshop on Materials for Advanced Metallization* (2007)
- [Brue195] : M.Brue1, « Silicon On Insulator material technology » *Electronics Letters*, 31, 14, p1201 (1995)

C

- [Chan 01 d]: V-W-C Chan, P-C-H Chan, M. Chan; "3D integrated circuit using large grain polysilicon film"; *Proceedings of the International conference on Solid-State and Integrated-Circuit Technology*, pp58 (2001)

- [Chan 01a]: V-W-C Chan, P-C-H Chan, M. Chan; "Multiple layers of CMOS integrated circuits using recrystallized silicon film"; IEEE Electron Device Letters, Volume 22, Issue 2, pp 77 (2001)
- [Chan 01b]: V-W-C Chan, P-C-H Chan, M. Chan, "Three-dimensional CMOS SOI integrated circuit using high-temperature metal-induced lateral crystallization"; IEEE Transactions on Electron Devices, Volume 48, Issue 7, pp 1394 (2001).
- [Chan 01c]: M. Chan; "The potential and realization of multi-layers three-dimensional integrated circuit"; Proceedings of the Solid-State and Integrated-Circuit Technology conference pp 40 (2001)
- [Chan00]: V-W-C. Chan, P-C-H. Chan, M. Chan; "Three dimensional CMOS integrated circuits on large grain polysilicon films"; Technical Digest of International Electron Devices Meeting (IEDM) pp161 (2000)
- [Chen83]: C-E. Chen, H-W. lam, S-D-S. Malhi, R-F. Pinizzotto; "Stacked CMOS SRAM cell", IEEE Electron Device Letters, VOL EDL-4, NO.8, pp 272 (1983).
- [Chen07]: C.K. Chen, K. Warner, D.R.W. Yost, J.M. Knecht, V. Suntharalingam, C.L. Chen, J.A. Burns, C.L. Keast; "Scaling Three-Dimensional SOI Integrated-Circuit Technology"; Proceedings of the IEEE International SOI Conference pp87 (2007)
- [Cheng00] L. W. Cheng, S. L. Cheng, L. J. Chena, H. C. Chien, H. L. Lee F. M. Pan, "Formation of Ni silicides on (001)Si with a thin interposing Pt layer", J. Vac. Sci. Technol. A Volume 18, Issue 4, pp. 1176 (2000)
- [Cheng07]: Kuan-Lun Cheng; Wu, C.C.; Wang, Y.P.; Lin, D.W.; Chu, C.M.; Tarn, Y.Y.; Lu, S.Y.; Yang, S.J.; Hsieh, M.H.; Liu, C.M.; Fu, S.P.; Chen, J.H.; Lin, C.T.; Lien, W.Y.; Huang, H.Y.; Wang, P.W.; Lin, H.H.; Lee, D.Y.; Huang, M.J.; Nieh, C.F.; Lin, L.T.; Chen, C.C.; Chang, W.; Chiu, Y.H.; Wang, M.Y.; Yeh, C.H.; Chen, F.C.; Chang, Y.H.; Wang, S.C.; Hsieh, H.C.; Lei, M.D.; Goto, K.; Tao, H.J.; Cao, M.; Tuan, H.C.; Diaz, C.H.; Mii, Y.J.; Wu, C.M.; "Highly Scaled, High Performance 45nm Bulk Logic CMOS Technology with 0.242 μm^2 SRAM Cell"; IEEE International Electron Devices Meeting Technical Digest (IEDM) , pp 243 (2007)
- [Chui03]: C-O. Chui, H. Kim, D. Chi, B-B. Triplett, P-C McIntyre, K-C. Saraswat; "A sub-400°C Germanium MOSFET Technology with High-K dielectric and Metal Gate"; IEEE International Electron Device Meeting (IEDM), pp437 (2002)
- [Colinge 81a]: J-P. Colinge, E. Demoulin, "A high Density Inverter with stacked Transistors", IEEE Electron device Letters, VOL. EDL-2 NO.10 (1981)
- [Colinge 81b]: J-P. Colinge, E. Demoulin, "ST-CMOS (Stacked transistors CMOS): a double poly nMOS compatible CMOS technology", Technical Digest of International Electron Devices Meeting pp 557 (1981)
- [Colinge04]: J-P. Colinge, Silicon-On Insulator Technology: Materials to VLSI », 3rd edition, Kluwer academic publishers. (2004)

- [Colinge82a]: J-P. Colinge, E. Demoulin, M. Lobet: "Stacked Transistors CMOS (ST-MOS), an NMOS technology modified to CMOS", IEEE transactions on Electron Devices, VOL. ED-29, NO4 (1982)
- [Coudrain 08] : P Coudrain, P. Batude, X. Gagnard, C. Leyris, S. Ricq, M. Vinet, A. Pouydebasque, N. Moussy, Y. Cazaux, B. Giffard, P. Magnan, P. Ancey ; « Setting up 3D Sequential Integration for Back-Illuminated CMOS Image Sensors with Highly Miniaturized Pixels with low temperature Fully Depleted SOI transistors" IEEE International Electron Device Meeting (IEDM),PP 271 (2008)
- [Coudrain 09] : P. Coudrain, P. Batude, X. Gagnard, C. Leyris, M. Vinet, A. Castex, C. Lagahe-Blanchard, A. Pouydebasque, Y. Cazaux, B. Giffard, P. Ancey, P. Magnan; "Enabling 3D-Integrated Back-Illuminated CMOS Image Sensors for Highly Miniaturized Pixels"; accepté dans Transactions on Electron devices (2009)
- [CTRIM] : M. Posselt, Radiation Effects and Defects in Solids,Vol. 130-131, pp. 87-119, 1994

D

- [Daschs03]: C. Dachs, R. Surdeanu, B. Pawlak, G. Doornbos, R. Duffy, A. Heringa, Y. Ponomarev, V. Venezia, M. Van Dal, P. Stolk, R. Lindsay, K. Henson, B. Dieu, L. Geenen, I. Hoflijik, O. Richard, T. Clarysse, B. Brijs, W. Vandervorst, X. Pagès; "CMOS scaling beyond the 90 nm CMOS technology node: shallow junction and integration challenges"; Ultra Shallow Junctions. 7th Int. Worksh. Fabrication, Characterization and Modeling of Ultra Shallow Doping Profiles in Semiconductor. pp15 (2003)
- [Deduytsche07] : D. Deduytsche, C. Detavernier, R. L. Van Meirhaeghe, J. L. Jordan-Sweet and C. Lavoie, "Formation and morphological stability of NiSi in the presence of W, Ti, and Ta alloying elements", , Volume 101 Issue 4 (2007)
- [Detavernier04] : C. Detavernier, C. Lavoie, "Influence of Pt addition on the texture of NiSi on Si(001)"Appl. Phys. Lett. 84, pp 3549, (2004)
- [Detavernier06]: C. Detavernier, D. deduytsche, J. Jordan-Sweet, C-J. Cabral, C. Lavoie; "Influence of alloying elements on the formation and stability of NiSi; ECS Transactions, 3(2) pp131 (2006)
- [Dumont07]: Benjamin Dumont, "Etude et intégration de jonctions ultrafines pour les technologies CMOS 45 et en deçà", thèse de doctorat (2007)
- [Devoivre02]: T. Devoivre, M. Lunenburg, C. Julien, J-P. Carrere, P. Ferreira, W.J. Toren, A. VandeGoor, P. Gayet, T. Berger, O. Hinsinger, P. Vannier, Y. Trouiller, Y. Rody, P-J. Goirand, R. Palla, I. Thomas, F. Guyader, D. Roy, B. Borot, N. Planes, S. Naudet, F. Pico, D. Duca, F. Lalanne, D. Heslinga, M. Haond; "Validated 90nm CMOS technology platform with low-k copper interconnects for advanced system-on-chip (SoC)", IEEE International Workshop on Memory Technology, Design and Testing. (MTDT), pp 157 (2002)

F

[Fazzini08]: P. F. Fazzini, F. Cristiano, C. Dupré, A. Claverie, T. Ernst, M. Gavelle ; « Defect evolution after germanium preamorphization in silicon on insulator structures »; Journal of Vacuum Science and Technology, pp342 (2008)

[Feng 06]: J. Feng; Y. Liu, P-B. Griffin, J-D. Plummer; “Integration of Germanium-on-Insulator and Silicon MOSFETs on a Silicon Substrate”; IEEE Electron Device Letters, Volume 27, Issue 11, pp 911 (2006)

[Fleury 09]: D. Fleury, A. Cros, G. Bidal, J. Rosa, G. Ghibaudo; “Technique to Extract the Source/Drain Series Resistance of MOSFETs”; IEEE Electron Device Letter 0741-3106 (2009)

G

[Garrou08]: Philip Garrou; Christopher Bower; Peter Ramm « 3D Integration : Technology and Applications », John Wiley & Sons, 2008

[Garros08]: X. Garros, M. Casse, G. Reimbold, F. Martin, C. Leroux, A. Fanton, O. Renault, V. Cosnier, F. Boulanger; “Guidelines to improve mobility performances and BTI reliability of advanced high-k/metal gate stacks”; Proceedings of the Symposium on Very Large Scale Integration (VLSI), pp68 (2008)

[Ghani03]: T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, M. Bohr; “A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors” IEEE International Electron Devices Meeting Technical Digest (IEDM) , pp 11.6.1 (2003)

[Ghibaudo88] : G. Ghibaudo, new method for the extraction of MOSFET parameters, electronic Letters, vol.24,no.9,p.543, 1988.

