



HAL
open science

Fabrication de CMOS à basse température pour l'intégration 3D séquentielle

Cao-Minh Lu

► **To cite this version:**

Cao-Minh Lu. Fabrication de CMOS à basse température pour l'intégration 3D séquentielle. Micro et nanotechnologies/Microélectronique. Université Grenoble Alpes, 2017. Français. NNT: 2017GREAT109 . tel-01824340

HAL Id: tel-01824340

<https://theses.hal.science/tel-01824340>

Submitted on 27 Jun 2018

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Nano électronique et Nano technologie**

Arrêté ministériel : 25 mai 2016

Présentée par

Cao-Minh Vincent LU

Thèse dirigée par **Claire Fenouillet-Béranger, HDR, CEA-LETI,**
et
codirigée par **Thomas Skotnicki, HDR, STMicroelectronics**

préparée au sein du **CEA-LETI et STMicroelectronics**
dans **l'École Doctorale Electronique, Electrotechnique,
Automation et Traitement du Signal**

Fabrication de CMOS à basse température pour l'intégration 3D séquentielle

Thèse soutenue publiquement le **24 Octobre 2017,**
devant le jury composé de :

Mr Francis Calmon

Professeur, INSA de Lyon, Président

Mr Adrian IONESCU

Professeur, EPFL, Rapporteur

Mr Jean FOMPEYRINE

Chercheur, IBM Zürich, Rapporteur

Mr Jean-Pierre COLINGE

Directeur de recherche, CEA-LETI, Examineur

Mlle Perrine BATUDE

Ingénieur-Chercheur, CEA-LETI, Co-Encadrante

Mr Jean-Baptiste PIN

Ingénieur, Applied Materials, Invité



Sommaire

SOMMAIRE.....	3
CHAPITRE 1 – INTRODUCTION A LA 3D SEQUENTIELLE.....	7
1.1 Contexte et intérêt de la 3D Séquentielle	8
1.2 Problématiques de la 3D Séquentielle	15
1.3 Développement de briques technologiques à bas budget thermique	21
CHAPITRE 2 – MATERIAUX POUR LES ESPACEURS DE GRILLE.....	33
2.1 Étude des matériaux.....	36
2.1.1 Caractérisation des matériaux.....	36
2.1.2 Résistance aux chimies humides	42
2.2 Étude morphologique	44
2.1.3 Uniformité et conformité du dépôt.....	44
2.1.4 Gravure anisotropique des espaceurs	45
2.1.5 Reprise de l'épitaxie source et drain	48
2.3 Étude électrique	49
2.1.6 Intégration du SiOCN	49
2.1.7 Intégration du SiCO.....	53
CHAPITRE 3 – EPITAXIE A BASSE TEMPERATURE POUR LA SURELEVATION DES SOURCES ET DRAINS.....	57
3.1 Préparation de surface	59
3.1.1 Préparation de surface classique (HF-last+Recuit H ₂)	59
3.1.2 Nettoyage par plasma : Siconi	61
3.2 Croissance épitaxiale	65
3.1.3 Epitaxie sélective vs dépôt/gravure cyclée.....	65
3.1.4 Epitaxie Si intrinsèque	70
3.1.5 Epitaxie dopée <i>in-situ</i>	74
CHAPITRE 4 – PROPRIETES DE L'EMPILEMENT DE GRILLE.....	83
4.1 Procédés expérimentaux	89
4.2 Influence du budget thermique	92
4.3 Procédés leviers à bas budget thermique	97
4.1.1 Nitruration du HfO ₂	97
4.1.2 Densification de l'oxyde chimique.....	102
4.1.3 Passivation par deutérium	104
4.4 Recuits alternatifs.....	108
4.1.4 Recuit micro-ondes.....	108
4.1.5 Recuit laser	110
4.5 Pistes pour l'amélioration de la fiabilité NBTI	112
CHAPITRE 5 – CONCLUSIONS DU MANUSCRIT	119
REFERENCES.....	127
LISTE DE PUBLICATIONS.....	143

Remerciements

Pour commencer, j'aimerais remercier tout le jury, Francis Calmon, Jean Fompeyrine, Adrian Ionescu, Jean-Pierre Colinge et Jean-Baptiste Pin pour avoir eu le courage de lire tout ce manuscrit, pour leurs retours constructifs et surtout de ne pas m'avoir humilié devant amis et famille pendant la soutenance. Je remercie également Claire et Thomas de m'avoir donné l'opportunité d'effectuer cette thèse malgré des réponses douteuses sur ce qu'est un transistor lors de l'entretien, et donc d'avoir finalement pu (à peu près) comprendre comment un transistor MOS marche.

Ceci dit, j'aimerais ensuite remercier toute l'équipe CoolCube, à commencer par Claire pour son suivi et sa disponibilité tout du long – surtout lors des débarquements aléatoires mais réguliers dans son bureau pour des questions – ainsi qu pour son support et sa bonté. Je remercie également Perrine pour sa participation à mon régime en me faisant sauter un repas lors de la rédaction du dernier VLSI, bien que cela m'ait fait manquer mon grand chelem/cuillère en bois. Je remercie aussi à Laurent, dit « La Bible », pour avoir maintenu mon taux de caféine à flot ainsi que pour les nombreuses questions/réponses parfois métaphysiques. Merci également à Luca pour m'avoir appris l'importance de l'organisation, Jessy qu'on peut travailler en étant stressé de la vie, Fabien car, il faut pas déconner non plus, la zenitude, Alex pour ses interventions régulières et pertinentes en réunion d'équipe ainsi que Paul et Louis pour leur aide.

Je remercie aussi tout LICL : Maud grâce à qui j'ai pu rester dans cette merveilleuse ville de Grenoble, Marie-Pierre maître es clafoutis, Bernard maître es SEM, Claude maître es Siam, Laurent B. maître es SNOW1BIS, Nils et Valérie et leurs gestions de 193485 lots à Crolles, Louis pour ne pas nous avoir snobé aux US, Phuong, Laurent G., Hervé et Zdenek grâce à qui j'ai pu travailler sur le fameux projet III-V, Sylvain, Yves, Cyrille, François... pour leur bonne humeur (pour la plupart) de tous les jours et d'avoir fait vivre le labo. Merci également aux équipes DTSI : Virginie B., Céline, Chloé et Anne pour avoir testé à peu près tous les diélectriques basse température possibles, Jean-Michel pour la première épitaxie FAR du Leti, Philippe, Rémy pour les nombreuses discussions et théories – parfois farfelues, Virginie L., Pascal, Marine et Noélie pour les heures passées pour retirer 30nm de SiN, Sébastien et Pablo pour leur aide à la démonstration de la technique désormais connue de « crème brûlée », Véro pour les 14238 heures de recuit sur les lots MOCA, Frédérique, Benoit, Xavier et Mikael pour l'apprentissage de la combinaison « F9+PGDN », Charles prophète de la capa MOS, Alain, Giovanni, Fabienne et Rabah pour les 432183 puces testées, Christiane, André et Amy pour avoir montré qu'une P2 ne sert pas toujours, Carmelo, Christian A. pour avoir montré à peu près tous les profils d'espaceurs possibles, Christian V. « oui 20nm de CD c'est OK », Nicolas, Maxime et Julie & Lamine pour le karaoké mémorable, ainsi que Mélanie et Olivier grâce à qui j'ai pu rédiger un papier sans rien débiter au sujet. Je remercie aussi AMAT, merveilleuse entreprise, leader des équipementiers en semiconducteur et proposant des équipements à la pointe ainsi que mes désormais collègues David, J-B et Vincent, Nicolas de LAM et Karim et Fulvio d'Excico pour les nombreuses démos.

Je remercie ensuite tous les thésards et amis pas encore cités : Issam pour le piège à bulles à l'entrée du bureau, « miam-miam » Mathilde, « pas toujours poker face » Coco, Julien P., Mathias & Loïc à réunir pour un repas mouvementé, Anouar, Lina rebelle de Miribel, Giulia pour ne pas avoir volé mon déjeuner à la récré, Carlos, Julien D., José la gente esta muy loca, Julien B. pour le fast-food challenge rempli mais perdu, Benoit, Thilo, Julia, Adam... Désolé pour les personnes que j'aurais oublié lors de l'écriture de ces remerciements pas du tout fait en retard sous la deadline des 3 mois.

Je remercie enfin toute ma famille pour leur support sans qui je n'en serais sûrement pas arrivé là.

Merci à vous tous,

VLU

Remerciements

Chapitre 1 – Introduction à la 3D Séquentielle

Depuis plus de 60ans, la micro- puis nano-électronique s’est rapidement développée de telle manière qu’elle est aujourd’hui présente de partout : téléphonie, ordinateurs, voitures, domotique, etc. Toutes ces applications nécessitent notamment une puissance et une rapidité de calculs élémentaires (e.g. pour les processeurs des ordinateurs) voire plus complexes (e.g. systèmes GPS) et/ou des capacités de stockage d’informations (i.e. mémoires) de plus en plus importantes et qui se traduisent par une augmentation de la densité de composants sur une puce. Cette augmentation se fait suivant la loi empirique de Moore (aussi appelée « More Moore ») stipulant que le nombre de composants par unité de surface doit doubler tous les deux ans environ (Figure 1.1).

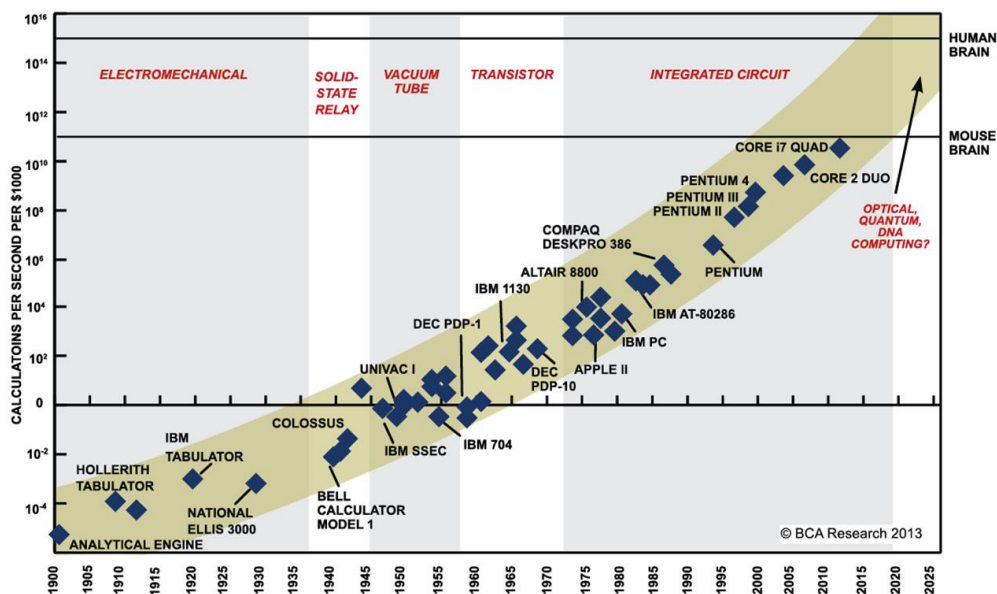


Figure 1.1 : Loi de Moore pour la capacité de calcul au fil des années (pris de [Kurzweil06]).

Dans ce manuscrit, nous discuterons plus particulièrement des transistors à effet de champ métal-oxyde-semiconducteur (MOSFET – *Metal Oxide Semiconductor Field Effect Transistors*), composants élémentaires basés sur la formation d’un canal conducteur entre deux électrodes (appelées source et drain) contrôlée par une tension appliquée sur une troisième électrode (grille). Dans ce cas-là, la loi de Moore se traduit essentiellement par une réduction des dimensions caractéristiques des transistors permettant de réduire leur surface sur une puce et de leur tension d’alimentation afin de limiter leur consommation d’énergie. En particulier, chaque diminution de longueur de grille donne son nom à un nœud technologique associé (e.g. 90nm, 65nm, 32nm...). Néanmoins, nous verrons par la suite que la simple réduction de la taille des transistors ne suffit plus pour des nœuds technologiques avancés.

Dans ce premier chapitre, nous verrons tout d’abord dans la Section 1.1 que la loi de Moore arrive (voire est déjà arrivée) à ses limites et qu’il est nécessaire de trouver une autre voie pour améliorer les performances des puces. De là, nous introduirons et expliquerons les intérêts de l’intégration 3D Séquentielle, avant de détailler dans la Section 1.2 les principales difficultés auxquelles cette option est confrontée. En particulier, nous verrons dans la Section 1.3 qu’il est nécessaire de développer des procédés de fabrication à bas budget thermique pour la production d’un transistor, ce qui est le sujet principal de cette thèse.

1.1 Contexte et intérêt de la 3D Séquentielle

La réduction des dimensions caractéristiques et l'augmentation exponentielles des composants MOSFET pendant plusieurs décennies et sur plusieurs nœuds technologiques ont été permises essentiellement grâce aux propriétés exceptionnelles du silicium utilisé comme semiconducteur et de son oxyde (SiO₂) comme isolant pour la grille. Néanmoins, la miniaturisation des transistors suivant la loi de Moore comporte plusieurs limites que l'on peut regrouper dans trois catégories.

Limites physique

De manière intuitive, on peut fixer la fin de la loi de Moore lorsque les dimensions des transistors atteindront la taille de l'atome ou plutôt d'une monocouche d'atome (de l'ordre de l'Angström). Mais bien avant d'en arriver là, des effets parasites apparaissent déjà à l'échelle du nanomètre. Par exemple, lorsque les dimensions du transistor sont réduites, la grille atteint des longueurs comparables aux zones de déplétion qui apparaissent entre les zones sources et drains fortement dopées et le canal. Lorsque les deux zones de déplétion de charge se rejoignent, des effets dits de canaux courts (SCE – *Short Channel Effects*) apparaissent qui diminuent plus ou moins la hauteur de barrière entre la source et le drain (Figure 1.2a) en fonction entre autres de la longueur de grille (L_g), de l'oxyde de grille (permittivité ϵ_{ox} , épaisseur T_{ox}) et de l'épaisseur de la zone de déplétion induite par la grille dans le canal (T_{dep}) [Skotnicki03]. Cet effet est de plus amplifié lorsqu'une tension est appliquée sur le drain (soit un gradient de potentiel V_{DS}) qui diminue davantage la hauteur de barrière par l'effet DIBL (*Drain Induced Barrier Lowering*). Ces deux phénomènes ont alors pour effet de diminuer la tension de seuil du transistor (i.e. tension nécessaire pour établir le canal de conduction) lorsque les longueurs de grille diminuent et on peut alors parler de mauvais contrôle électrostatique du canal par la grille. Une conséquence directe est alors une augmentation du courant à l'état « off », c'est-à-dire quand aucune tension n'est appliquée sur la grille, et donc de la consommation électrique d'une puce.