[Gibbons80] : J-F. Gibbons, K-F. Lee, ‘One gate Wide CMOS inverter on Laser-Recrystallized Polysilicon’, IEEE Electron Device Letters, VOL1, NO6 (1980)

[Gibbons 82]: J-F. Gibbons, K-F. Lee, “A folding principle for generating three-dimensional MOSFET device structures in beam-recrystallized polysilicon films” Technical Digest of International Electron Devices Meeting pp111 (1982)

[Goeloe91]: G-T. Goeloe, E-W. Maby, D-J. Silversmith, R-W. Mountain, D-A. Antoniadis; “Vertical single-gate CMOS inverters on laser-processed multilayer substrates”; Technical Digest of International Electron Devices Meeting pp 554 (1981)

H

- [Haran08] : B. S. Haran, A. Kumara, L. Adam, J. Changa, V. Basker, S. Kanakasabapathy, D. Horak, S. Fan, J. Chen, J. Faltermeier, S. Seo, M. Burkhardt, S. Burns, S. Halle, S. Holmes, R. Johnson, E. McLellan, T. M. Levin, Y. Zhu, J. Kuss, A. Ebert, J. Cummings, D. Canaperi, S. Paparao, J. Arnold, T. Sparks, C. S. Koay, T. Kanarsky, S. Schmitz, K. Petrillo, R. H. Kim, J. Demarest, L. F. Edge, H. Jagannathan, M. Smalley, N. Berliner, K. Cheng, D. LaTulipe, C. Koburger, S. Mehta, M. Raymond, M. Colburn, T. Spooner, V. Paruchuri, W. Haenscha, D. McHerron, and B. Doris; "22 nm Technology Compatible Fully Functional 0.1 μm^2 6T-SRAM Cell"; IEEE International Electron Device Meeting (IEDM), pp 1 (2008)
- [Hamilton06] : J. J. Hamilton, N. E. B. Cowern, J. A. Sharp, K. J. Kirkby, E. J. H. Collart, B. Colombeau, M. Bersani, D. Giubertoni, A. Parisini; "Diffusion and activation of ultrashallow B implants in silicon on insulator: End-of-range defect dissolution and the buried Si/SiO₂ interface; Applied Physics Letters, 89 (042111), (2006)
- [Hammes07] : M. Hammes, C. Kranz, J. Kissing, D. Seippel, P-H. Bonnaud, E. Pelos; "A GSM Baseband Radio in 0.13 μm CMOS with Fully Integrated Power-Management"; Technical Digest of the International Solid State Circuit conference (ISSCC), pp264 (2007)
- [Hartmann08a] : J-M. Hartmann, F. Gonzatti, T. Billon; "Growth kinetics and boron-doping of very high Ge content SiGe for source/drain engineering; Journal of Crystal Growth 310, 62 (2008)
- [Hartmann08b] : J-M. Hartmann, F. Anduier, D. Lafond, T. Ernst, Y. Bogumilowicz, V. Delaye, O. Weber, D. Rouchon, A-M. Papon; Material Science engineering, B 76 pp154 (2008)
- [Hartmann09] : J-M. Hartmann, A. Abbadie, N. Cherkashin, H. Grampeix, L. Clavelier; "Epitaxial growth of Ge thick layers on nominal and 6° off So (110); Ge surface passivation by Si"; Journal of Semiconductor Science and Technology, vol 24, pp1 (2009)
- [Hasegawa08] : S. Hasegawa, Y. Kitamura, K. Takahata, H. Okamoto, T. Hirai, K. Miyashita, T. Ishida, H. Aizawa, S. Aota, A. Azuma, T. Fukushima, H. Harakawa, E. Hasegawa, M. Inohara, S. Inumiya, T. Ishizuka, T. Iwamoto, N. Kariya, K. Kojima, T. Komukai, N. Matsunaga, S. Mimotogi, S. Muramatsu, K. Nagatomo, S. Nagahara, Y. Nakahara#, K. Nakajima, K. Nakatsuka, M. Nishigoori, A. Nomachi, R. Ogawa, N. Okada, S. Okamoto, K. Okano, T. Oki, H. Onoda, T. Sasaki, M. Satake, T. Suzuki, Y. Suzuki, M. Tagami, K. Takeda, M. Tanaka, K. Taniguchi, M. Tominaga, G. Tsutsui, K. Utsumi, S. Watanabe, T. Watanabe, Y. Yoshimizu, T. Kitano, H. Naruse, Y. Goto, T. Nakayama, N. Nakamura, F. Matsuoka; "A Cost-Conscious 32nm CMOS Platform Technology with Advanced Single Exposure Lithography and Gate-First Metal Gate/High-K Process"; IEEE International Electron Device Meeting (IEDM), pp 1 (2008)
- [Hayakawa90] : S. Hayakawa, M. Kakumu, H. Takeuchi, K. Sato, T. Ohtani, T. Yoshida, T. Nakayama, S. Morita, M. Kinugawa, K. Maeguchi, K. Ochi, J. Matsunaga, A. Aono, K. Noguchi, T. Asami; "A 1 μm A retention 4 Mb SRAM with a thin-film-transistor load cell"; Digest of Technical Papers of Solid-State Circuits Conference (ISSCC), pp128 (1990)
- [Hirano08] : Y. Hirano, M. Tsujiuchi, Y. Maki, T. Iwamatsu, Y. Ishii, A. Miyanishi, Y. Tsukamoto, K. Nii, T. Ipposhi, H. Oda, S. Maegawa, Y. Inoue, "Novel Low-Power and High-Speed SOI SRAM With Actively Body-Bias Controlled (ABC) Technology for Emerging Generations", IEEE Transactions on Electron Devices, Volume: 55, Issue: 1, pp 365 (2008).
- [Huan03] : C. H. Huang, D. S. Yu, A. Chin, C. H. Wu, W. J. Chen, C. X. Zhu, M. F. Li, B. J. Cho, and D. L. Kwong, "Fully silicided NiSi and germanided NiGe dual gates on SiO₂/Si and Al₂O₃/Ge-on-insulator MOSFETs," 2003 IEEE International Electron Devices Meeting, Technical Digest, pp. 319-322, 2003.