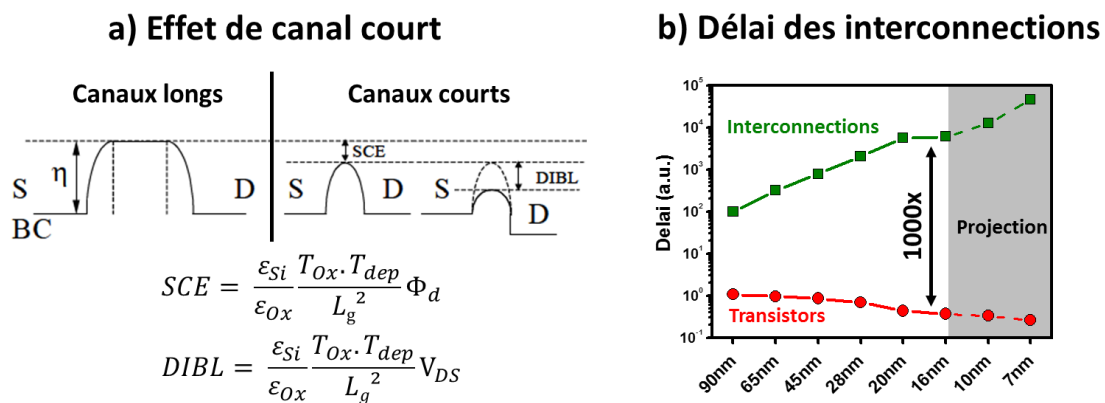


Figure 1.2 : Effets néfastes apparaissant lors de la miniaturisation des transistors : a) Effet de canal court diminuant la tension de seuil des transistors (pris de [Skotnicki03]) et b) augmentation du délai de propagation des signaux dans les interconnexions [Yeap13].

En outre, l'augmentation exponentielle du nombre de composants implique également que le nombre d'interconnexions électriques entre les transistors doit également augmenter. Cependant, il devient alors nécessaire d'avoir des lignes métalliques plus petites et avec une longueur moyenne plus élevée voire d'empiler plusieurs niveaux de métaux à la suite afin d'interconnecter la grande densité de composants. Une conséquence directe est que le temps de propagation des signaux dans les interconnexions, qui est proportionnel à la résistance (avec $R \propto L/S$) des lignes, augmente drastiquement contrairement aux délais dans les transistors qui tendent à diminuer grâce à leur miniaturisation (Figure 1.2b) [Yeap13].

Afin de poursuivre la loi de Moore, plusieurs innovations technologiques ont donc été introduites. Par exemple, de nouveaux matériaux plus ou moins exotiques ont été intégrés dans les transistors, comme les diélectriques à haute permittivité au niveau de l'empilement de grille pour améliorer le contrôle du canal (sujet sur lequel nous reviendrons plus en détails dans le Chapitre 4) ou bien de faible permittivité entre les interconnexions afin de réduire les capacités parasites au niveau du *back-end-of-line* (BEOL) et réduire les délais de propagation [Havemann02]. De plus, l'aluminium utilisé initialement pour les lignes métalliques a été remplacé par le cuivre plus conducteur pour les mêmes raisons. L'introduction du cuivre a néanmoins – entre autres – nécessité une meilleure gestion de la contamination, celui-ci ayant tendance à facilement diffuser dans le silicium lors d'un recuit, causant une dégradation des propriétés des composants. L'utilisation du cuivre est donc totalement prohibée dans le *front-end-of-line* (FEOL), c'est-à-dire lors de la fabrication des transistors [Taylor02]. Les recherches pour des technologies plus avancées voient également l'apparition d'éléments des colonnes III (e.g. gallium, indium) et V (e.g. arsenic) du tableau périodique dont l'alliage offre d'excellentes propriétés pour la micro-électronique (mobilité des porteurs plus élevée dans le canal) ou pour la photonique (ajustement de la bande interdite) [Chau05]. Néanmoins, deux freins actuels à l'utilisation des III-V sont un prix relativement élevé ainsi que la difficulté de trouver un oxyde de bonne qualité pour les applications MOSFET comparé au silicium et au SiO_2 . Enfin, les matériaux dits bidimensionnels ou 2D constitués de quelques monocouches d'atomes tels que le graphène ou les dichalcogénures de métaux de transitions (e.g. MoS_2 , WSe_2) sont également investigués afin de profiter de leurs faibles dimensions et de leurs propriétés (e.g. conductivité supérieure du graphène pour les contacts) [Wu13].

Par ailleurs, de nouvelles architectures ont également été introduites afin de résoudre le problème d'effet de canal court et améliorer le contrôle électrostatique de la grille, l'idée principale étant d'augmenter la surface effective de la grille sur le canal (Figure 1.3).

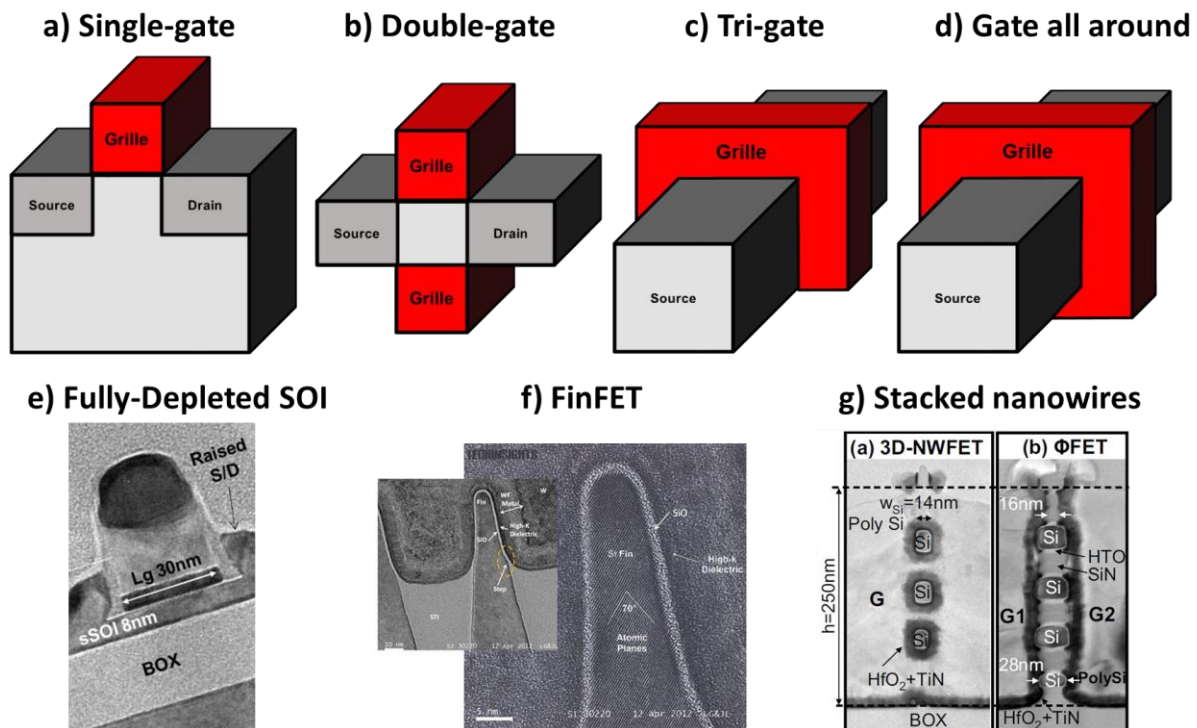


Figure 1.3 : Différentes architectures pour améliorer le contrôle électrostatique de la grille sur le canal : a) Une seule grille « planaire », b) double-grille, c) triple-grille et d) grille enrobante. Exemples d'architectures développées : e) FDSOI [Fenuillet12], f) FinFET [Das12] et g) Nanofils empilés [Dupre08].

Ainsi, plutôt que de contrôler le canal de manière planaire (*single-gate*), des grilles sur d'autres faces du canal sont rajoutées jusqu'à l'architecture dite *gate-all-around* enrobant alors totalement le canal (Figure 1.3d). De ces architectures, on peut plus particulièrement citer la technologie FinFET, à mi-chemin entre un *double-gate* et un *tri-gate* autour d'un ruban de silicium (appelé *fin*) qui constitue le canal (Figure 1.3f). Initialement proposée par l'université de Berkeley [Hisamoto00], c'est essentiellement l'option FinFET qui a été choisie pour les nœuds technologiques inférieurs à 14nm (e.g. IBM [Tsutsui16] ou Samsung [Bae16]) grâce à l'excellent contrôle électrostatique de la grille sur le *fin* permettant de diminuer les effets de canaux courts. Enfin, la technologie nanofils empilés (*stacked nanowires*) semble être l'architecture ultime avec un canal quasiment unidimensionnel contrôlé par la grille tout autour et est investiguée pour les nœuds technologiques inférieures à 7nm [Kuhn12].

Plutôt que d'augmenter la capacité surfacique de la grille, une autre alternative est de diminuer la valeur de T_{dep} afin de réduire les effets de canaux courts (d'après les formules données dans la Figure 1.2a). Une première possibilité pour cela est de doper fortement le canal, la profondeur de déplétion étant d'autant plus faible que le niveau de dopage est élevée. En contrepartie, des effets de *scattering* des porteurs sur les atomes dopants peuvent apparaître, dégradant fortement leur mobilité. Une autre possibilité est de limiter physiquement la zone de charge d'espace en diminuant l'épaisseur de la zone semiconductrice. Pour cela, le canal doit être séparé du substrat, par exemple par l'insertion d'une couche de SiO_2 (alors appelée BOX – *Buried Oxide*). Dans la pratique, cela passe par la fabrication du transistor sur un substrat possédant un film mince de silicium sur un oxyde, substrat alors appelé *Silicon-on-Insulator* (SOI). Ces substrats SOI peuvent être obtenus par exemple grâce à la technologie SmartCut™ développée par Soitec et qui repose sur une implantation d'ions hydrogène à travers un oxyde (Figure 1.4) [Brue195]. La plaque A oxydée est ensuite collée au-dessus d'une autre plaque B puis un recuit est effectué permettant de séparer le substrat en deux au niveau de la couche d'hydrogène implantée, permettant alors d'obtenir un film mince de silicium monocristallin (dépendant de la position de la couche d'hydrogène implantée) sur un BOX d'épaisseur variable (dépendant de l'oxydation effectuée).

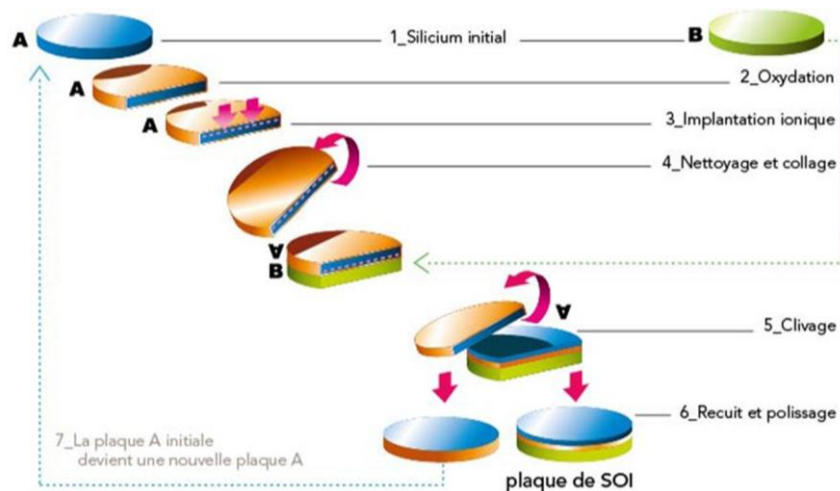


Figure 1.4 : Principe de la technologie SmartCut™ pour l'obtention d'un substrat SOI (pris de <http://soitec.com>).

Lorsque l'épaisseur de silicium est inférieure à une vingtaine de nanomètres, le canal est totalement en déplétion et on parle alors de *Fully-Depleted SOI* (FDSOI - Figure 1.3e) aussi appelé *Ultra-Thin Body and Buried Box* (UTBB). Un avantage supplémentaire de la technologie FDSOI est qu'une polarisation peut également être appliquée sous le BOX (aussi appelé *back biasing*) agissant comme une seconde grille. Par cette polarisation, la tension de seuil du dispositif peut être modulée et par exemple être diminuée pour avoir de meilleures performances, ou augmentée pour diminuer la consommation électrique [Fenuillet10]. La technologie a également été développée jusqu'au nœud 14nm (alors

appelé 14FDSOI chez STMicroelectronics [Weber15a]) mais de manière générale exhibe des performances électriques moins bonnes que la technologie FinFET. Nous verrons toutefois plus loin que cette technologie peut être utilisée pour d'autres applications où la performance n'est pas primordiale.

Enfin, d'autres architectures reposant sur des phénomènes quantiques sont également étudiées, par exemple le tunnel-FET (TFET) basée sur une conduction non plus enclenchée par la formation d'un canal mais sur des effets tunnel entre les bandes d'énergie des sources et drains dopées opposément permettant d'obtenir une commutation beaucoup plus rapide des transistors [Bhuwalka04]. Pour aller plus loin, la logique utilisant des objets quantiques (e.g. spin d'un électron, polarisation d'un photon) pourrait largement surpasser la logique numérique en bénéficiant de leurs propriétés particulières (e.g. superposition et intrication quantique). La création et la manipulation d'un qubit basée sur le spin d'un électron a notamment été démontrée en adaptant une plateforme FDSOI [Hutin16].

Un dernier levier technologique afin d'améliorer les performances des transistors est l'introduction de contraintes sur le canal. Une contrainte permet en effet d'améliorer la mobilité des trous si elle est compressive, ou des électrons si elle est en tension, augmentant de fait les performances du transistor (Figure 1.5) [Horstmann05]. Cette contrainte peut être amenée de différentes manières, par exemple par l'épitaxie d'une couche de maille cristalline différente au-dessus du canal (e.g. SiGe [LeRoyer12]) ou au niveau des sources et drains (que nous traiterons dans le Chapitre 3), ou bien par le *capping* du transistor par une couche contrainte (ce que nous reverrons plus loin dans ce chapitre).