I

- [Im00]: S. Im, K. Banerjee; "Full Chip Thermal Analysis of Planar (2-D) and Vertically Integrated (3-D) High Performance Ics"; IEEE International Electron Device Meeting (IEDM), pp727 (2000)
- [Imbert05] : Bruno Imbert, « Etude de la formation du siliciure de nickel-platine intégré dans la fabrication de transistors CMOS pour les technologies 65 et 45 nm », thèse de doctorat (Institut polytechnique de Grenoble, 2005)
- [[Imbert09] : Bruno Imbert, « Etude de la formation du siliciure de nickel-platine intégré dans la fabrication de transistors CMOS pour les technologies 65 et 45 nm », thèse de doctorat (Institut polytechnique de Grenoble, 2009)
- [Inoue86]: Y. Inoue, K. Sugahara, S. Kusunoki, M. Nakaya, T. Nishimura, Y. Horiba, Y. Akasaka, Y. H. Nakata; "A three dimensional static RAM"; Electron Device Letter, Volume: 7, Issue: 5, pp327 (1986)
- [Inoue89]: Y. Inoue, "4PMOS-2NMOS vertically stacked CMOS-SRAM with 0.6 μm design rule, VLSI pp39 (1989)
- [Ishigaki08] : T. Ishigaki¹, R. Tsuchiya¹, Y. Morita¹, H. Yoshimoto¹, N. Sugii¹, T. Iwamatsu², H. Oda², Y. Inoue², T. Ohtou³, T. Hiramoto³, and S. Kimura, "Silicon on Thin BOX (SOTB) CMOS for Ultralow Standby Power with Forward-biasing Performance Booster" , Solid-State Device Research Conference ESSDERC, page 198, (2008)

[ITRS]: <http://www.itrs.net/>

J

- [Jagar99]: S. Jagar, M. Chan, M-C. Poon, W. Hongmei, Q. Ming, P-K. Ko, W. Yangyuan; "Single grain thin-film-transistor (TFT) with SOI CMOS performance formed by metal-induced-lateral-crystallization"; Technical Digest of International Electron Devices Meeting (IEDM) pp293 (1999)
- [Jang 04]: J-H. Jang, S-M. Jung, Y-H. Kang, W-S. Cho, J-H. Moon, C-D. Yeo, K-H. Kwak, B-H. Choi, B-J. Hwang, W-R. Jung, S-J. Kim, J-R. Kim, J-H. Na, H. Lim, J-H. Jeong, K. Kim; "Novel 3-dimensional 4F² SRAM technology with 0.294 μm^2 S³ (stacked single-crystal Si) cell and SSTFT (stacked single-crystal thin film transistor)"; Proceedings of Solid-State Device Research conference (ESSDERC), pp 445 (2004).
- [Josse06] : Josse, E.; Parihar, S.; Callen, O.; Ferreira, P.; Monget, C.; Farcy, A.; Zaleski, M.; Villanueva, D.; Ranica, R.; Bidaud, M.; Barge, D.; Laviron, C.; Auriac, N.; Le Cam, C.; Harrison, S.; Warrick, S.; Leverd, F.; Gouraud, P.; Zoll, S.; Guyader, F.; Perrin, E.; Baylac, E.; Belledent, J.; Icard, B.; Minghetti, B.; Manakli, S.; Pain, L.; Huard, V.; Ribes, G.; Rochereau, K.; Bordez, S.; Blanc, C.; Margain, A.; Delille, D.; Pantel, R.; Barla, K.; Cave, N.; Haond, M.; Cost-Effective Low Power Platform for the 45-nm Technology Node IEEE International Electron Devices Meeting Technical Digest (IEDM) , pp 1 (2006)

- [Jung 04]: S-M. Jung, J. Jang, W. Cho, J. Moon, K. Kwak, B. Choi, B. Hwang, H. Lim, J. Jeong, J. Kim, K. Kim; "The revolutionary and truly 3-dimensional 25F² SRAM technology with the smallest S³ (stacked single-crystal Si) cell, 0.16um², and SSTFT (stacked single-crystal thin film transistor) for ultra high density SRAM"; Digest of Technical Papers of Symposium on VLSI Technology pp228 (2004)
- [Jung 05]: S-M. Jung, Y. Rah, T. Ha, H. Park, C. Chang, S. Lee, J. Yun, W. Cho, H. Lim, J. Park, J. Jeong, B. Son, J. Jang, B. Choi, H. Cho, K. Kim; "Highly cost effective and high performance 65nm S³ (stacked single-crystal Si) SRAM technology with 25F², 0.16um² cell and doubly stacked SSTFT cell transistors for ultra high density and high speed applications", Digest of Technical Papers of Symposium on VLSI Technology pp220 (2005).
- [Jung04]: S-M Jung, H. Lim, W. Cho, H. Cho, C. Yeo, Y. Kang, D. Bae, J. Na, K. Kwak, B. Choi, S. Kim, J. Jeong, Y. Chang, J. Jang, J. Kim, K. Kim, B-I. Ryu; "Highly area efficient and cost effective double stacked S³ (stacked single-crystal Si) peripheral CMOS SSTFT and SRAM cell technology for 512M bit density SRAM"; Technical Digest of International Electron Devices Meeting (IEDM), pp265 (2004).
- [Jung06]: S-M. Jung, J. Jang, W. Cho, H. Cho, J. Jeong, Y. Chang, J. Kim, Y. Rah, Y. Son, J. Park, M-S. Song, K-H. Kim, J-S. Lim, K. Kim; "Three Dimensionally Stacked NAND Flash Memory Technology Using Stacking Single Crystal Si Layers on ILD and TANOS Structure for Beyond 30nm Node"; Technical Digest of International Electron Devices Meeting (IEDM), pp1 (2006).

K

- [Kamata09]: Y. Kamata, A. Takashima, Y. Kamimuta, T. Tezuka, "New Approach to Form EOT-Scalable Gate Stack with Strontium Germanide Interlayer for High-k/Ge MISFETs", VLSI pp79 (2009)
- [Kang04]: Y.H. Kang, S-M. Jung, J-H. Jang, J-H. Moon, W-S. Cho, C-D. Yeo, K-H. Kwak, B-H. Choi, B. J. Hwang, W-R. Jung, S-J. Kim, J-H. Kim, J-H. Na, H. Lim, J-H. Jeong, K. Kim; "Fabrication and characteristics of novel load PMOS SSTFT (Stacked Single-crystal Thin Film Transistor) for 3-Dimensional SRAM memory cell"; Proceedings of IEEE International SOI Conference, pp 127 (2004)
- [Kawamura83]: S. Kawamura, N. Sasaki, T. Iwai, M. Nakano, M. Takagi, « Three dimensional CMOS IC's Fabricated by using beam recrystallization », IEEE Electron Device Letters, VOL.4, NO.10 (1983)
- [Kawamura84]: S. Kawamura, N. Sasaki, T. Iwai, R. Mukai, M. Nakano, M. Takagi; "3-dimensional gate array with vertically stacked dual SOI/ CMOS structures fabricated by beam-recrystallization"; Digest of Technical Papers of Symposium on VLSI Technology pp44 (1984)

- [Kawamura87]: S. Kawamura, N. Sasaki, S. Kawai, T. Shirato, N. Aneha, N. M. Nakano; "3-D high-voltage CMOS ICs by recrystallized SOI merged with bulk control-unit"; Technical Digest of International Electron Devices Meeting pp758 (1987)
- [Kim04] : S-D. Kim, J6B. Johnson, J. Yuan, J-C-S. Woo; "Optimization of Recessed and Elevated Silicide Source/Drain Contact Structure Using Physical Compact Resistance Modeling and Simulation in Ultra-Thin Body SOI MOSFETs"; International conference on SIMulation of Semiconductor Processes And Devices SISPAD (2004).
- [Koomen73]: J. Koomen, "Investigation of the MOST channel conductance in weakinversion," Solid State Electron., vol. 16, pp. 801–810, 1973.
- [Koyanagi01] M. Koyanagi, Y. Nakagawa, K.-W. Lee, T. Nakamura, Y. Yamada, K. Inamura, K.-T. Park, H. Kurino, "Neuromorphic vision chip fabricated using three-dimensional integration technology", Proceedings of the IEEE International Solid-State Circuit Conference (ISSCC), pp 270 (2001).
- [Kumar01]: M. Kumar, H. Liu, J-K-O. Sin., J. Wan, K-L. Wang; A 3-D BiCMOS technology using selective epitaxial growth (SEG) and lateral solid phase epitaxy (LSPE); Technical Digest of International Electron Devices Meeting (IEDM) pp729 (2001).
- [Kunio89]: T. Kunio, K. Oyama, Y. Hayashi, M. Morimoto, "Three dimensional ICs, having four stacked active device layers"; Technical Digest of International Electron Devices Meeting pp 838 (1989)
- L**
- [Lai 06 b]: E-K. Lai, H-T. Lue, Y-H. Hsiao, J-Y. Hsieh, S-C. Lee, C-P. Lu; S-Y. Wang, L-W. Yang, K-C. Chen, J. Gong, K-Y. Hsieh, J. Ku, R. Liu, C-Y Lu; "A Highly Stackable Thin-Film Transistor (TFT) NAND-Type Flash Memory"; Digest of Technical Papers of Symposium on VLSI Technology, pp 46 (2006)
- [Lai 06a]: E-K.Lai, H-T. Lue, Y-H. Hsiao, J-Y. Hsieh, C-P. Lu; S-Y. Wang, L-W. Yang, T. Yang, K-C. Chen, J. Gong, K-Y. Hsieh, R. Liu, C-Y Lu; "A Multi-Layer Stackable Thin-Film Transistor (TFT) NAND-Type Flash Memory"; Technical Digest of International Electron Devices Meeting (IEDM) pp1 (2006)
- [Lau82]: C-K. Lau, Y-C. See, D-B. Scott, J-M. Bridges, S-M. Perna, R-A. Davies, "Titanium disilicide self aligned source/ drain +gate technology" IEEE International Electron Device Meeting (IEDM), Volume: 28, pp 714 (1982).
- [LeRoyer07] : C. Le Royer, L. Clavelier, C. Tabone, C. Deguet, L. Sanchez, J.-M. Hartmann, M.-C. Roure, H. Grampeix, S. Deleonibus, "0.12 μ m p-MOSFETs with High-K and Metal Gate Fabricated in a Si Process Line on 200mm GeOI Wafers", ESSDERC 2007, p.458-461
- [Lee03]: M-L. Lee, E-A. Fitzgerald; "Optimized Strained Si / Strained Ge Dual-Channel Heterostructures for High Mobility P- and N-MOSFETs"; IEEE Inetrnational Electron Device Meeting (IEDM), pp429 (2003)
- [Leyris 06] C. Leyris, F. Martinez, M. Valenza, A. Hoffmann, J. Vildeuil, and F. Roy, "Impact of random telegraph signal in cmos image sensors for lowlight levels," Proceedings of the 32nd European Solid-State Circuits Conference (ESSCIRC) pp. 376 (2006)