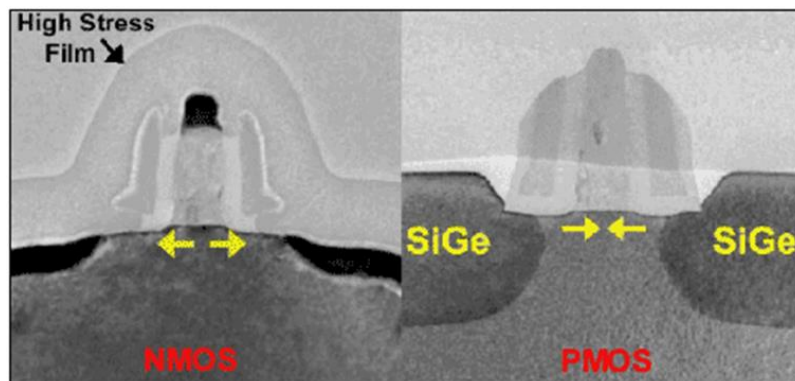


Figure 1.5 : Génération d'une contrainte tensile pour l'amélioration de la mobilité des électrons dans les nMOS (gauche) ou compressive pour l'amélioration de la mobilité des trous dans les pMOS (droite) (pris de [Horstmann05]).

Ainsi, la miniaturisation des transistors a nécessité plusieurs innovations technologiques (nouveaux matériaux, nouvelles architectures, génération de contraintes) afin de faire face aux exigences des nœuds avancés. En particulier, les nanofils empilés semblent être l'architecture ultime pour le contrôle électrostatique et peuvent être associés à des « boosters » de performance (e.g. matériaux III-V, contrainte).

Limites technologique

Par ailleurs, il devient de plus en plus difficile de fabriquer, manipuler ou tout simplement de voir des objets de plus en plus petits. D'une part, la miniaturisation et l'agrandissement du diamètre des plaques (200nm, 300nm voire 450nm) accentue l'importance de la variabilité, c'est-à-dire la différence de propriétés des dispositifs entre différentes plaques, entre différentes puces d'une plaque ou à l'intérieur d'une puce, et qui influe directement sur le rendement de fabrication. Un meilleur contrôle de l'uniformité des procédés est alors nécessaire et passe par des optimisations de l'équipement et des conditions de procédé (e.g. contrôle de la température et/ou de la pression de la chambre). D'autre part, de nouveaux procédés doivent être développés afin de déposer des couches plus minces tout en

conservant un contrôle de l'épaisseur. Par exemple, le dépôt par ALD (*Atomic Layer Deposition*) a été développé permettant de déposer des films à l'échelle de la mono-couche d'atomes et est notamment utilisé pour le dépôt des diélectriques de grille.

Néanmoins, la photolithographie, qui permet de définir les formes des composants par l'insolation et la gravure d'une résine photosensible, reste probablement l'étape la plus critique pour la miniaturisation des composants. En effet, la résolution de la photolithographie dépend principalement de la définition du masque et de la longueur d'onde de la lumière utilisée pour insoler la résine. Cette dernière a donc été progressivement diminuée jusqu'à 193nm (longueur d'onde d'un laser argon-fluor). Réduire davantage la longueur d'onde pose tout d'abord le problème de l'absorption par l'air d'une partie du faisceau. Cette problématique peut être contournée par l'utilisation d'un vide poussé ou d'un liquide transparent dans le cas de la photolithographie à immersion, ce dernier permettant également de diminuer la longueur d'onde grâce à un indice de réfraction différent (Figure 1.6a) [FabianPease08]. La diminution de la longueur d'onde entraîne également une augmentation inversement proportionnelle de l'énergie du faisceau qui peut causer une ionisation du liquide ou des couches sous la résine. La photolithographie dite extrême-UV ($\lambda \sim 10-15\text{nm}$) reste tout de même en développement avec des difficultés de rendement et pour trouver une résine adaptée. D'autres alternatives sont le *multi-patterning* où la photolithographie et la gravure sont répétées plusieurs fois pour atteindre les dimensions voulues (Figure 1.6b) [Yaegashi15], ou bien le *Directed Self-Assembly* (DSA) qui repose sur l'utilisation de copolymères à bloc qui sous certaines conditions peuvent s'ordonner linéairement (Figure 1.6c), permettant de diviser par un facteur deux à quatre le pitch (i.e. l'espacement minimal entre deux motifs) [Black04].

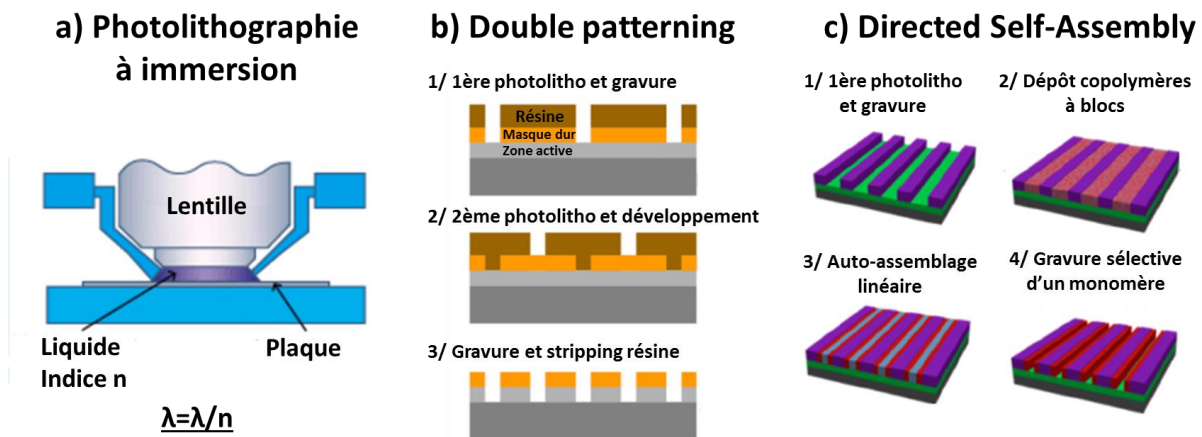


Figure 1.6 : Principes de techniques de patterning avancées : a) Photolithographie à immersion, b) Multi-patterning (cas du double patterning) et c) Directed self-assembly.

Par ailleurs, la métrologie devient également prépondérante dans le nombre d'étapes pour la fabrication d'un dispositif puisque presque chaque procédé doit être contrôlé afin de détecter les plaques, puces ou transistors défectueux le plus tôt possible. Les développements de la microscopie, et en particulier la SEM (*Scanning*) et la TEM (*Transmission Electron Microscopy*) sont primordiales pour la recherche afin de visualiser les dispositifs. L'ellipsométrie et la mesure CD (*Critical Dimension*) sont également régulièrement utilisées après le dépôt ou la gravure de couches. Cependant, l'ellipsométrie requiert une surface relativement importante de l'ordre de la centaine de μm^2 et est effectuée sur des boîtes de mesures dédiées en périphérie des puces. Par conséquent, la valeur mesurée peut différer de l'épaisseur déposée sur des motifs beaucoup plus petits. La mesure CD permet elle de mesurer la largeur d'un motif « vue du dessus » et est effectuée directement sur les dispositifs. L'introduction de nouvelles architectures dense et hors du plan telle que le FinFET complique toutefois l'évaluation fine des motifs. De nouvelles techniques sont donc proposées

combinant par exemple des techniques unidimensionnelles (e.g. SIMS – *Secondary Ion Mass Spectroscopy* – permettant d’analyser la composition d’un échantillon) et 2D (e.g. *Atome Probe Microscopy* permettant de détecter une particule) voire de nouvelles méthodes sont développées telles que le Scalpel SPM (*Scanning Probe Microscopy*) [Vanderworst16]. Ces techniques, bien que destructives, permettent de mieux optimiser les procédés sur des architectures plus complexes.

Limites économique

Enfin, le coût de fabrication explose avec la miniaturisation essentiellement dû à la complexité croissante de fabrication nécessitant plus d’étapes, plus de masques de photolithographie ainsi que plus d’équipements qui sont également plus coûteux [Haron10]. Par conséquent, très peu d’entreprises couvrent aujourd’hui la totalité de la ligne d’un produit, celle-ci étant fragmentée entre la recherche et le développement (e.g. CEA-Leti, IMEC), la production (e.g. GlobalFoundries, Qualcomm) et le design et la vente de produits (e.g. AMD, Apple). On assiste alors souvent à des alliances entre les différentes parties, par exemple entre le CEA-Leti et STMicroelectronics à Grenoble, ou entre TSMC et le National Nano Device Laboratories (NDL) à Hsinchu. Au final, Intel et Samsung sont les deux seules entreprises restantes actives depuis la recherche de nouvelles technologies à la commercialisation de produits.

Cette dernière limite est probablement la plus critique et mène à une alternative à la miniaturisation : le « More Than Moore ». L’idée est alors non plus de mettre plus de composants dans une puce, mais de diversifier les fonctions apportées. Ces fonctions sont multiples et les domaines d’applications conséquents : système de navigation, voiture intelligente, biologie... L’accent n’est alors plus forcément mis sur la performance électrique et la rapidité des dispositifs mais sur la nécessité du marché. Ainsi, selon les applications, on cherchera à avoir des dispositifs qui consomment très peu (e.g. smartphone, montre connectée), qui fonctionnent à des fréquences variables (e.g. télécommunications) ou bien qui interagissent avec d’autres domaines scientifiques (e.g. optique pour la photonique, mécanique pour les MEMS – *Microelectromechanical systems*). Pour cela, les développements effectués pour la miniaturisation des transistors peuvent être adaptés, par exemple avec l’utilisation de matériaux III-V pour l’électronique de puissance ou de l’architecture FDSOI pour les applications à faible consommation. Ainsi, certains fondeurs tels que GlobalFoundries proposent des produits à base de FinFET et de FDSOI pour diversifier leur offre. Dans notre étude, nous utiliserons essentiellement la technologie FDSOI développée conjointement par le CEA-Leti et Soitec et utilisée dans des produits entre autres par STMicroelectronics.

Quelle que soit la voie prise, l’intégration 3D, basée sur l’utilisation de la dimension verticale en empilant une ou plusieurs couches de dispositifs, peut s’avérer être une solution prometteuse. D’une part, l’empilement de composants permet d’augmenter la densité des composants tout en conservant la même surface « vue du dessus ». D’autre part, l’interconnexion verticale des dispositifs permet de diminuer la longueur moyenne des lignes métalliques, diminuant de fait les délais de propagation des signaux. Enfin, elle est également adaptée à une intégration hétérogène de composants (e.g. logique sur logique, mémoire sur logique, III-V sur silicium...) en séparant la fabrication des différents niveaux de la puce. De là, on peut alors distinguer deux types d’intégration 3D : l’intégration dite parallèle où les différentes couches sont fabriquées séparément avant d’être empilées, et l’intégration séquentielle où les couches sont fabriquées successivement les unes sur les autres sur un seul et même substrat (Figure 1.7).

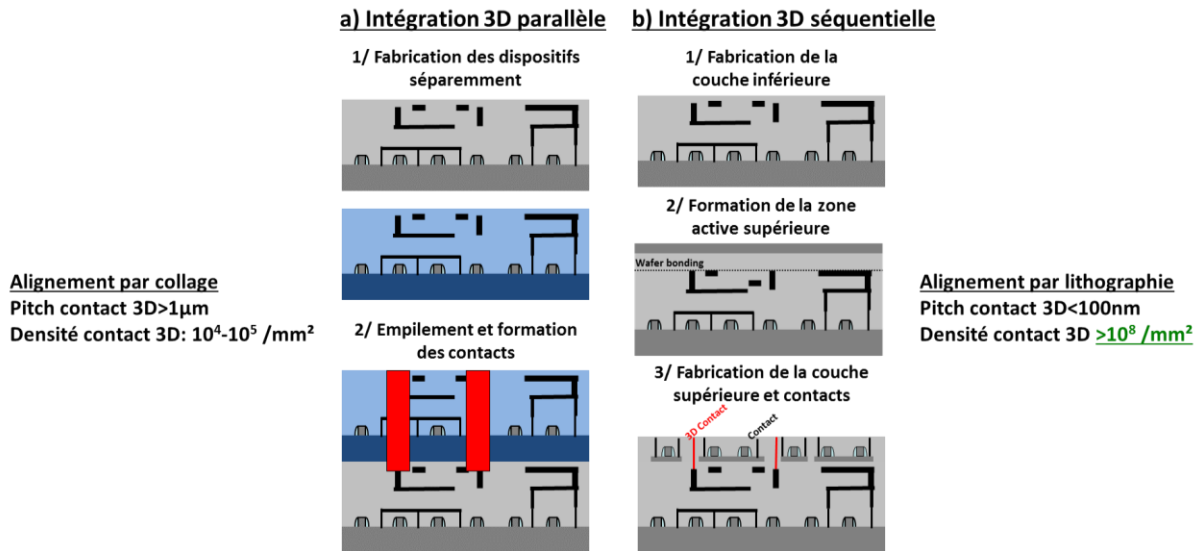


Figure 1.7 : Différences entre les intégrations 3D a) parallèle et b) séquentielle.

Une différence essentielle entre ces deux options est que dans le cas de l'intégration parallèle, l'alignement entre les deux niveaux de composants est effectué lors du collage entre les deux plaques, tandis qu'il se situe au niveau de la fabrication du second étage dans le cas de l'intégration séquentielle. Dans le premier cas, l'alignement se fait donc entre – au moins – deux niveaux de lithographies (première et deuxième couche) ce qui est nécessairement moins précis qu'un alignement sur un seul niveau qui est donc proche d'une étape de lithographie classique. Par conséquent, l'espacement minimal entre les contacts qui est dépendant du désalignement entre les deux niveaux est beaucoup plus faible pour l'intégration 3D Séquentielle. Une conséquence directe est que la densité de contacts 3D réalisable avec cette dernière est de plusieurs ordres supérieurs, pouvant par exemple atteindre 10^8 contacts/mm² en prenant en compte les règles de dessin du 14nm (Figure 1.8).

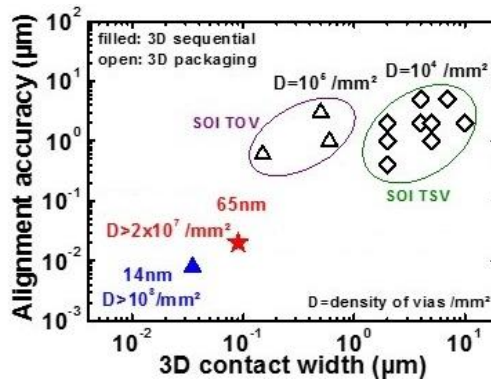


Figure 1.8 : Densité de contacts 3D/via par mm² atteignable selon la technologie (intégration parallèle ou séquentielle) et le nœud technologique (pris de [Brunet16]).