- [Lim 05]: H. Lim, S-M. Jung, Y. Rah, T. Ha, H. Park, C. Chang, W. Cho, J. Park, B. Son, J. Jeong, H. Cho, B. Choi, K. Kim; "65nm high performance SRAM technology with 25F2 0.16 μm^2 S³ (stacked single-crystal Si) SRAM cell, and stacked peripheral SSTFT for ultra high density and high speed applications", Proceedings of Solid-State Device Research conference (ESSDERC), pp 549 (2005).
- [Lin06] : X. Lin, S. Zhang, X. Wu, M. Chan; "Local clustering 3-D stacked CMOS technology for interconnect loading reduction"; IEEE Transactions on Electron Devices, Volume 53, Issue 6, pp 1405(2006).
- [Lin07]: Mingjie Lin, A. El Gamal, Y-C. Lu, S. Wong; "Performance Benefits of Monolithically Stacked 3-D FPGA"; IEEE Transactions on computer aided design of integrated circuits and systems, Vol 26, no2, pp216 (2007)
- [Lindsay04]: R. Lindsay, K. Henson, W. Vandervorst, K. Maex, B. J. Pawlak, R. Duffy, R. Surdeanu, P. Stolk, J. A. Kittl, S. Giangrandi, X. Pages, K. van der Jeugd; "Leakage optimization of ultra-shallow junctions formed by solid phase epitaxial regrowth"; Journal of. Vacuum Science and Technology B Volume 22, Issue 1, pp306 (2004)
- [Lin08]: . Lin, J. Luo, Y. Ma; "A Low-Power Monolithically Stacked 3D-TCAM Department of Electrical Engineering"; T IEEE International Symposium on Circuit and Systems (ISCAS) , pp3318 (2008)
- [Liu01]: H.Liu, M.Kumar, J-K-O.Sin, "A novel BiCMOS technology using selective epitaxy growth and lateral solid phase epitaxial", IEEE International Electron Device Meeting (IEDM), pp 729 (2001).
- [Liu02]: H. Liu, M. Kumar, J-K-O Sin; "A novel 3-D BiCMOS technology using selective epitaxy growth (SEG) and lateral solid phase epitaxial (LSPE)"; IEEE Electron Device Letter, Volume 23, Issue 3, pp 151 – 153 (2002)
- [Liu02b]: C. Liu, S. Tiwari, "Appliation of 3D CMOS technology to SRAMs", IEEE International SOI conference, pp68 (2002)
- [Liu06]: C-C. Liu, S. Tiwari; "Performance Advantages of 3-D Digital Integrated Circuits in a Mixed SOI and Bulk CMOS Design Space"; IEEE Transactions on Circuits and Systems, Vol56, No3, pp 207 (2006)
- [Lolivier05] : J. Lolivier, Manuscrit de thèse « Etude, réalisation et caractérisation de transistors silicium sur isolant complètement désertés de longueur de grille inférieure à 25nm » soutenue en 2005, ecole doctorale EEATS.
- [Loi06] : G-L. Loi, B. Agrawal, N. Srivastava, S-C. Lin, T. Sherwood, K. Banerjee; A Thermally-Aware Performance Analysis of Vertically Integrated (3-D) Processor-Memory Hierarchy; Proceedings of the 43rd annual Design Automation Conference, Session 56, pp991 (2006)
- [Luo04]: Z. Luo, A. Steegen, M. Eller, R. Mann, C. Baiocco, P. Nguyen, L. Lim, M. Hoinkis, V. Ku, V. Klee, V. Jamin, F. Wrschka, P. Shafer, P. Lin, W. Fang, S. Ajmera, A. Tan, W. Park, D. Mo, J. Lian, D. Vietzke, C. Coppock, A. Vayshenker, T. Hook, V. Chan, K. Kim, A. Cowley, S. Kim, E. Kaltalioglu, B. Zhang, S. Marokkey, Y. Lin, K. Lee, H. Zhu, M. Weybright, R. Rengarajan, J. Ku, T. Schiml, J. Sudijono, I. Yang, C. Wann; "High performance and low power transistors integrated in 65nm bulk CMOS technology"; IEEE International Electron Devices Meeting Technical Digest (IEDM) , pp 661 (2004)

M

- [Mangelinck99] : D. Mangelinck, J. Y. Dai, J. S. Pan, S. K. Lahiri, “ Enhancement of thermal stability of NiSi films on (100)Si and (111)Si by Pt addition, Appl. Phys. Lett. VOL 75, NO12, pp1736 (1999)
- [Maszara88]: W. P. Maszara, G. Goetz, A. Caviglia, J. B. McKitterick, « Bonding of silicon wafers for silicon on insulator, , Journal of Applied Physics. 64(10), pp4943 (1988)
- [Meindl01] : J.D. Meindl, R.Venkatesan, J-A.Davis, J. Joyner, A.Naeemi, P. Zarkesh-Ha, M. Bakir, T. Mule, P-A. Kohl, K-P. Martin ; “Interconnecting device opportunities for gigascale integration”; Technical Digest of International Electron Devices Meeting (IEDM) pp525 (2001).
- [Meindl03] : J.D. Meindl; “Interconnect opportunities for gigascale integration”; IEEE Micro, volume 23 issue 3 pp28 (2003).
- [Mitard08] : J. Mitard, B. De Jaeger, F.E. Leys, G. Hellings, K. Martens, G. Eneman, D.P. Brunco, R. Loo1, J.C. Lin, D. Shamiryan, T. Vandeweyer, G. Winderickx, E. Vrancken, C.H.Yu, K. De Meyer1, M. Caymax, L. Pantisano, M. Meuris, M.M. Heyns. “Record ION/IOFF performance for 65nm Ge pMOSFET and novel Si passivation scheme for improved EOT scalability”, IEDM p873
- [Molas04]: G. Molas, D. Deleruyelle, B. De Salvo, G. Ghibaudo, M. Gely, S. Jacob, D. Lafond, S. Deleonibus, “Impact of few electron phenomena on floating-gate memory reliability”, IEEE International Electron device Meeting (IEDM), pp.877 (2004)
- [Mofrad09]: M-R-T. Mofrad, J. Derakhshandeh, R. Ishihara; “Stacking of Single-Grain Thin-Film Transistors” 14th International Workshop on Active-Matrix Flatpanel Displays and Devices, JAPANESE JOURNAL OF APPLIED PHYSICS, Volume: 48 Issue: 3 Article Number: 03B015 (2009)
- [Mofrad09]: MRT. Mofrad, R. Ishihara, J. Derakhshandeh; “Monolithic 3D Integration of Single-Grain Si TFTs” Symposium on Amorphous and Polycrystalline Thin-Film Silicon Science and Technology held at the 2008 MRS Spring Meeting, 2008, pp 483 (2008)
- [Moriceau01]: H. Moriceau, F. Fournel, O. Rayssac, A-M. Cartier, C. Morales, S. Pocas, M.Zussy, E. Jalaguier, B. Biasse, B. Bataillou, A-M. Papon, C. Lagahe, B. Aspar, C. Maleville, F. Letertre, B. Ghyselen, T. Barbe ; “Overview on some recent advances in wafer bonding technologies”; Electrochemical Society Proceedings, vol27 p1 (2001)
- [Mukhopadhyay08]: S. Mukhopadhyay, K. Kim, C-T Chuang, “Device Design and optimization metgology for leakage and variability reduction in sub-45-nm FD/SOI SRAM”, IEEE Transactions on Electron Devices, Volume: 55, Issue: 1,pp 152 (2008).