Ainsi, l'intégration 3D séquentielle, parfois également appelée 3D monolithique ou 3D VLSI (*very large scale integration*) est une option intéressante pour augmenter la densité de composants sans diminuer leurs dimensions. Cette option pourrait de fait permettre de gagner un nœud technologique en empilant deux niveaux à un nœud technologique précédent (e.g. 14nm sur 14nm=10nm) tout en évitant les complexités de la miniaturisation citées précédemment. De plus, l'intégration 3D Séquentielle peut s'avérer économiquement plus favorable que le développement d'un nouveau nœud [Gitlin16]. Enfin, cette option retient également les avantages de l'intégration verticale, à savoir une longueur moyenne d'interconnexions plus faible et une certaine robustesse vis-à-vis d'une intégration hétérogène. Néanmoins, nous allons voir dans la section suivante que l'intégration 3D séquentielle reste relativement complexe et comporte quelques difficultés technologiques.

1.2 Problématiques de la 3D Séquentielle

L'intégration 3D Séquentielle est essentiellement confrontée à deux difficultés technologiques : l'obtention d'une zone active supérieure sur laquelle les composants sont fabriqués et la préservation de l'intégrité de toutes les couches inférieures lors de l'intégration des niveaux supérieures.

Zone active du niveau supérieur

La première difficulté est la formation d'une zone active (de silicium par exemple) pour la fabrication de la couche supérieure de composants au-dessus de la première couche de dispositifs. Plusieurs solutions ont été proposées que l'on peut classer dans trois catégories.

- Recristallisation d'une couche amorphe

Après la fabrication des composants du niveau inférieur, un oxyde est déposé et planarisé afin d'encapsuler les composants. Ensuite, une couche de silicium amorphe est déposée avant d'être recristallisée par un recuit laser nanosecondes (Figure 1.9a) [Yang15]. Ce recuit étant extrêmement court, des températures importantes peuvent être atteintes sans endommager la couche inférieure comme nous le verrons par la suite. L'avantage de cette solution est qu'elle est simple et peu coûteuse. En revanche, la zone active supérieure est alors composée de grains de polysilicium orientés aléatoirement dont les propriétés électriques sont inférieures à celles d'un silicium monocristallin, diminuant alors les performances des transistors.

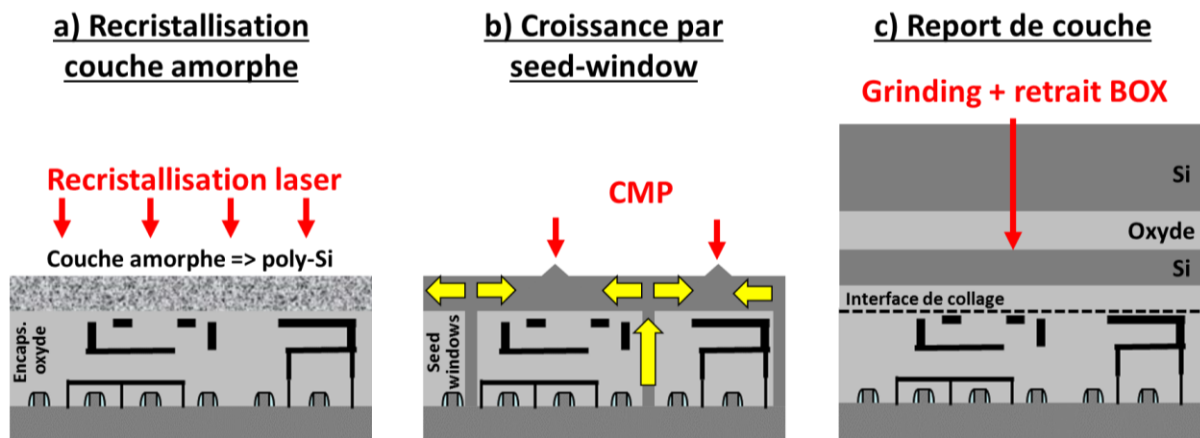


Figure 1.9 : Différentes techniques pour l'obtention d'une zone active au-dessus d'une couche de dispositifs : a) Recristallisation d'une couche laser, b) croissance à travers des cavités (seed-window) et c) report de couche.

- « Seed-Window »

Dans ce cas-ci, une fois l'oxyde d'encapsulation déposé et planarisé, celui-ci est gravé jusqu'à déboucher sur la zone active inférieure. Une croissance épitaxiale permet alors de remonter l'information cristalline à travers les cavités (aussi appelés *seed windows*) puis de former une couche monocristalline via une croissance latérale (Figure 1.9b). Cette méthode comporte néanmoins trois désavantages. Tout d'abord, lorsque les couches épitaxiales de deux cavités se rejoignent, des mottes peuvent se former, nécessitant alors une nouvelle étape de planarisation pour ôter ces zones défectueuses. Une manière de contourner ce problème est de réaliser une épitaxie jusqu'au niveau du trou, puis de déposer une couche amorphe plane qui est ensuite recristallisée par laser par exemple [Son07]. D'autre part, la présence des *seed windows* occupe une zone qui ne peut être utilisée pour la fabrication de dispositifs, réduisant donc la densité de composants atteignables. Enfin, bien que l'épitaxie permette la croissance d'une couche d'excellente qualité cristalline, celle-ci n'est pas totalement parfaite et peut comporter des défauts qui dégradent les performances des dispositifs.

- Report de couche

Une dernière méthode est de reporter une plaque SOI au-dessus du substrat initial par un collage moléculaire entre deux oxydes [Brunet14]. Le silicium bulk ainsi que le BOX sont ensuite retirés par voie mécanique (grinding) et humide afin d’obtenir une zone active monocristalline (Figure 1.9c). Si cette méthode est plus coûteuse (car elle nécessite la destruction d’un substrat SOI complet) et plus complexe, elle permet toutefois d’obtenir une zone active supérieure de qualité parfaite et d’épaisseur bien contrôlée puisque dépendantes uniquement du substrat de départ. De plus, le collage permet une plus grande robustesse dans le type de zone active supérieure, avec la possibilité d’obtenir une couche orientée suivant une autre direction ou d’un autre matériau (e.g. Ge ou matériau III-V) par exemple.

Préservation de l’intégrité niveau inférieur

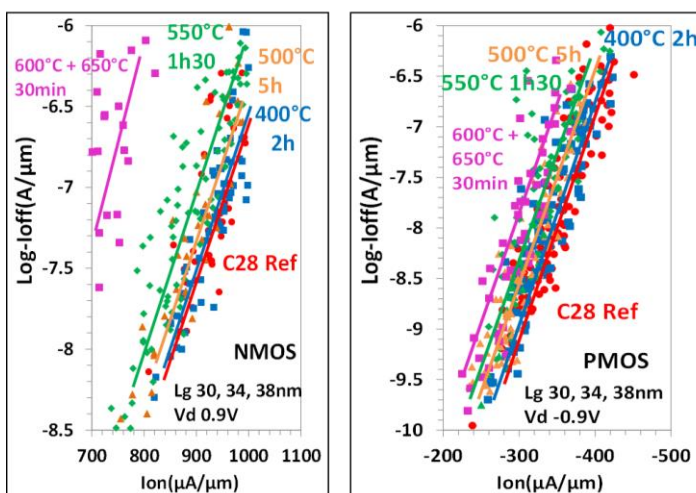
Une seconde problématique de l’intégration 3D Séquentielle est qu’il faut pouvoir fabriquer les composants du niveau supérieur sans endommager tout ce qu’il y a en-dessous. Par en-dessous, nous entendons les transistors mais aussi éventuellement des lignes métalliques (alors appelées BEOL intermédiaire). Ces lignes intermédiaires peuvent être obligatoires afin d’adresser unitairement une très grande densité de composants, sans quoi un problème de congestion de routage peut apparaître. Cette seconde problématique se traduit alors en une limitation du budget thermique (i.e. couple température + temps des procédés) pour la fabrication du niveau supérieur qui est le sujet principal de cette thèse.

Avant tout, il est nécessaire de déterminer le budget thermique limite pour le niveau du haut, soit le budget thermique admissible par le niveau bas (i.e. transistors & BEOL intermédiaire) avant une dégradation de ses performances électriques.

- Stabilité thermique des transistors

Tout d’abord, afin de déterminer ce budget thermique seuil, des études ont été réalisées sur la stabilité thermique des transistors [Fenouillet14a]. Pour cela, différents recuits four ont été effectués sur des dispositifs après leur fabrication mais avant la formation des contacts pour décorrélérer le comportement des transistors et du BEOL. Ces recuits permettent de simuler le budget thermique du niveau supérieur et correspondent à différentes étapes de fabrication d’un transistor qui seront décrites plus en détails dans la section suivante. La Figure 1.10a montre des résultats électriques sur les deux types de transistors (n et pMOS) avant et après recuit.

a) Performances électriques



b) Observations morphologiques

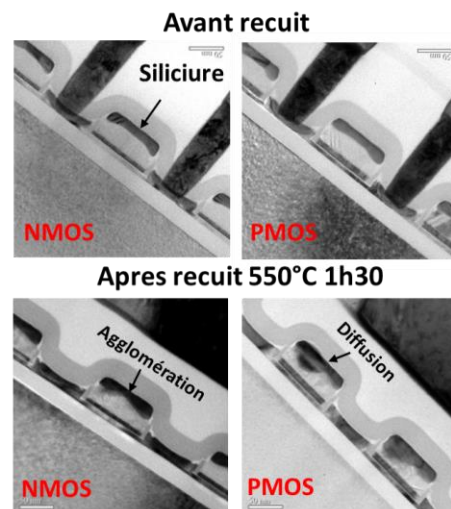


Figure 1.10 : a) Compromis Ion/Ioff pour des nMOS (gauche) et pMOS (droite) d’une technologie 28FDSOI et b) Observations TEM avant et après recuit des silicures sur la grille (pris de [Fenouillet14b]).

Le compromis I_{on}/I_{off} est une figure de mérite permettant d'évaluer les performances d'un transistor à une tension d'utilisation V_{DD} donnée. Ainsi, le I_{on} correspond au courant débité lorsque le transistor est actif ($V_{drain}=V_{grille}=V_{DD}=0.9V$ ici) et le I_{off} correspond au courant de drain lorsque le transistor est passif ($V_{grille}=0V$ et $V_{drain}=V_{DD}$), ce dernier représentant donc la consommation du transistor lorsqu'il n'est pas utilisé. La Figure 1.10a montre un début de dégradation du compromis I_{on}/I_{off} (i.e. déplacement de la pente vers la gauche) lorsque le recuit est supérieur à $550^{\circ}C$ 1h30, dégradation beaucoup plus visible sur les transistors de type N. Cette dégradation est en fait essentiellement due aux siliciures (ici, un alliage nickel-platine) des transistors qui sont utilisés pour améliorer la conductivité des accès. En effet, les observations TEM sur la Figure 1.10b montrent une dégradation de la morphologie des siliciures avec soit un effet d'agglomération, soit une diffusion du siliciure. Dans les deux cas, la résistance des siliciures augmente et dégrade donc les performances des transistors. Ainsi, un premier budget thermique seuil peut être fixé à $500^{\circ}C$ 5h. Il est toutefois important de noter que la stabilité thermique du transistor va dépendre du métal utilisé pour la siliciuration. Ainsi, des études sur une technologie 14FDSOI avec des siliciures comprenant un pourcentage de platine plus élevé (15% au lieu de 10% précédemment) montrent que les transistors sont cette fois stables jusqu'à $550^{\circ}C$ 2h [Fenouillet14b]. De la même manière, la stabilité des transistors peut être optimisée en implantant du tungstène ou du fluor dans le siliciure NiPt [Carron14], ou bien en changeant l'alliage pour du nickel-cobalt par exemple [Deprat15].

Par ailleurs, des températures plus élevées peuvent être atteintes lorsque la durée du recuit est diminuée. Ainsi, l'intégrité des composants du niveau inférieur est préservée après un recuit à $600^{\circ}C$ pendant une ou deux minutes [Batude15]. Afin de réduire davantage la durée de recuit – et donc augmenter la température limite, de nouveaux procédés de traitement thermique doivent être mis au point, les techniques de recuits conventionnelles pouvant difficilement descendre en dessous de temps de l'ordre de la seconde. Pour cela, des recuits laser ont été développés et peuvent être divisés en deux catégories : les recuits laser millisecondes et nanosecondes.

Dans le premier cas, on peut notamment citer le DSA (*Dynamic Surface Anneal*) développé par Applied Materials (AMAT) qui est basé sur le balayage d'un faisceau laser sur la plaque. Des températures maximales proches de $1300^{\circ}C$ peuvent alors être atteintes sur des durées de l'ordre de la milliseconde selon la vitesse de balayage. Concernant la stabilité thermique des transistors, un budget thermique de $800^{\circ}C$ 0.5ms peut être atteint sans dégrader les performances des composants (Figure 1.11a) [Batude15].

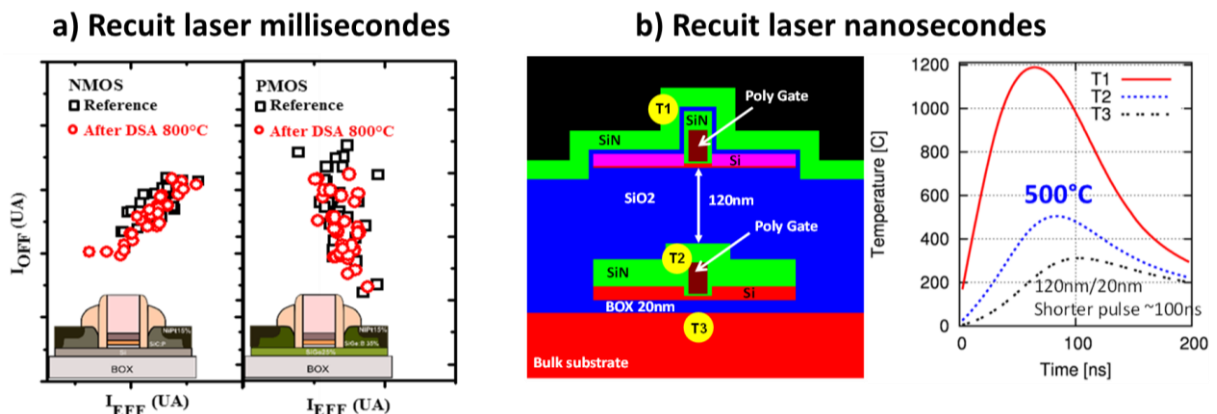


Figure 1.11 : a) Stabilité thermique des transistors après recuit DSA $800^{\circ}C$ 0.5ms (pris de [Batude15]) et b) Simulations d'un recuit laser nanoseconde permettant d'atteindre $1200^{\circ}C$ sur le transistor supérieur (T1) tout en conservant le transistor inférieur (T2) sous $500^{\circ}C$ (pris de [Fenouillet14b]).

Dans le deuxième cas, on peut citer deux types de recuits laser nanosecondes : le laser UV ($\lambda=308nm$) développé par LASSE (*Laser Systems & Solutions of Europe*) et le laser vert ($\lambda=532nm$) d'AMAT. Ces

deux procédés reposent sur l’envoi du faisceau laser par pulse sur une zone variable (e.g. la taille d’une puce) et sur des durées de l’ordre de la centaine de nanosecondes. En revanche, ils diffèrent sur la longueur d’onde utilisée ce qui impacte alors le comportement des matériaux recuits dont l’absorptivité peut varier. Par exemple, dans le cas du laser UV, un silicium amorphe absorbe beaucoup plus l’énergie du faisceau laser qu’un silicium monocristallin, tandis que leurs absorptivités sont relativement équivalentes pour un laser vert. Néanmoins, quelle que soit la longueur d’onde, la faible durée de recuit limite fortement la diffusion de l’énergie thermique de la surface recuite vers les couches inférieures, permettant potentiellement d’atteindre des températures supérieures à 1000°C pour la couche supérieure. Des simulations montrent par exemple qu’une température de 1200°C peut être atteinte au niveau du haut tout en conservant la température du niveau bas sous 500°C, assurant son intégrité (Figure 1.11b). C’est également la raison pour laquelle un recuit laser nanosecondes était utilisé pour la recristallisation d’une couche amorphe au-dessus d’une couche de composants (cf. Section 1.1). A noter tout de même que du fait de la faible longueur d’onde, le recuit laser nanoseconde est fortement affectée par les motifs présents, ce qui peut avoir une influence sur l’uniformité du recuit.

- Stabilité du BEOL intermédiaire

De la même manière, la stabilité du BEOL intermédiaire a été évaluée dans des études précédentes. Comme nous l’avons vu dans la Section 1.1, les matériaux utilisés pour le BEOL standard sont le cuivre pour les lignes métalliques couplé avec un oxyde de faible permittivité (e.g. SiOCH poreux avec $k \sim 2.7$) qui sépare les lignes. Cependant, nous avons également vu que le cuivre est un élément ayant tendance à diffuser lors d’un recuit. C’est la raison pour laquelle le dépôt de cuivre est généralement précédé du dépôt d’une barrière de diffusion, typiquement de Ta/TaN. Un second problème est que le cuivre est prohibé dans le FEOL et pose le risque d’une contamination lors de la fabrication des composants du niveau supérieur par exemple si la plaque vient à se casser. Par conséquent, le tungstène, bien qu’ayant une résistivité 6 fois supérieure à celle du cuivre, a également été évalué pour le BEOL intermédiaire, celui-ci étant moins problématique en terme de contamination. La Figure 1.12a montre que dans le cas du cuivre, le budget thermique le plus haut évalué (500°C 2h) ne dégrade pas la résistance des lignes métalliques [Fenouillet15]. De la même manière, l’intégrité du tungstène est préservée jusqu’à un budget thermique de 550°C 5h. Toutefois, la capacité du SiOCH entre les lignes métalliques augmente d’environ 5-7% après recuit, dégradant légèrement les délais de propagation des signaux dans le BEOL. Cette augmentation de capacité peut être attribuée à une contraction du diélectrique poreux sous l’effet d’un budget thermique et d’une légère modification de la permittivité (Figure 1.12b) [Deprat16]. Une alternative serait alors d’utiliser un autre oxyde (e.g. SiO₂) plus stable thermiquement mais avec une permittivité plus élevée ($k \sim 3.9$).

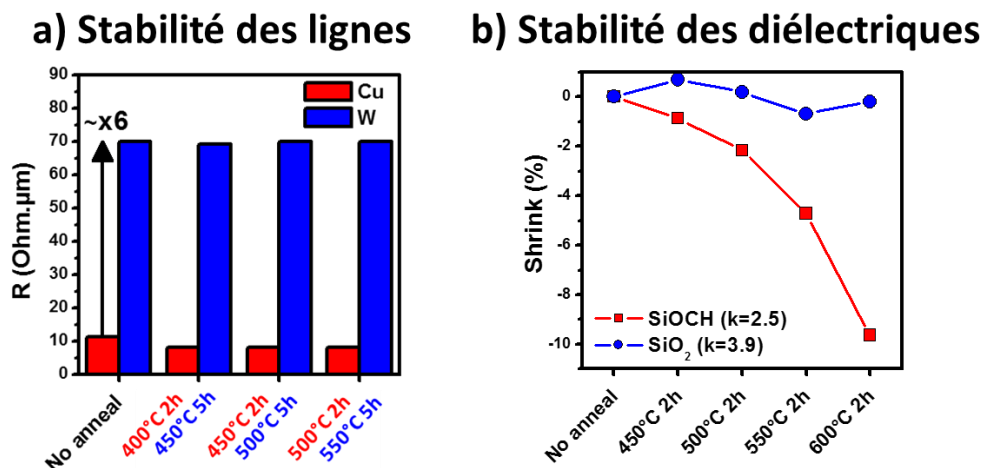


Figure 1.12 : Etudes de stabilité thermique sur a) la variation de résistance des lignes métalliques (pris de [Fenouillet15] et b) la variation d’épaisseur des diélectriques interlignes (pris de [Deprat16]).

Au final, le budget thermique limite pour la stabilité des transistors et du BEOL intermédiaire a été fixé dans notre étude à 500°C pour des temps de procédés allant jusqu'à 2 heures afin de garder une certaine marge (Figure 1.13). Les procédés DSA et recuit laser nanosecondes seront également étudiés pour certaines applications dans ce travail.

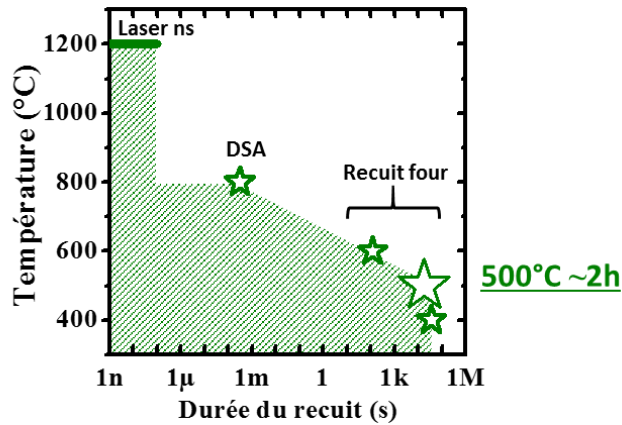


Figure 1.13 : Graphe représentant la fenêtre de budget thermique acceptable pour les couches inférieures. Les étoiles représentent les budgets thermiques où la stabilité des transistors et/ou du BEOL intermédiaire est validée (pris de [Batude15]).

Ainsi, les deux problématiques les plus importantes de l'intégration 3D Séquentielle sont l'obtention d'une zone active de bonne qualité ainsi que la fabrication à bas budget thermique des composants du niveau supérieur. Des groupes de recherche ont proposé différentes solutions à ces deux problèmes selon l'application voulue. Le groupe AIST (National Institute of Advanced Industrial Science and Technologies) [Irisawa14] et IBM Zurich [Deshpande15] proposent par exemple une co-intégration de transistors III-V (InGaAs) sur transistors SiGe (i.e. canal du transistor en SiGe), profitant de l'hétérogénéité de l'intégration pour optimiser séparément le niveau de pMOS en-dessous et le niveau de nMOS au-dessus, en particulierité en termes de mobilités des trous et des électrons respectivement (Figure 1.14a).

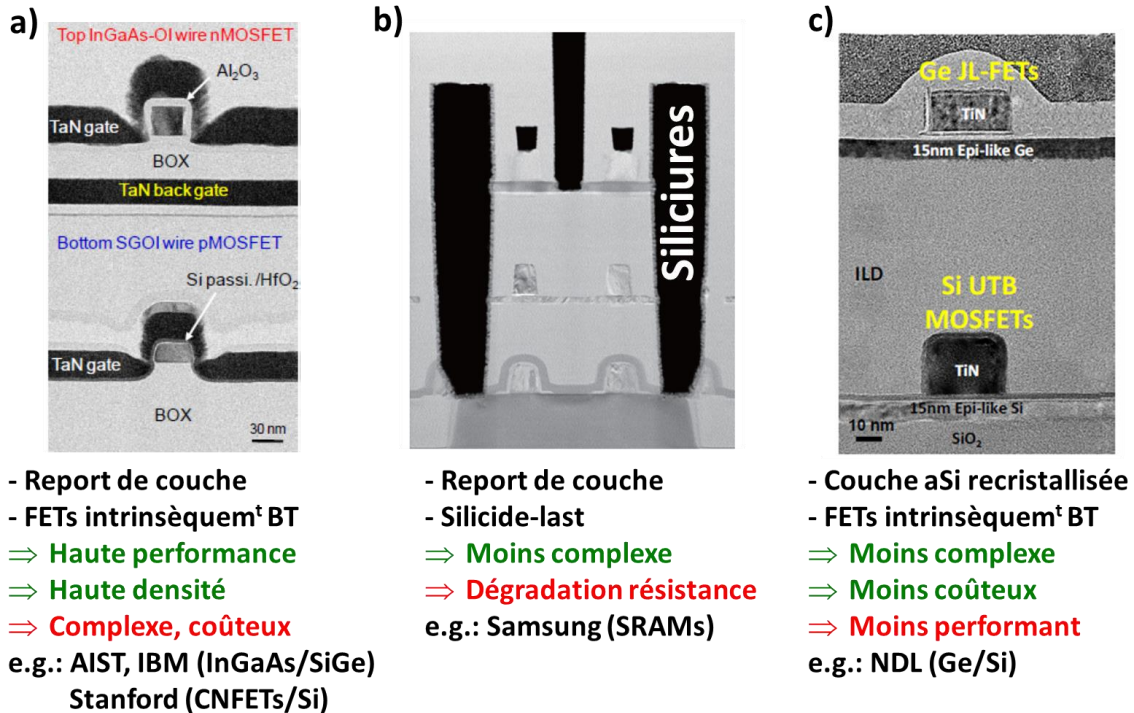
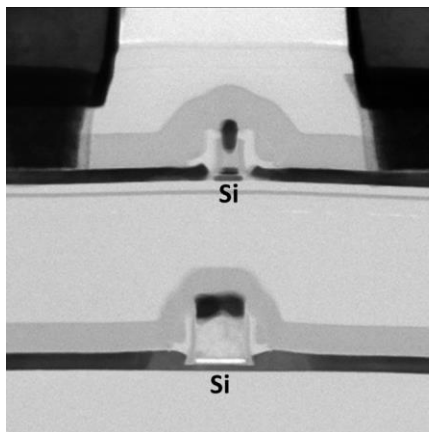


Figure 1.14 : Synthèse non exhaustive des différentes intégrations 3D Séquentielles proposées dans la littérature (pris de [Irisawa14], [Jung07], [Shen14]).

Dans les deux cas, la zone active supérieure est obtenue par report de couche tandis que la fabrication d'un transistor III-V est intrinsèquement faite à basse température ($T < 400^\circ\text{C}$) afin de ne pas endommager ces matériaux. Cette architecture 3D Séquentielle possiblement ultime offre donc des performances optimales, mais la maturité des matériaux III-V et le coût de fabrication peuvent compliquer l'industrialisation de cette technologie. Le groupe de Stanford propose lui une intégration de transistors à base de nanotubes de carbone (CNFETs – *Carbon Nanotubes FETs*) également intrinsèquement fabriqués à basse température et sur une couche reportée au-dessus de transistors silicium [Shulaker14], technologie qui souffre néanmoins également du manque de maturité des CNFETs. Samsung a également proposé sa version de l'intégration 3D Séquentielle en empilant des circuits mémoires (SRAM – *Static Random Access Memory*) [Jung07]. Afin de s'affranchir de la problématique de la stabilité des composants inférieurs, les siliciures ont été formés à la fin de l'intégration (dite *silicide-last*), résultant en une siliciuration latérale (Figure 1.14b). L'inconvénient de cette option est alors une résistance d'accès plus élevée au niveau des zones non métallisées proches de la grille. Enfin, une autre technologie notable utilisant l'intégration 3D Séquentielle est celle proposée par NDL où la zone active supérieure est obtenue par recristallisation laser d'une couche amorphe (alors appelée *epi-like*) permettant l'intégration de transistors Ge sur Si par exemple (Figure 1.14c) [Shen14]. Cette intégration est moins complexe et moins coûteuse mais mène aussi à des performances électriques moins bonnes dues à la présence d'une zone active polycristalline.

Au CEA-Leti, l'intégration CoolCube™ a été développée et sera l'option étudiée dans ce manuscrit. L'intégration CoolCube™ repose initialement sur un empilement de composants logiques sur composants logiques (e.g. Si sur Si, Si sur SiGe...). Les composants du niveau inférieurs sont alors d'abord fabriqués à haute température et en intégrant les siliciures (*silicide-first*). La zone active supérieure est ensuite obtenue par report de couche grâce à un collage oxyde-oxyde suivie de la fabrication à bas budget thermique d'un second niveau de transistors (Figure 1.15) [Brunet16].



Intégration CoolCube™

- Report de couche
- Silicide-first
- ⇒ Haute densité
- ⇒ Relativement peu complexe
- ⇒ Relativement moins coûteux
- ⇒ Relativement performant
- ⇒ Nécessite la fabrication d'un transistor Si à bas budget thermique

Figure 1.15 : Principes et observation TEM de l'intégration CoolCube™ (pris de [Brunet16]).

Cette intégration permet également une utilisation optimale de la troisième dimension en termes de densité, est légèrement moins complexe et moins coûteuse que les technologies proposées par AIST, IBM ou Stanford, mais aussi potentiellement plus performantes que les options proposées par Samsung et NDL. L'intégration CoolCube™ reprend aussi sensiblement les mêmes étapes de fabrication que celles utilisées dans les fonderies et est donc plus compatible avec le milieu industriel. En revanche, elle nécessite donc la fabrication d'un transistor supérieur à bas budget thermique. Nous allons ainsi voir dans la prochaine section que le process flow (i.e. toutes les étapes de fabrication nécessaires) d'un transistor à base de silicium comporte plusieurs points critiques qu'il faut résoudre afin de rendre possible cette intégration.