N

- [Natarajan08]: S. Natarajan, M. Armstrong, M. Bost, R. Brain, M. Brazier, C-H Chang, V. Chikarmane, M. Childs, H. Deshpande, K. Dev, G. Ding, T. Ghani, O. Golonzka, W. Han, J. He*, R. Heussner, R. James, I. Jin, C. Kenyon, S. Klopcic, S-H. Lee, M. Liu, S. Lodha, B. McFadden, A. Murthy, L. Neiberg, J. Neiryneck, P. Packan, S. Pae*, C. Parker, C. Pelto, L. Pipes, J. Sebastian, J. Seiple, B. Sell, S. Sivakumar, B. Song, K. Tone, T. Troeger, C. Weber**, M. Yang, A. Yeoh, K. Zhang; “32nm Logic Technology Featuring 2nd-Generation High-k + Metal-Gate Transistors, Enhanced Channel Strain and 0.171 μ m² SRAM Cell Size in a 291Mb Array”, IEEE International Electron Device Meeting (IEDM), pp1 (2008)
- [Nii06]: Nii, H.; Sanuki, T.; Okayama, Y.; Ota, K.; Iwamoto, T.; Fujimaki, T.; Kimura, T.; Watanabe, R.; Komoda, T.; Eiho, A.; Aikawa, K.; Yamaguchi, H.; Morimoto, R.; Ohshima, K.; Yokoyama, T.; Matsumoto, T.; Hachimine, K.; Sogo, Y.; Shino, S.; Kanai, S.; Yamazaki, T.; Takahashi, S.; Maeda, H.; Iwata, T.; Ohno, K.; Takegawa, Y.; Oishi, A.; Togo, M.; Fukasaku, K.; Takasu, Y.; Yamasaki, H.; Inokuma, H.; Matsuo, K.; Sato, T.; Nakazawa, M.; Katagiri, T.; Nakazawa, K.; Shinyama, T.; Tetsuka, T.; Fujita, S.; Kagawa, Y.; Nagaoka, K.; Muramatsu, S.; Iwasa, S.; Mimotogi, S.; Yoshida, K.; Sunouchi, K.; Iwai, M.; Saito, M.; Ikeda, M.; Enomoto, Y.; Naruse, H.; Imai, K.; Yamada, S.; Nagashima, N.; Kuwata, T.; Matsuoka, F; “A 45nm High Performance Bulk Logic Platform Technology (CMOS6) using Ultra High NA(1.07) Immersion Lithography with Hybrid Dual-Damascene Structure and Porous Low-k BEOL”; IEEE International Electron Devices Meeting Technical Digest (IEDM) , pp 1 (2006)
- [Nishimura87]: T. Nishimura, Y. Inoue, K. Sugahara, S. Kusunoki, T. Kumamoto, S. Nakagawa, M. Nakaya, Y. Horiba, Y. Akasaka; “Three dimensional IC for High Performance Image Signal Processor”; Technical Digest of International Electron Devices Meeting pp 112 (1987)

O

- [Ohtake86]: K. Ohtake, K. Shirakawa, M. Koba, K. Awane, Y. Ohta, D. Azuma, S. Miyata; “Triple layered SOI dynamic memory”, Technical Digest of International Electron Devices Meeting pp 148 (1986)
- [Oishi05]: Oishi, A., Fujii, O., Yokoyama, T., Ota, K., Sanuki, T., Inokuma, H., Eda, K., Idaka, T., Miyajima, H., Iwasa, S., Yamasaki, H., Oouchi, K., Matsuo, K., Nagano, H., Komoda, T., Okayama, Y., Matsumoto, T., Fukasaku, K., Shimizu, T., Miyano, K., Suzuki, T., Yahashi, K., Horiuchi, A., Takegawa, Y., Saki, K., Mori, S., Ohno, K., Mizushima, L., Saito, M., Iwai, M., Yamada, S., Nagashima, N., Matsuoka, F, “High performance CMOSFET technology for 45nm generation and scalability of stress-induced mobility enhancement technique”: IEEE International Electron Device Meeting (IEDM), pp 229 (2005)

P

- [Packan08]: P. Packan, S. Cea, H. Deshpande, T. Ghani, M. Giles, O. Golonzka, M. Hattendorf, R. Kotlyar, K. Kuhn, A. Murthy, P. Ranade, L. Shinfren, C. Weber and K. Zawadzki, “High Performance Hi-K+ Metal Gate Strain Enhanced Transistors on (110) Silicon”, IEEE International Electron Device Meeting (IEDM 2008)
- [Pawlak04]: B. J. Pawlak, R. Lindsay, R. Surdeanu, B. Dieu, L. Geenen, I. Hoflijk, O. Richard, R. Duffy, T. Clarysse, B. Brijs, W. Vandervorst, C. J. J. Dachs; “Chemical and electrical dopants profile evolution during solid phase epitaxial regrowth”; Journal of. Vacuum Science and Technology B Volume 22, Issue 1, pp297 (2004)
- [Pouydebasque05]: A. Pouydebasque, B. Dumont, R. El-Farhane, A. Halimaouit, C. Laviron, D. Lenoble, C. Rossato, N. Casanova, V. Carron, S. Pokrant, T. Skotnicki, “CMOS integration of solid phase epitaxy for sub-50 nm devices”, IEEE European Solid State Device Research Conference, pp 419 (2005)

R

- [Rafhay08]: Q. Rafhay, “Modelling of nano nMOSFET with alternative channel materials in the fully and quasi ballistic regimes”, Manuscrit de thèse, école doctorale EEATS (2008)
- [Rahman00]: A. Rahman, R. Reif; “System level performance evaluation of three dimensional integrated circuits; IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol8, pp671 (2000)
- [Rajendran 07]: B. Rajendran, R-S. Shenoy, D-J. Witte, N-S. Chokshi, R-L. De Leon, G-S. Tompa, R. Fabian; “Low Thermal Budget Processing for Sequential 3-D IC Fabrication”; IEEE Transactions on Electron Devices, Volume: 54, Issue: 4, pp 707 (2007)
- [Romanjek08]: K. Romanjek, L. Hutin, C. Le Royer, A. Pouydebasque, M.-A. Jaud, C. Tabone, E. Augendre, L. Sanchez, J.-M. Hartmann, H. Grampeix, V. Mazzocchi, S. Soliveres, R. Truche, L. Clavelier, P. Scheiblin, X. Garros, G. Reibold, M. Vinet, F. Boulanger, S. Deleonibus, “High performance 70nm Gate Length Germanium-On-Insulator pMOSFETs With High- κ /Metal Gate”; ESSDERC, pp75 (2008)
- [Roos91]: G. Roos, B. Hofflinger, M. Schubert, R. Zingg “High-Density ASICs with three dimensional CMOS process” VLSI –TSA pp297 (1991)

[Roos92]: G. roos, B. Hoefflinger, "Complex 3D CMOS Circuits based on Triple-Decker cell", IIE Journal of Solid-State Circuits, VOL 27, NO 7 (1992)

[Roos93]: G. Roos, B. Hoefflinger, "Three-Dimensional CMOS NAND with three Stacked Channels" Electronics letters VOL.29 NO. 24 (1993)

S

[Sellier08] : M. Sellier, J-M. Portal, B. Borot, S. Colquhoun, R.Ferrant, F. Boeuf, A. Farcy, "Predictive Delay Evaluation on Emerging CMOS Technologies: A Simulation Framework", 9th International Symposium on Quality Electronic Design (ISQED), pp492 (2008)

[Shah 84]: A-H. Shah, L-R. Hite, S-S. M. Shetti, P-K. Chatterjee, H-E. Davis, R-K. Hester, S-D-S. Malhi, R. Karnaugh, C-D. Gosmeyer, R-S. Sundaresan, C-E. Chen, H-W. Lam, R-A. Haken; "A 2 μ m Stacked CMOS 64K SRAM"; Digest of Technical Papers of Symposium on VLSI Technology pp8 (1984)

[Sheraw05]: C. D. Sheraw, M. Yang, D. M. Fried, G. Costrini, T. Kanarsky, W-H. Lee, V. Chan, M. V. Fischetti, J. Holt, L. Black, M. Naeem, S. Panda, L. Economikos, J. Groschopf, A. Kapur, Y. Li, R. T. Mo, A. Bonnoit, D. Degraw, S. Luning, D. Chidambarrao, X. Wang, A. Bryant, D. Brown, C-Y. Sung, P. Agnello, M. Jeong, S-F. Huang, X. Chen, M. Khare, "Dual Stress Liner Enhancement in Hybrid Orientation Technology", Digest of Technical Papers Symposium on VLSI Technology, pp12 (2005)

[Skotnicki88] : T.Skotnicki, G. Merckel, T.Pedron ; "The Voltage-Doping Transformation: A New Approach to the Modeling of MOSFET Short-Channel Effects"; Electron Device Letter, vol9, no3 (1988)

[Sohn06]: K. Sohn, H-S. Mo, Y-H. Suh, H-G. Byun, H-J. Yoo; "An Autonomous SRAM With On-Chip Sensors in an 80-nm Double Stacked Cell Technology"; IEEE journal od Solid-State Circuit, vol 41, Issue 4, pp823 (2006)

[Sohn08]: K. Sohn, Y-H. Suh, Y-J. Son, D-S. Yim, K-Y. Kim, D-G. Bae, T. Kang, H. Lim, S-M. Jung, H-G. Byun, Y-H. Jun, K. Kim; "A 100nm Double-Stacked 500MHz 72Mb Separate-I/O Synchronous SRAM with Automatic Cell-Bias Scheme and Adaptive Block Redundancy"; Digest of Technical Papers of the IEEE International Solid-State Circuits Conference (ISSCC), pp386 (2008).

[Son07]: Y-H. Son, J-W. Lee, P. Kang, M-G. Kang, J- B. Kim, S- H. Lee, Y-P. Kim, I- S. Jung, B- C. Lee, S-Y. Choi, U-I. Chung, J-T. Moon, B-I. Ryu; "Laser-induced Epitaxial Growth (LEG) Technology for High Density 3-D Stacked Memory with High Productivity", Digest of Technical Papers of Symposium on VLSI Technology pp80 (2007).

[Jung07]: S-M. Jung, H. Lim, C. Yeo, K. Kwak, B. Son, H. Park, J. Na, J-J. Shim, C. Hon, K. Kim; "High Speed and Highly Cost effective 72M bit density S³ SRAM Technology with Doubly Stacked Si Layers, Peripheral only CoSix layers and Tungsten Shunt W/L Scheme for Standalone and Embedded Memory", Digest of Technical Papers of Symposium on VLSI Technology pp 82 (2007).

- [Steen07]: S. Steen, D. LaTulipe, A-W. Topol, D-J. Frank, K. Belote, D. Posillico; "Overlay as the key to drive wafer scale 3D integration"; *Microelectronics Engineering* 84 p1412 (2007).
- [Strane07]: J. Strane, D. Brown, C. Lavoie, J. Suenaga, B. Haran, P. Press, P. Besser, P. Flaitz, M. Gribelyuk, T. Kammler, I. Peidous, H. Chen, S. Waidmann, A. Frye, P. DeHaven, A. Domenicucci, C. Murray, R. Knarr, H.J. Engelmann, C. Streck, V. Kahlert, S. Deshpande, E. Leobandung, J. Pellerin, J. Jagannathan; "Implementation of Robust Nickel Alloy Salicide Process for High-Performance 65nm SOI CMOS Manufacturing"; *Technical digest of the VLSI Technologie System and Application (VLSI-TSA)*, p1 (2007)
- [Subramanian98]: V. Subramanian, K-C. Saraswat; "High-performance germanium-seeded laterally crystallized TFTs for vertical device integration"; *IEEE Transactions on Electron Devices*, Volume 45, Issue 9, pp 1934 (1998)
- [Subramanian99]: V. Subramanian, M. Toita, N-R. Ibrahim, S-J. Souri, K-C. Saraswat; "Low-leakage germanium-seeded laterally-crystallized single-grain 100-nm TFTs for vertical integration applications"; *IEEE Electron Device Letters*, Volume 20, Issue 7, pp341 (1999).
- [Suguro04] : K. Suguro, T. Ito, K. Matsuo, T. Iinuma, K-T. Nishinohara, "Overview of the Prospects of Ultra-Rapid Thermal Process for Advanced CMOSFETs", *International Workshop on Junction Technology, IWJT* pp 18 (2004)
- [Sung05]: C-Y. Sung, H. Yin, H-Y. Ng, K-L. Saenger, V. Chan, S-W. Crowder, J. Li, J-A. Ott, R. Bendernagel, J-J. Kempisty, V. Ku, H-K. Lee, Z. Luo, A. Madan, R-T. Mo, P-Y. Nguyen, G. Pfeiffer, M. Raccioppo, N. Rovedo, D. Sadana, J-P. de Souza, R. Zhang, Z. Ren, C-H. Wann; "High Performance CMOS Bulk Technology Using Direct Silicon Bond (DSB) Mixed"; *IEEE International Electron Device Meeting (IEDM)* pp225 (2005)
- [Sung08]: Chun-Yung Sung, Haizhou Yin, Hung Y. Ng, Katherine L. Saenger, Victor Chan, Scott W. Crowder, Jinghong Li, John A. Ott, Robert Bendernagel, Jeremy J. Kempisty, Victor Ku, H.K. Lee, Zhijiong Luo, Anita Madan, Renee T. Mo, Phung Y. Nguyen, Gerd Pfeiffer, Michael Raccioppo, Nivo Rovedo, Devendra Sadana, Joel P. de Souza, Rong Zhang, Zhibin Ren and Clement H. Wann, "High Performance CMOS Bulk Technology Using Direct Silicon Bond (DSB) Mixed Crystal Orientation Substrates", *IEEE International Electron Device Meeting (IEDM)*, pp225 (2008)
- [Suntharalingam05]: V. Suntharalingam, R. Berger, J.A. Burns, C.K. Chen, C.L. Keast, J.M. Knecht, R.D. Lambert, K.L. Newcomb, D.M. O'Mara, D.D. Rathman, D.C. Shaver, A.M. Soares, C.N. Stevenson, B.M. Tyrrell, K. Warner, B.D. Wheeler, D.-R.W. Yost, D.J. Young, "Megapixel CMOS image sensor fabricated in three-dimensional integrated circuit technology", *Proceedings of the IEEE International Solid-State Circuit Conference (ISSCC)*, pp356 (2005)
- [Sodini82]: C. Sodini, T. Ekstedt, J. Moll, "Charge accumulation and mobility in thin dielectric MOS transistors," *Solid State Electron.*, vol. 25, pp.833, 1982.

T

- [Takao91]: Y. Takao, H. Shimada, N. Suzuki, Y. Matsukawa, Y. Kobayashi, N. Sasaki; "A low power SRAM utilizing high ION-I/OFF ratio laser recrystallized SOI pMOSFETs load, Digest of Technical Papers of Symposium on VLSI Technology pp95 (1991)