1.3 Développement de briques technologiques à bas budget thermique

Nous allons ici détailler le process flow d'un transistor afin d'identifier les points critiques où un développement à bas budget thermique est nécessaire. La fabrication d'un transistor peut en général être découpée en 4 étapes principales : 1) formation de la zone active, 2) formation de la grille, 3) formation des jonctions et 4) formation des contacts. Nous prendrons ici le cas particulier d'une technologie Tri-Gate FDSOI avec des règles de dessin 65nm développée au Leti. Comme nous l'avons vu auparavant, d'autres technologies existent et peuvent différer dans le process flow. Néanmoins, les étapes ou briques technologiques critiques que nous allons identifier sont généralement également utilisées dans les autres technologies.

Formation de la zone active

La fabrication d'un transistor FDSOI commence donc par une plaque SOI (Si 16nm sur un BOX 145nm) obtenue grâce à la technologie SmartCut™ (Figure 1.16a).

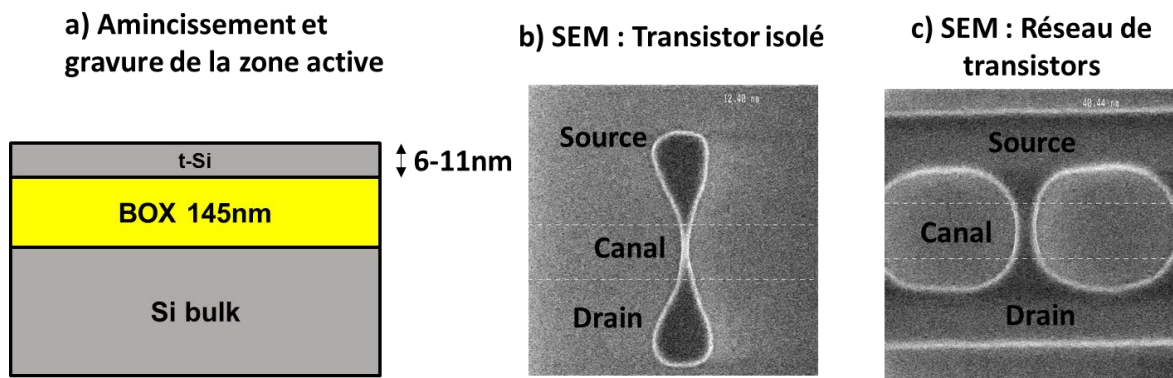


Figure 1.16 : Etapes de formation de la zone active. a) Amincissement de la plaque SOI et gravure des zones actives. Observations SEM après gravure b) d'une zone active isolée ou c) en réseau.

Le film de silicium est tout d'abord aminci en oxydant la surface puis en retirant l'oxyde formé par voie humide, afin d'obtenir un film mince (t-Si) et avec une épaisseur bien contrôlée allant de 6nm à 12nm. Ce film est ensuite patterné grâce à une étape de photolithographie des zones actives suivie d'une gravure anisotrope par plasma. Le motif obtenu pour un transistor isolé ressemble alors à des ailes de libellules, les deux excroissances permettant d'avoir un accès facile aux sources et drains lors de la formation des contacts (Figure 1.16b). Des motifs en réseau (i.e. à plusieurs grilles) avec des sources et drains communes seront également utilisés (Figure 1.16c).

Formation de la grille

L'empilement de grille est ensuite déposé sur toute la plaque (Figure 1.17a). Un masque dur (*hard mask* en anglais – HM) composé d'une bicouche de SiN et de TEOS est également déposé qui est nécessaire pour l'étape de surélévation des sources et drains comme nous le verrons par la suite. L'empilement de grille est enfin gravé et, tout comme pour les sources et drains, la grille est déportée afin de faciliter l'accès à l'électrode (Figure 1.17b et c).

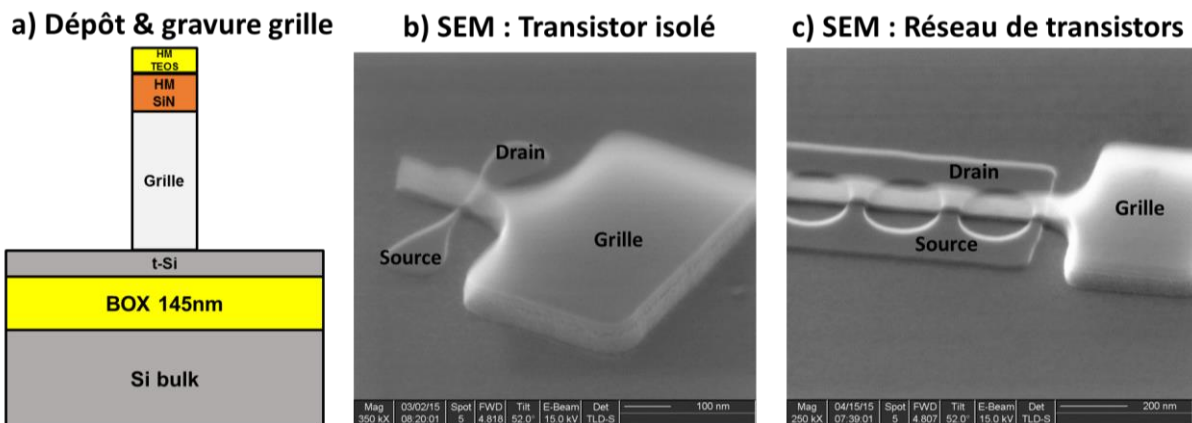


Figure 1.17 : Etapes de formation de l'empilement de grille. a) Dépôt et gravure de l'empilement de grille; Observations SEM après gravure b) d'un transistor isolé et c) d'un réseau de transistors.

Formation des jonctions

Après le dépôt et la gravure de l'empilement de grille, les jonctions doivent être formées. Celles-ci correspondent aux zones fortement dopées d'un transistor MOSFET, c'est-à-dire les sources et drains. Pour cela, un diélectrique est tout d'abord déposé et gravé afin de former des espaceurs (appelés *Spacer0*) qui permettent d'isoler électriquement la grille des futurs accès source et drain (Figure 1.18a).

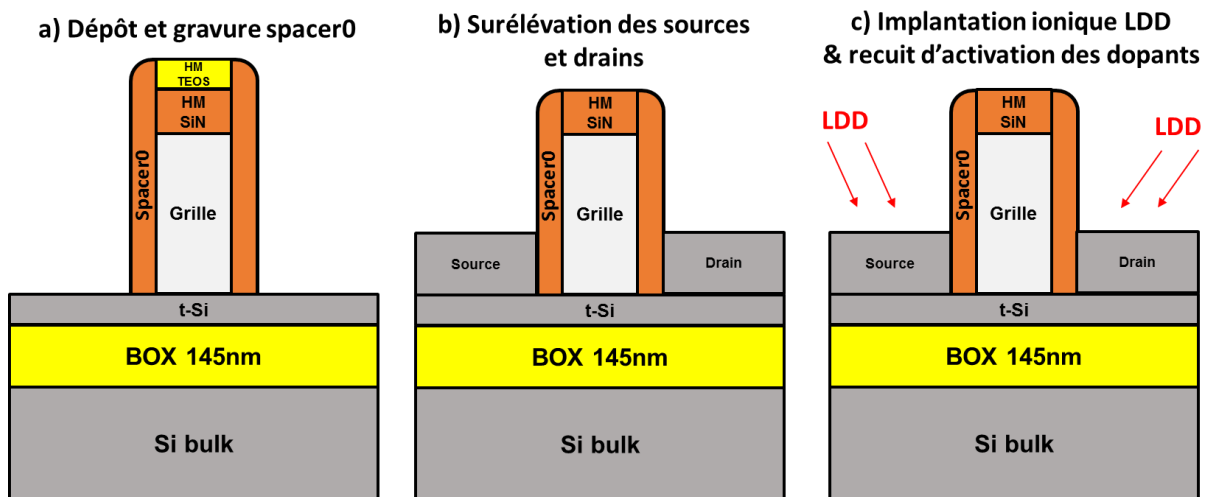


Figure 1.18 : Etapes de formation des jonctions sources et drains. a) Dépôt et gravure des espaceurs de grille, b) Epitaxie de surélévation des sources et drains et c) Implantation ionique et recuit d'activation des dopants LDD.

Le film de silicium étant trop fin pour pouvoir former les jonctions correctement, une étape de surélévation des sources et drains est nécessaire afin d'épaissir le film et est réalisée par croissance épitaxiale (Figure 1.18b). Le masque dur permet alors d'éviter une croissance parasite sur la grille qui créerait des courts-circuits. À ce niveau, le masque dur TEOS est normalement totalement consommé après les différentes étapes de gravure et d'épitaxie. Une première implantation dite LDD (pour *lightly-doped sources and drains*) est alors effectuée permettant de doper les sources et drains près du canal afin de diminuer la résistance électrique des accès (Figure 1.18c). C'est à ce moment que le type de transistor est déterminé : implantation de donneurs (typiquement phosphore, arsenic) pour des nMOS ou d'accepteurs (bore, BF_2) pour des pMOS. Un recuit est ensuite réalisé afin de fournir l'énergie nécessaire aux dopants pour substituer un atome du semiconducteur et les activer électriquement. Ce recuit d'activation peut également permettre de guérir les défauts créés par l'énergie de l'implantation ionique, par exemple en recristallisant les zones amorphisées.

Formation des contacts

Avant de former les contacts, le masque dur SiN restant au-dessus de la grille doit tout d'abord être retiré, celui-ci étant un isolant. Pour cela, un jeu d'espaceurs temporaires (appelés *Spacer1*) en oxyde de silicium est déposé et gravé anisotropiquement, ce dernier étant nécessaire afin de protéger les espaceurs de grille lors du retrait masque dur (Figure 1.19a).

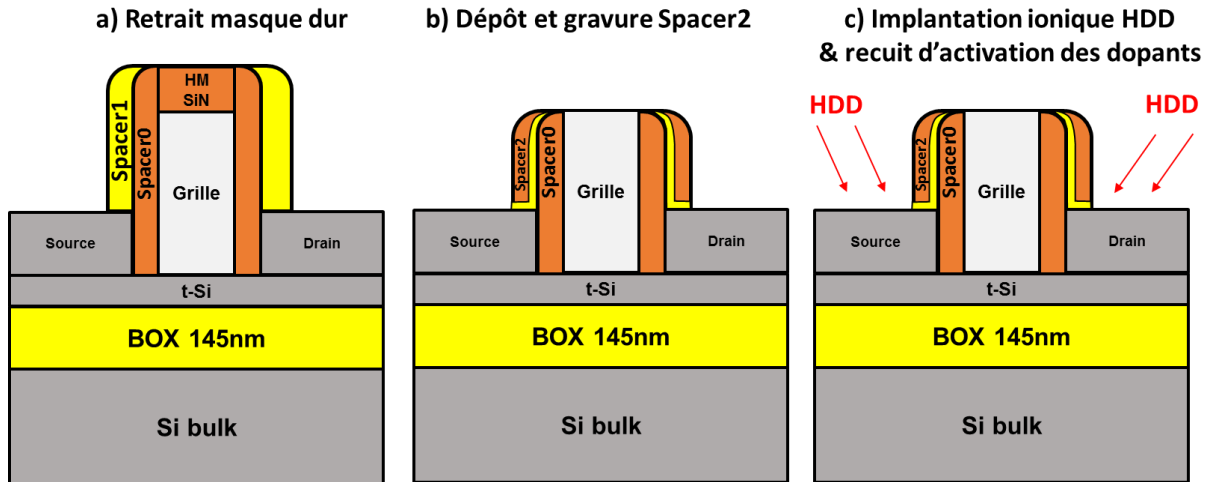


Figure 1.19 : Etapes de formation des accès. a) Retrait du masque dur nitrure sur la grille, b) dépôt et gravure des espaceurs de contact et c) implantation ionique et recuit d'activation des dopants HDD.

Le SiN est ensuite retiré sélectivement par rapport aux espaceurs de protection, typiquement dans une chimie H_3PO_4 (acide phosphorique) qui n'attaque pas ou peu le SiO_2 . Enfin, le *Spacer1* est retiré par voie humide sélectivement par rapport aux espaceurs de grille, par exemple dans une chimie HF (acide fluorhydrique). Afin d'avoir une prise de contact la moins résistive possible sur les trois accès, une siliciuration (i.e. formation d'un alliage entre le silicium et un métal) doit être effectuée. Avant cela, un troisième jeu d'espaceurs (appelés *Spacer2*) est déposé et gravé, celui-ci permettant d'éloigner les futurs contacts des sources, drains et grilles et d'éviter des couplages parasites entre (Figure 1.19b). Une deuxième implantation HDD (pour *highly-doped sources and drains*) est ensuite faite suivie d'un recuit d'activation encore une fois pour diminuer la résistance dans ces régions (Figure 1.19c). La siliciuration des accès sources, drains et grilles est ensuite réalisée en déposant un alliage (e.g. NiPt) sur toute la plaque. Un premier recuit est alors effectué afin de faire réagir le silicium avec le métal et former des siliciures (e.g. NiPtSi). Un retrait sélectif du métal n'ayant pas réagi par voie humide est ensuite effectué, suivi d'un deuxième recuit permettant d'optimiser la résistivité du siliciure (Figure 1.20a).

Une couche CESL (*Contact Etch Stop Layer*) est ensuite déposée, et le tout encapsulé dans un oxyde dit IMD (*Inter-Metal Dielectric*). Comme son nom l'indique, le CESL est une couche d'arrêt permettant un meilleur contrôle de la profondeur des contacts lors de leur gravure en présence d'une topographie. En pratique, l'oxyde IMD est gravé avec une chimie très sélective au matériau du CESL (nitrure de silicium) jusqu'à la couche d'arrêt, l'épaisseur gravée d'IMD variant donc selon la zone (grille, source ou drain) tandis que l'épaisseur de CESL restant à graver est identique pour les trois accès. En outre, le CESL peut également permettre d'introduire de la contrainte sur le transistor s'il est compressif ou tensile, améliorant la mobilité des porteurs dans le canal. Les contacts gravés, ceux-ci sont remplis par un métal conducteur, typiquement du tungstène précédé d'une barrière Ti/TiN (Figure 1.20b). Les contacts peuvent ensuite être redirigés par des lignes métalliques sur plusieurs niveaux afin de former des circuits avec d'autres composants, avant d'arriver à la surface où des pads d'aluminium permettent de prendre effectivement contact pour mesurer électriquement les dispositifs (Figure 1.20c).