- [Takao92] : Y. takao, H. Shimada, N. Suzuki, Y. Matsukawa, N. Sasaki, "Low-power and High Stability SRAM technology Using a Laser-Recrystallized p-channel SOI MOSFET", IEEE Transactions on Electron Devices, VOI 39, NO 9, (1992)
- [Takeda00] : K. Takeda, Y. Aimoto, N.Nakamura, H. Toyoshima, T. Iwasaki, K. Noda, K. Matsui, S. Itoh, S. Masuoka, T. Horiuchi, A. Nakagawa, K. Shimogawa, H. Takahashi, "A 16-Mb 400-MHz Loadless CMOS Four-Transistor SRAM Macro" IEEE journal of Solid-State Circuits , Volume: 35, Issue: 11, pp 1631 (2000)
- [Tan08] : C-S. Tan, R-J. Gutmann, L. Rafael Reif, "Wafer Level 3-D ICs process technology", Springer, 2008
- [Tateshita06] : Y. Tateshita, J. Wang, K. Nagano, T. Hirano, Y. Miyanami, T. Ikuta, T. Kataoka, Y. Kikuchi, S. Yamaguchi, T. Ando, K. Tai, R. Matsumoto, S. Fujita, C. Yamane, R. Yamamoto, S. Kanda, K. Kugimiya, T. Kimura, T. Ohchi, Y. Yamamoto, Y. Nagahama, Y. Hagimoto, H. Wakabayashi, Y. Tagawa, M. Tsukamoto, H. Iwamoto, M. Saito, S. Kadomura and N. Nagashima, "High-Performance and Low-Power CMOS Device Technologies Featuring Metal/High-k Gate Stacks with Uniaxial Strained Silicon Channels on (100) and (110) Substrates", IEEE International Electron Device Meeting (IEDM), (2006)
- [Taur92]: Y. Taur, D-S. Zicherman, D-R. Lombardi, P-J. Restle, C-H. Hsu, H-I. Nanafi, M-R. Wordeman, B. Davari, G-G. Shahidi.; "New 'shift and ratio' method for MOSFET channel-length extraction"; IEEE International Electron Device Letters, vol. 13, issue 5, pp. 267 (1992)
- [Thomas03] : O. Thomas, A. Amara "An SOI 4 Transistors Self-Refresh Ultra-Low Voltage memory cell", Proceedings of the 2003 International Symposium on Circuits and Systems, Volume:5, pp401 (2003)
- [Thomas09]: O. Thomas, M. Vinet, O. Rozeau, P. Batude, A. Valentian ; "Compact 6T SRAM cell with robust Read/Write stabilizing design in 45nm Monolithic 3D IC technology"; soumis à IEEE International Conference on Integrated Circuit Design and Technology and Tutorial (2009)
- [Tiwari02]: S. Tiwari, H-S. Kim, S. Kim, A. Kumar, C-C. Liu, L. Xue; "Three-dimensional integration in silicon electronics"; Proceedings IEEE Lester Eastman Conference on High Performance Devices, pp24 (2002)
- [Topol05]: A-W. Topol, D-C. La Tulipe, L. Shi, S-M. Alam, D-J. Frank, S-E. Steen, J. Vichiconti, D. Posillico, M. Cobb, S. Medd, J. Patel, S. Goma, D. DiMilia, M. T. Robson, E. Duch, M. Farinelli, C. Wang, R. A. Conti, D. M. Canaperi, L. Deligianni, A. Kumar, K-T. Kwietniak, C. D'Emic, J. Ott, A. M. Young , K. W. Guarini, and M. Jeong; "Enabling SOI-Based Assembly Technology for Three-Dimensional (3D) Integrated Circuits (ICs)"; Technical Digest of International Electron Devices Meeting (IEDM) pp 352 (2005).
- [Topol06] : A-W. Topol ,D.-C. La Tulipe, L. Shi, D-J. Frank, K. Bernstein, S-E. Steen, A. Kumar, G-U. Singco, A-M. Young, K-W. Guarini, M. Jeong; "Three-dimensional integrated circuits"; Source IBM Journal of Research and Development archive Volume 50 , Issue 4/5 (July 2006)

[Toshiba04]: System-in-Package, Toshiba system catalog (2004), disponible en ligne/: <http://www.semicon.toshiba.co.jp/eng/prd/common/pdf/sce0010a.pdf>

[Tyagi05]: S. Tyagi, C. Auth, P. Bai, G. Curello, H. Deshpande, S. Gannavaram, O. Golonzka, R. Heussner, R. James, C. Kenyon, S-H. Lee, N. Lindert, M. Liu, R. Nagisetty, S. Natarajan, C. Parker, J. Sebastian, B. Sell, S. Sivakumar, A. St Amour, K. Tone; "Advanced low power, high performance, strained channel 65nm technology" IEEE International Electron Devices Meeting Technical Digest (IEDM) , pp 245 (2005)

U

[Uemoto90]: Y. uemoto, R. Fujii, A. Nakamura, K. Senda, "A High performance Stacked-CMOS SRAM Cell by solid Phase Growth technique", Digest of Technical Papers of Symposium on VLSI Technology, pp 21 (1990)

V

[Villanueva03]: D. Villanueva, A. Pouydebasque, E. Robilliart, T. Skotnieki, E. Fuchs, H. Jaouen ; « Impact of the lateral Source / Drain Abruptness on MOSFET Characteristics and Transport Properties»; IEEE International Electron Device Meeting (IEDM) p237 (2003)

W

[Whan04] S. J. Whang, S. J. Lee, F. Gao, N. Wu, C. X. Zhu, J. S. Pan, L. J. Tang, and D. L. Kwong, "Germanium p- & n-MOSFETs fabricated with novel surface passivation (plasma-PH3 and thin AlN) and TaN/HfO2 Gate stack," Ieee International Electron Devices Meeting 2004, Technical Digest, pp. 307-310, 2004.

[Wang00a]: H. Wang, M. Chan, S. Jagar, Y. Wang, K-K. Ping; "Submicron super TFTs for 3-D VLSI applications"; IEEE Electron Device Letters, Volume 21, Issue 9, pp439 (2000)

- [Wang00b]: H. Wang, M. Chan, S. Jagar, V-M-C. Poon, M. Qin, Y. Wang, K-K Ping; "Super thin-film transistor with SOI CMOS performance formed by a novel grain enhancement method"; IEEE Transactions on Electron Devices, Volume 47, Issue 8, pp 1580 (2000)
- [Wang06]: Howard C.-H. Wang, Shih-Hian Huang, Ching-Wei Tsai, Hsien-Hsin Lin, Tze-Liang Lee, Shih-Chang Chen, Carlos H. Diaz, Mong-Song Liang and Jack Y.-C. Sun, "High-Performance PMOS Devices on (110)/<111'> Substrate/Channel with Multiple Stressors", IEEE International Electron Device Meeting (IEDM), pp1 (2006)
- [Weber05]: O. Weber, Y. Bogumilowicz, T. Ernst, J.-M. Hartmann, F. Ducroquet, F. Andrieu, C. Dupré, L. Clavelier, C. Le Royer, N. Cherkashin, M. Hytch, D. Rouchon, H. Dansas, A.-M. Papon, V. Carron, C. Tabone, and S. Deleonibus; "Strained Si and Ge MOSFETs with High-K/Metal Gate Stack for High Mobility Dual Channel CMOS"; IEEE International Electron Device Meeting (IEDM), pp 137(2005)
- [Weber07] : O. Weber, T. Irisawa, T. Numata, M. Harada, N. Taoka, Y. Yamashita, T. Yamamoto, N. Sugiyama, M. Takenaka and S. Takagi; "Examination of Additive Mobility Enhancements for Uniaxial Stress Combined with Biaxially Strained Si, Biaxially Strained SiGe and Ge Channel MOSFETs"; IEEE International Electron Device Meeting (IEDM), p719 (2007)
- [Weber08]: O. Weber, O. Faynot, F. Andrieu, C. Buj-Dufournet, F. Allain, P. Scheiblin, J. Foucher, N. Daval, D. Lafond, L. Tosti, L. Brevard, O. Rozeau, C. Fenouillet-Beranger, M. Marin, F. Boeuf ,D. Delprat, K. Bourdelle, B.-Y. Nguyen, S. Deleonibus; "High Immunity to Threshold Voltage Variability in Undoped Ultra-Thin FDSOI MOSFETs and its Physical Understanding"; IEEE International Electron Device Meeting, IEDM technical digest, (2008)
- [Wei02] : L. Wei, , R. Zhang, K. Roy, ,Z.Chen, D-B. Janes; "Vertically Integrated SOI Circuits for Low-Power and High-Performance Applications"; IEEE Transactions on Very Large Scale Integration (VLSI) Systems, VOL10, NO3 (2002)
- [Widiez05]: Julie Widiez, « Etude, fabrication et caractérisation de transistor CMOS double grille planaires déca-nanométriques » Manuscrit de thèse école doctorale EEATS Institut National Polytechnique de grenoble (2005)
- [Wong 07]: S. Wong, A. El-Gamal, P. Griffin, Y. Nishi, F. Pease, J. Plummer; "Monolithic 3D Integrated Circuits"; International Symposium on Technology, Systems and Applications, (VLSI-TSA), p1 (2007)
- [Woodard06]: E-M. Woodard, R-G. Manley, G. Fenger, R-L. Saxer, K-D. Hirschman, D. Dawson-Elli, J. G. Couillard; "Low Temperature Dopant Activation for Integrated Electronics Applications"; IEEE 16th Biennial University/Government/Industry Microelectronics Symposium, pp161 (2006)
- [WuEDL07] N. Wu, Q. C. Zhang, D. S. H. Chan, N. Balasubramanian, and C. X. Zhu, "Gate-first germanium nMOSFET with CVD HfO₂ gate dielectric and silicon surface passivation," Ieee Electron Device Letters, vol. 27, pp. 479-481, 2006.
- [Wu05]: X. Wu, P-C-H. Chan, S. Zhang, C. Feng, M. Chan; "A three-dimensional stacked fin-CMOS technology for high-density ULSI circuits"; IEEE Transactions on Electron Devices, volume: 52, Issue: 9,pp 1998(2005)
- [Wu08] : Shien-Yang Wu, C.W. Chou, C.Y. Lin, M.C. Chiang, C.K. Yang, M.Y. Liu, L.C. Hu, C.H. Chang, P.H. Wu, C.I. Lin, H.F. Chen, S.Y. Chang, S.H. Wang, P.Y. Tong, Y.L. Hsieh, P.Y. Tong, J.J. Liaw, K.H. Pan, C.H. Hsieh, C.H. Chen, J.Y. Cheng, C.H. Yao, W.K. Wan, T.L. Lee, K.T. Huang, C.C. Chen, K.C. Lin, L.Y. Yeh, K.C. Ku, S.C. Chen, C.W. Chang, H.J. Lin, S.M. Jang, Y.C. Lu, J.H. Shieh, M.H. Tsai, J.Y. Song, K.S. Chen, V. Chang, S.M. Cheng, S.H. Yang, C.H. Diaz, Y.C. See, M.S. Liang; "32nm CMOS Low Power SoC Platform Technology for Foundry Applications with Functional High Density SRAM": IEEE International Electron Device Meeting (IEDM), pp 263 (2008)

Y

- [Yako08]: K.Yako, K. Uejima, T.Yamamoto, A. Mineji, T.Nagumo, T. Ikezawa, N. Matsuzaka, S. Shishiguchi, T. Hase, M. Hane; "Aggressive Design of Millisecond Annealing Junctions for Near-Scaling-Limit Bulk CMOS using Raised Source/Drain Extensions"; IEEE International Electron Device Meeting (IEDM), pp1 (2008)
- [Yamazaki 86]: K. Yamazaki, M. Yoneda, S. Ogawa, M. Ueda, S. Akiyama, Y Terui, "Fabrication Technologies for dual 4Kbit Stacked SRAM"; Technical Digest of International Electron Devices Meeting pp436 (1986)
- [Yamazaki90]: K. Yamasaki, Y.Itoh, A. wada, K. Morimoto, Y.Tomita; "4-Layer 3D IC Technologies for Parallel Signal Processing", Technical Digest of International Electron Devices Meeting pp599 (1990)
- [Yang04]: M. Yang, V. Chan, S. H. Ku, M. Jeong, L. Shi, K. K. Chan, C. S. Murthy, R. T. MO, H. S. Yang, E. A. Lehnef, Y. Surprid, F. F. Jamin, P. Oldiges, Y. Zhang, B. N. To, J. R. Holt, S. E. Steen, M. P. Chudzik, D. M. Fried, K. Bemstein, H. Zhu, C. Y. Sung, 1. A. OR, D. C. Boyd, andN. Rovedo, "On the Integration of CMOS with Hybrid Crystal Orientations", Digest of Technical Papers Symposium on VLSI Technology, pp160 (2004)
- [Yang06a]: Min Yang, Victor W. C. Chan, Kevin K. Chan, Leathen Shi, David M. Fried, James H. Stathis, Anthony I. Chou, Evgeni Gusev, John A. Ott, Lindsay E. Burns, Massimo V. Fischetti, and Meikei Jeong, "Hybrid-Orientation Technology (HOT): Opportunities and Challenges", IEEE Transactions on Electron Devices (TED), Volume: 53, Issue: 5, pp 965 (2006)
- [Yang06b]: M. Yang, K. Chan, A. Kumar, S.-H. Lo, J. Sleight, L. Chang, R. Rao, S. Bedell, A. Ray, J. Ott, J. Patel, C. D'Emic, J. Rubino, Y. Zhang, L. Shi, S. Steen, E. Sikorski, J. Newbury, R. Meyer, B. To, P. Kozlowski, W. Graham, S. Maurer, S. Medd, D. Canaperi, L. Deligianni, J. Tornello, G. Gibson, T. Dalton, M. Jeong, and G. Shahidi, "Silicon-on-Insulator MOSFETs with Hybrid Crystal Orientations", : Digest of Technical Papers Symposium on VLSI Technology, (2006)
- [Yang07]: B. Yang, K. Nummy, A. Waite, L. Black, H. Gossmann, H. Yin, Y. Liu, B. Kim, S. Narashimha, P. Fisher, H-V. Meer, J. Johnson, D. Chidambarrao, S-D. Kim, C. Sheraw, D. Wehella-gamage, J. Holt, X. Chen, D. Park, C-Y. Sung, D. Schepis, M. Khare, S. Luning, P. Agnello, "Stress dependence and poly-pitch scaling characteristics of (110) PMOS drive current", Symposium on Very Large Scale integration (VLSI) p126 (2007)
- [Yeo05] C. C. Yeo, B. J. Cho, E. Gao, S. J. Lee, A. H. Lee, C. Y. Yu, C. W. Liu, L. J. Tang, and T. W. Lee, "Electron mobility enhancement using ultrathin pure Ge on Si substrate," Ieee Electron Device Letters, vol. 26, pp. 761-763, 2005.
- [Yin06]: Haizhou Yin, C.Y. Sung, K.L. Saenger, M. Hamaguchi, R. Hasumi, K. Ohuchi, H. Ng, R. Zhang, K.J. Stein, T.A. Wallner, J. Li, J.A. Ott1, X. Chen, Z.J. Luo, N. Rovedo, K. Fogel, G. Pfeiffer, R. Kleinhenz, R. Bendernagel, D.K. Sadana1, M. Takayanagi, K. Ishimaru, S.W. Crowder, D. Park, M. Khare, and G. Shahidi, "Scalability of Direct Silicon Bonded (DSB) Technology for 32nm Node and Beyond", Digest of Technical Papers Symposium on VLSI Technology, pp222 (2006)

- [Yu 04]: D-S. Yu, A. Chin, C-C. Laio, C-F. Lee, C-F. Cheng, W-J. Chen, C. Zhu, M-F. Li, W-J. Yoo, S-P. McAlister, D-L. Kwong; "3D GOI CMOSFETs with novel IrO₂(Hf) dual gates and high-k dielectric on 1P6M-0.18 μm-CMOS" Technical Digest of International Electron Devices Meeting (IEDM) pp 181 (2004)
- [Yu04b]: D. S. Yu, Albert Chin, C. C. Laio, C. E. Lee, C. E. Cheng, W. J. Chen, C. Zhu, M.-F. Liz, W. J. Yoo, S. P. McAlister, D. L. Kwong; "3D GOI CMOSFETs with Novel IrO₂(Hf) Dual Gates and High-K Dielectric on 1P6M-0.18 μm-CMOS » ; IEEE International Electron Device Meeting (IEDM), pp181 (2004)
- [Yu05] : D-S. Yu, A. Chin, C-C. Liao, C-F. Lee, C-F. Cheng, M-F. Li, W-J. Yoo, S-P. McAlister; "Three-dimensional metal gate-high-κ-GOI CMOSFETs on 1-poly-6-metal 0.18-μm Si devices"; IEEE Electron Device Letters, Volume: 26, Issue: 2, pp118 (2005).

Z

- [Zhang99] : R. Zhang, K. Roy, D. B. Janes, "Architecture and Performance of 3-Dimensional SOI Circuits", IEEE International SOI Conference, pp44 (1999)
- [Zhang04]: S. Zhang, R. Han, X. Lin, X. Wu, M. Chan; "A stacked CMOS technology on SOI substrate"; IEEE Electron Device Letters, Volume 25, Issue 9, pp 661(2004)
- [Zingg89]: R-P. Zingg, B. Hofflinger, "Stacked CMOS Inverter with Symetric Device Performance, Technical Digest of International Electron Devices Meeting pp 989 (1989)
- [Zingg 90]: R-P. Zingg, J-A. Friedrich, G-W. Neudeck, B. Hofflinger, "Three-Dimensional Stacked MOS Transistors by localized Silicon Epitaxial Overgrowth", IEEE Transactions on Electron Devices, VOL 37. NO 6 (1990)