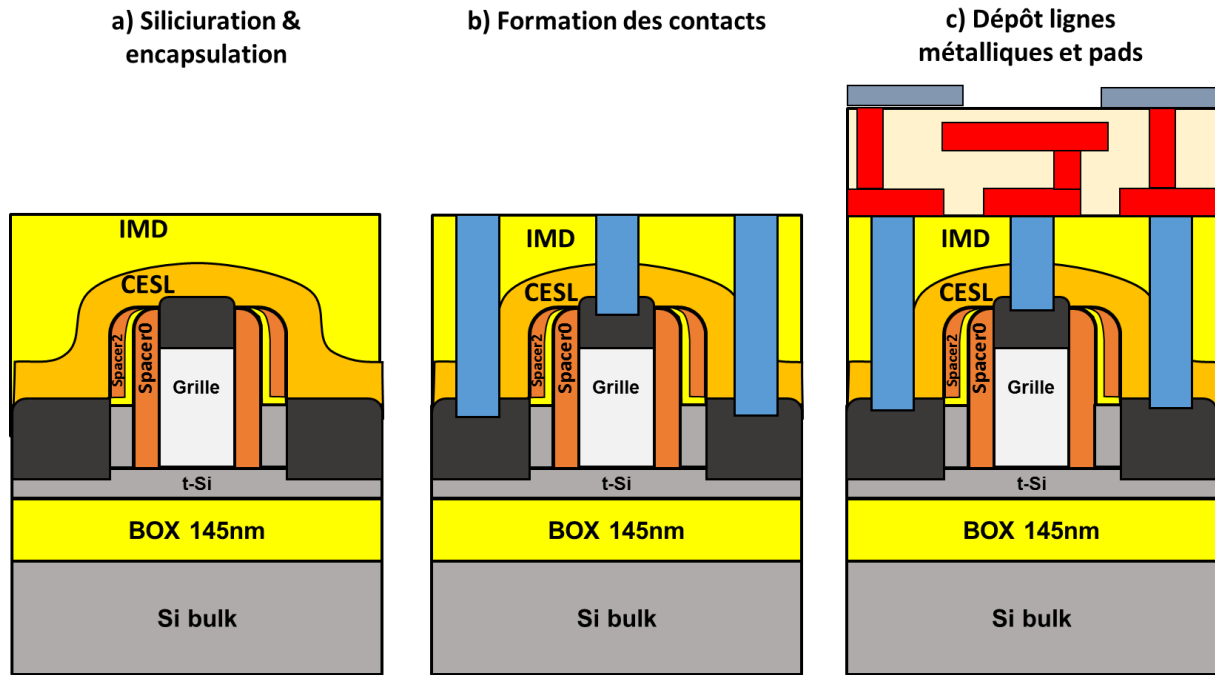


Figure 1.20 : Etapes de formations des contacts. a) Siliciuration des accès et encapsulation du transistor, b) gravure et remplissage des contacts et c) dépôt des lignes métalliques et des pads en aluminium.

Les dimensions caractéristiques du transistor (longueur et largeur du canal) sont alors déterminées par les dimensions de l'empilement de grille et de la zone active respectivement. La Figure 1.21 montre alors des observations TEM d'un transistor complet et d'une partie du BEOL. En particulier, la coupe selon la largeur du canal montre que celui-ci est entouré sur trois de ses côtés de l'empilement de grille, d'où l'architecture Tri-Gate. Ainsi, la largeur de grille effective est égale à la largeur du canal additionnée à deux fois l'épaisseur de la zone active. A noter toutefois que dans ce manuscrit, des largeurs de grille de l'ordre du micromètre seront souvent utilisées. Dans ce cas-là, la hauteur de la zone active est négligeable et le transistor se comporte donc comme un *single-gate*.

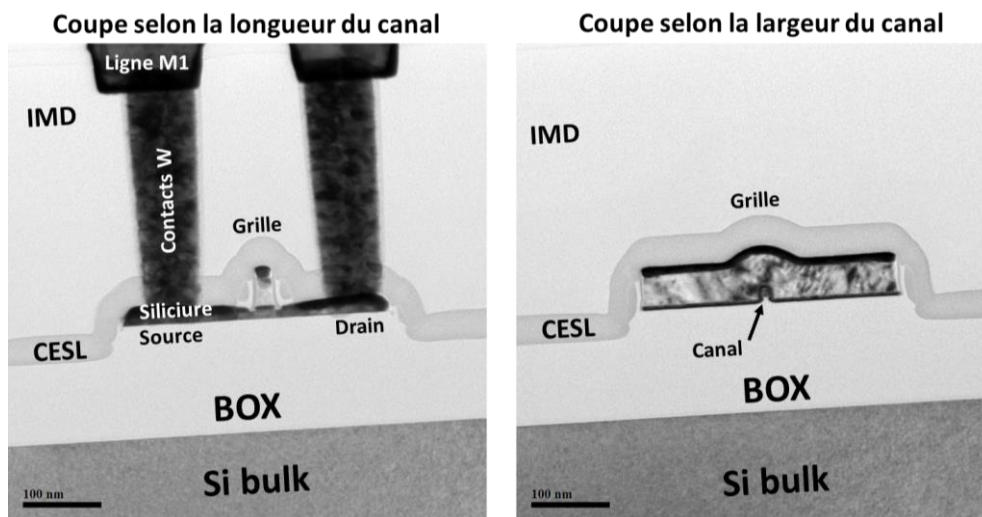


Figure 1.21 : Observations TEM d'un composant complet en coupe selon la longueur du canal (à gauche) ou la largeur du canal (à droite).

Le flow de fabrication (ou process flow) standard utilisé au Leti avec les budgets thermiques associés aux étapes importantes est résumé dans la Figure 1.22.

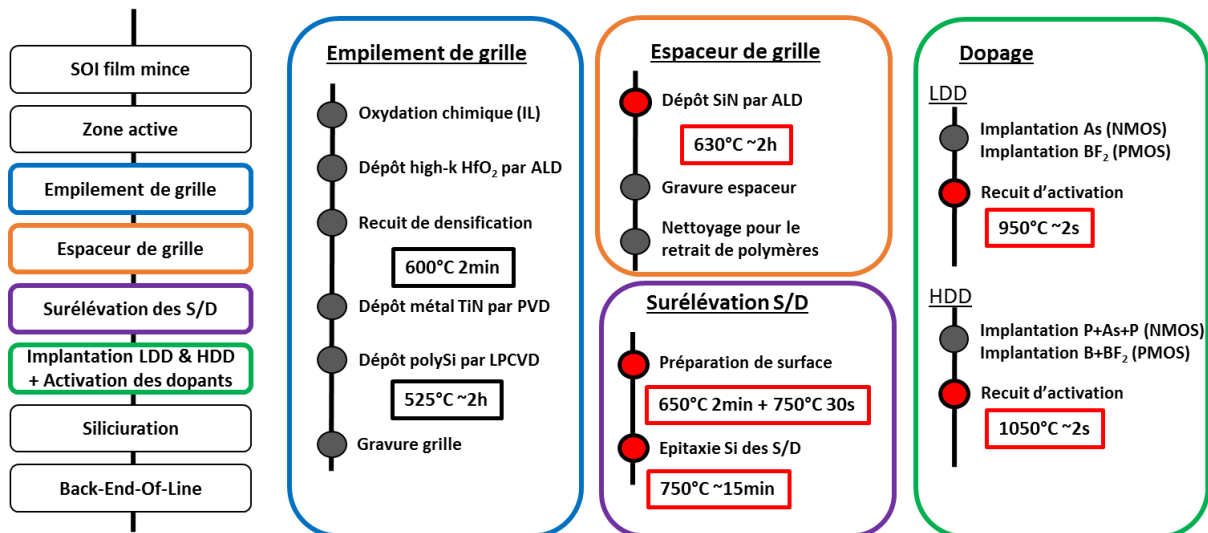


Figure 1.22 : Process flow d'un transistor standard au Leti et budgets thermiques associés. Les budgets thermiques qui doivent être réduits sont en rouge.

Dans la vue d'une intégration CoolCube™ à bas budget thermique, on remarque alors trois étapes critiques : le dépôt des espaceurs de grille, l'épitaxie de surélévation des sources et drains et l'activation des dopants. On peut également rajouter la brique de l'empilement de grille qui « voit » tous les budgets thermiques et qui en est impactée comme nous le verrons plus loin. Un développement à plus bas budget thermique pour ces briques technologiques est donc nécessaire afin de fabriquer les dispositifs du niveau supérieur, le tout sans dégrader leurs performances électriques par rapport aux références à haut budget thermique.

La problématique de l'activation des dopants au niveau des jonctions a déjà été le sujet de travaux précédents dont nous allons succinctement résumer les résultats. L'idée est donc de remplacer le recuit à haute température (1050°C ~2s) par une autre méthode d'activation à plus bas budget thermique.

Recristallisation épitaxiale en phase solide

Une première alternative au recuit à haute température est la technique SPER, soit recristallisation épitaxiale en phase solide (*Solid Phase Epitaxy Regrowth* en anglais). Cette technique repose sur la recristallisation d'un silicium amorphisé par implantation ionique des dopants (qui sont alors majoritairement en position interstitielle) à partir d'une couche monocristalline (appelée *seed*), permettant de garder la même orientation cristalline (Figure 1.23).

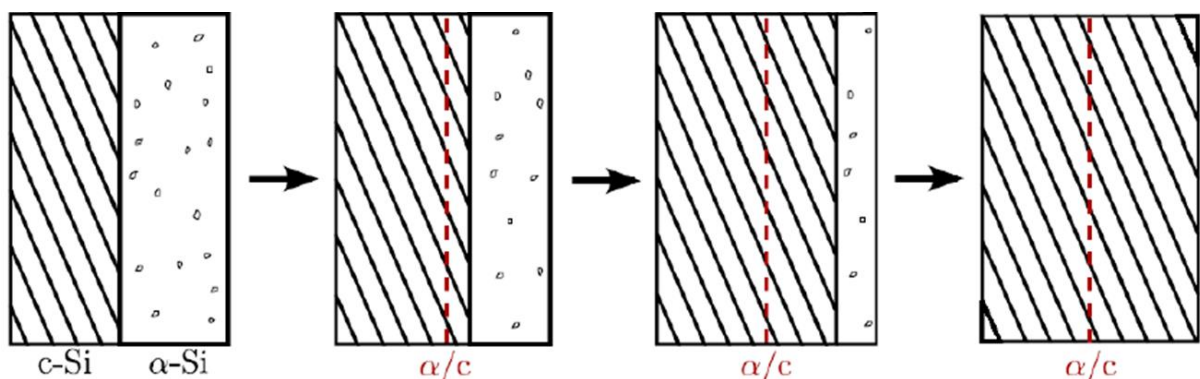


Figure 1.23 : Principe de la recristallisation en phase solide d'une couche amorphe à partir d'un seed cristallin (pris de [Pasini16a]).

Lors de cette recristallisation, les atomes se réorganisent et les dopants intègrent la matrice de silicium en position substitutionnelle d'où ils peuvent jouer leur rôle électrique (donner ou prendre un électron). L'intérêt majeur de la technique SPER est qu'une température plus faible (e.g. $T < 600^\circ\text{C}$) peut être utilisée afin de recristalliser la couche amorphisée et donc activer les dopants. Par ailleurs, un autre avantage est que la faible température limite la diffusion des dopants et permet d'obtenir des jonctions plus abruptes qu'un recuit à haute température. Cependant, plusieurs inconvénients apparaissent aussi avec l'utilisation de la SPER.

- Limite de clusterisation

Tout d'abord, comme indiqué précédemment, le recuit à haute température permet non seulement d'activer les dopants, mais aussi de guérir les défauts créés par l'implantation ionique. Outre l'amorphisation du silicium, un phénomène de clusterisation peut également apparaître où, au-dessus d'une concentration limite, les dopants tendent à former un agglomérat électriquement non actif qui dégrade la résistance de la couche dopée [Pasini14]. Cette concentration limite de clusterisation a été estimée pour une SPER à 600°C pour l'arsenic ($[\text{As}]_{600^\circ\text{C}} = 8 \times 10^{20} \text{ at/cm}^3$), le phosphore ($[\text{P}]_{600^\circ\text{C}} = 6 \times 10^{20} \text{ at/cm}^3$) et le bore ($[\text{B}]_{600^\circ\text{C}} = 3 \times 10^{20} \text{ at/cm}^3$). Ainsi, afin d'optimiser l'activation des dopants et donc minimiser la résistance électrique, une implantation uniforme sur toute la couche et contrôlée afin de rester en dessous de la limite de clusterisation est préconisée (Figure 1.24a).

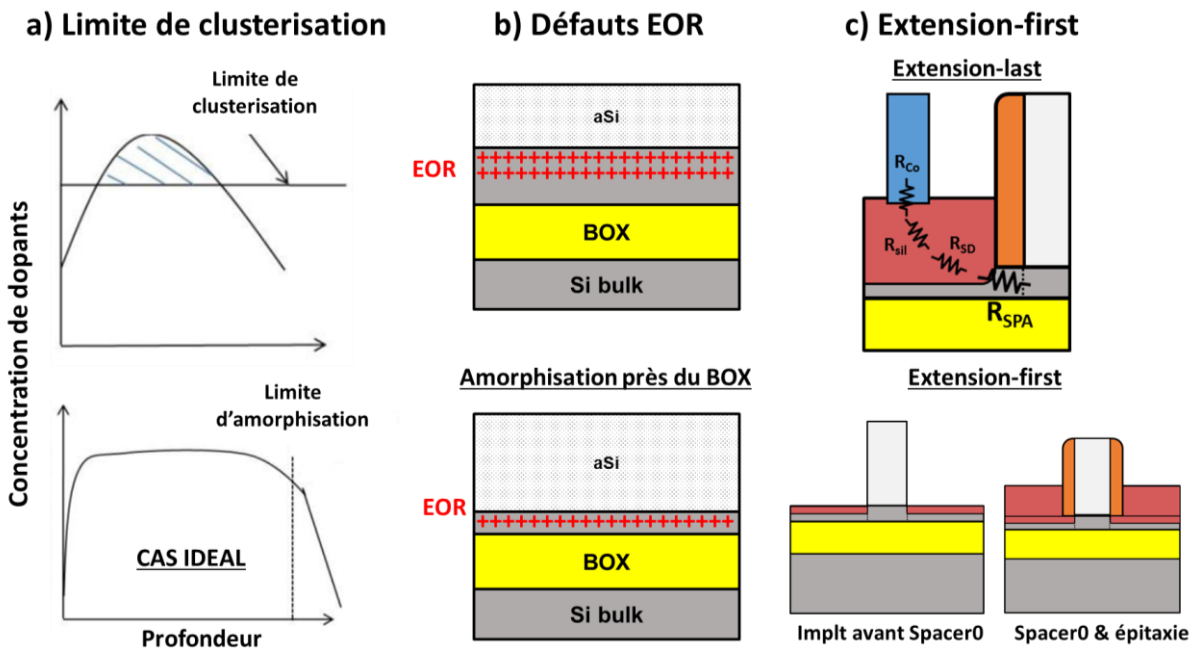


Figure 1.24 : Problématiques liées à une activation à basse température : a) limite de clusterisation (pris de [Pasini14]), b) défauts end-of-range et c) dopage sous les espaces de grille.

Dans la pratique, une seule étape d'implantation ionique est alors réalisée directement après surélévation des sources et drains (équivalent à l'implantation LDD), l'étape d'implantation HDD étant supprimée pour ne pas incorporer de dopants au-dessus de la concentration limite. Du côté des nMOS, une implantation de phosphore est préférée à l'arsenic malgré une limite de clusterisation plus basse essentiellement pour sa masse atomique plus faible qui permet d'une part d'obtenir une mobilité de porteurs plus élevée, et d'autre part pour mieux contrôler le profil de dopage sans amorphiser complètement les accès [Pasini15]. Du côté des pMOS, une triple implantation de bore est effectuée pour obtenir un profil de dopage plat. Le bore étant un atome léger et petit, une étape de pré-amorphisation par germanium est nécessaire pour amorphiser les accès. Les sources et drains (d'environ 30nm après surélévation) sont alors totalement recristallisées avec un recuit à 600°C pendant 2 minutes, budget thermique compatible avec l'intégration 3D Séquentielle. La température

peut être réduite moyennant un temps suffisant long pour recristalliser totalement la couche, mais d'autres problèmes liés par exemple à la présence de l'oxyde natif sur le silicium rentrent alors en jeu [Magee81].

- Défauts *end-of-range*

Par ailleurs, l'amorphisation par implantation ionique introduit également des défauts interstitiels sous l'interface amorphe/cristal qui ne sont pas totalement dissous à basse température. Lors d'un recuit, ces défauts dits *end-of-range* (EOR) peuvent interagir avec les atomes dopants et les désactiver électriquement, diminuant alors la résistance des accès. Une solution possible est de confiner ces défauts EOR en amorphisant plus près de l'oxyde enterré (Figure 1.24b) [Pawlak04]. Néanmoins, une épaisseur minimale de *seed* cristallin (~3nm) reste nécessaire afin de pouvoir correctement recristalliser la couche amorphe [Pasini15].

- Dopage sous les espaceurs de grille

Une dernière problématique importante de l'activation de dopants à basse température est la zone effectivement activée électriquement. Comme nous l'avons dit précédemment, la faible température ne permet pas de diffuser les dopants. En conséquence, seuls les dopants présents dans la couche initialement amorphisée sont activés lors de la SPER. Une première difficulté est alors de contrôler précisément la profondeur d'amorphisation lors de l'implantation ionique afin d'amorphiser une zone importante des sources et drains pour améliorer la résistance des accès tout en laissant un *seed* cristallin suffisant. Une autre difficulté est d'amorphiser la zone sous les espaceurs de grille à travers les sources et drains surélevés. Des simulations montrent en effet que si cette zone peut être amorphisée par exemple en augmentant la dose et/ou l'énergie d'implantation ou en inclinant l'angle de pénétration des ions, ces conditions mènent également à une amorphisation complète des sources et drains. Par conséquent, dans une intégration telle que présentée auparavant dans cette section (que l'on nommera « extension-last »), la zone sous les espaceurs de grille est très peu dopée et devient non négligeable dans la résistance totale des accès qui comprend également la résistance des sources et drains (R_{SD}), du silicium (R_{sil}) et des contacts (R_{co}). Une solution à ce problème est alors d'effectuer l'implantation de dopants avant la surélévation des sources et drains (intégration dite « extension-first ») voire avant la formation des espaceurs de grille, permettant de placer des dopants près du canal (Figure 1.24c) [Pasini16b]. Des optimisations sont alors nécessaires par exemple pour contrôler la zone amorphisée dans une couche de silicium mince (6 à 12nm) ou pour ne pas introduire de dopants parasites dans la grille ou l'endommager.

Pour résumer, la technique SPER permet donc d'activer efficacement les dopants dans une zone préalablement amorphisée et doit être optimisée en termes de profil de dopage et de profondeur d'amorphisation. Par ailleurs, le passage à une intégration dite *extension-first* où une première implantation de dopants est réalisée avant la surélévation des sources et drains voire avant la formation des espaceurs de grille semble obligatoire afin de s'affranchir d'une résistance élevée sous les espaceurs de grille. Cette dernière intégration a notamment été démontrée dans [Pasini16] sur une technologie 14FDSOI de STMicroelectronics permettant d'obtenir des performances proches (90-95%) d'une référence à haute température sur les nMOS (Figure 1.25a) et sur les pMOS (Figure 1.25b). A noter qu'un jeu d'espaceurs de grille très fin (3nm) a tout de même été déposé avant l'implantation *extension-first* afin de ne pas endommager l'empilement de grille. Aussi, des variantes de pré-amorphisation par germanium (PAI – *PreAmorphization Implant*) ont été utilisées pour faire varier la profondeur d'amorphisation.

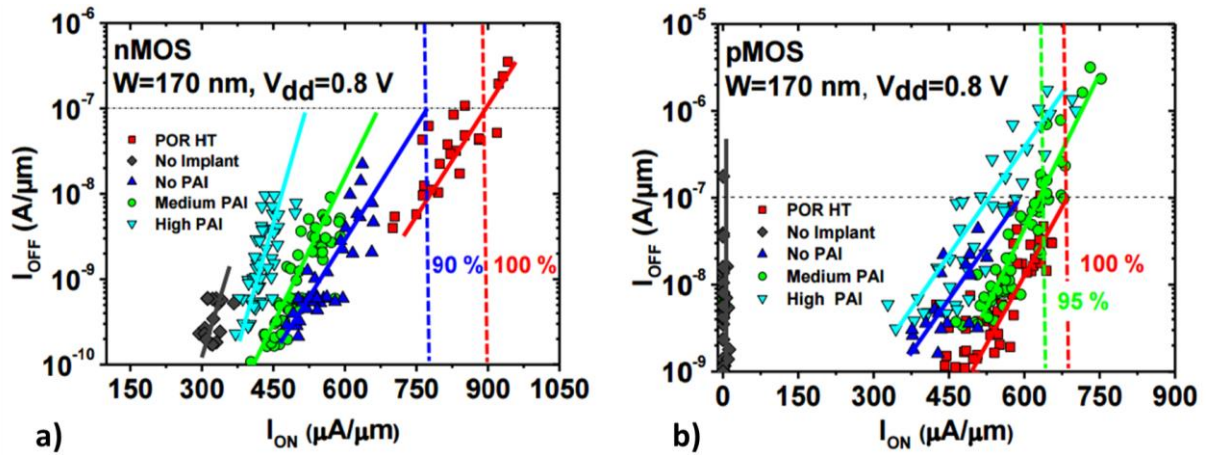


Figure 1.25 : Performances de transistors FDSOI activés à basse température avec une intégration extension-first pour a) des nMOS (basse température en bleu) et b) des pMOS (basse température en vert) (pris de [Pasini16]).

Activation par recuit laser

D’autre part, comme nous l’avons vu dans la Section 1.2, le recuit laser, et plus particulièrement celui nanoseconde, permet d’apporter une énergie très élevée permettant de chauffer à très haute température la surface de la plaque sur un temps extrêmement court (de l’ordre de la centaine de nanosecondes), temps qui ne permet pas de diffuser totalement l’énergie thermique jusqu’aux transistors du niveau bas [Fenouillet14b]. Ce recuit a également été étudié pour l’activation de dopants, par exemple le laser UV ($\lambda=308\text{nm}$) pour l’activation d’arsenic [Kerdiles16], du phosphore [Acosta16] et du bore [Fenouillet14b] sur pleine plaque. Le principe de l’activation par recuit laser est sensiblement similaire à la technique SPER dans la mesure où elle repose également sur la recristallisation d’une couche amorphisée par l’implantation de dopants. La Figure 1.26 montre alors que pour certaines densités d’énergie du faisceau laser, une résistance carrée (i.e. le niveau d’activation de dopants) équivalente à un recuit rapide à haute température (RTP – *Rapid Thermal Processing*) de $1050^\circ\text{C} \sim 2\text{s}$ est obtenue.

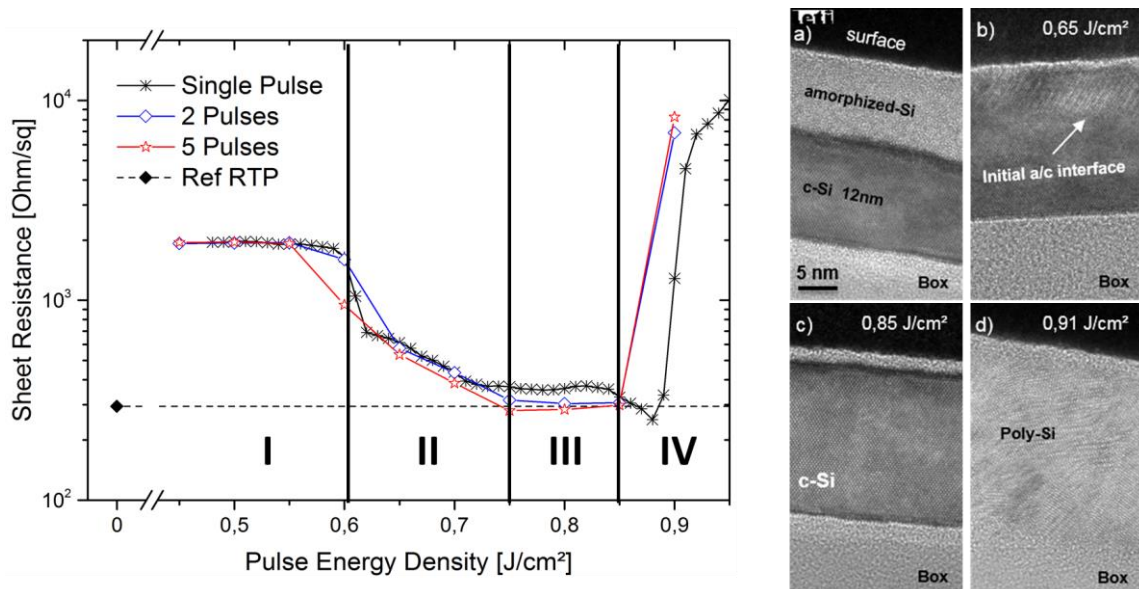


Figure 1.26 : Résistance carrée d’une couche de silicium amorphisée par implantation de phosphore en fonction de la densité d’énergie du recuit laser et observations TEM associées (pris de [Acosta16]).

On remarque notamment 4 régions. La région I ($J < 0.61\text{J}/\text{cm}^2$) correspond à une densité d’énergie trop faible pour activer les dopants, menant à aucune variation de la résistance carrée. Cette dernière commence à diminuer dans la région II ($0.61 < J < 0.75\text{J}/\text{cm}^2$) et correspond en réalité à un début de fonte

de la couche amorphe, signifiant que la température de fusion ($T \sim 1100^\circ\text{C}$) a été atteinte. Cependant, lorsque la couche liquide atteint l'interface avec le *seed* cristallin, une recristallisation dite explosive se produit menant à la formation d'une couche polycristalline et défectueuse comme le montre l'observation TEM (b). Au contraire, une densité d'énergie située dans la région III ($0.75 < J < 0.85 \text{ J/cm}^2$) mène à une couche parfaitement cristalline et à une résistance carrée minimale. Une explication possible est que l'énergie thermique supérieure apportée par une densité d'énergie plus élevée permet de fondre également la couche polycristalline obtenue par recristallisation explosive puis de recroître une couche de meilleure qualité cristalline. Enfin, lorsque la densité d'énergie atteint la région IV ($J > 0.85 \text{ J/cm}^2$), le *seed* cristallin est également fondu (soit $T \sim 1400^\circ\text{C}$) et mène à la formation d'une couche totalement polycristalline et fortement résistive [Acosta16].

Ainsi, le recuit laser nanosecondes est également une technique prometteuse pour l'activation de dopants afin de limiter l'endommagement des composants inférieurs. En revanche, le procédé doit être optimisé pour contrôler correctement la fonte et la recristallisation de la couche amorphisée. De plus, comme nous l'avons mentionné auparavant, la présence d'une topographie peut éventuellement poser des problèmes d'uniformité. Le recuit laser nanosecondes n'a néanmoins pas encore été démontré électriquement sur des transistors.

Activation par recuit micro-ondes

Un autre type de recuit alternatif au recuit four est le recuit par micro-ondes. Tout comme le recuit laser, le recuit micro-onde est basé sur l'émission d'ondes électromagnétiques mais avec des longueurs d'ondes beaucoup plus élevées (10^{-1} à 10^2 cm). Trois effets régissent alors le comportement d'un matériau sous l'effet des ondes électromagnétiques [Clark00].

- Effet Joule ou échauffement ohmique

Lorsque des particules chargées (i.e. les électrons dans notre cas) sont en présence d'un champ électrique, ces dernières se mettent en mouvement. Dans le cas particulier d'une onde, le champ électrique oscillant rapidement fait s'entrechoquer les électrons, créant de l'énergie thermique par effet Joule. Une condition importante à l'échauffement d'un matériau est alors la présence d'électrons qui ne sont pas liés à un atome et qui sont donc libres de se déplacer. La capacité d'un matériau à s'échauffer est donc directement liée à sa conductivité.

- Effet de relaxation

Sous l'effet d'un champ électromagnétique, les molécules d'un matériau peuvent se polariser et créer un dipôle. Lorsque le champ est oscillant, ce dipôle se met à tourner suivant l'orientation du champ et peut générer de la chaleur par friction ce qui ralentit le mouvement du dipôle (d'où le terme de relaxation). Cependant, pour les fréquences ($f < 10 \text{ GHz}$) et les matériaux généralement utilisés pour la micro-électronique, cet effet est très faible car la rotation des dipôles reste relativement lente.

- Effet de peau

Comme nous l'avons vu précédemment, le champ électrique oscillant des micro-ondes génère un courant électrique alternatif dans un matériau conducteur. Ce courant alternatif génère alors un champ magnétique alternatif qui à son tour génère un nouveau courant alternatif. La particularité de ce dernier courant est qu'il va dans le sens inverse du courant initial et compense ce dernier au centre du matériau, tandis que les deux courants s'additionnent en surface. Au final, l'effet Joule induit par le mouvement des électrons n'apparaît qu'en surface du matériau conducteur et donc seule cette zone est fortement chauffée. Cet effet dit de peau est alors d'autant plus important que la conductivité du matériau est élevée.

Ainsi, dans les conditions de recuits micro-ondes utilisées, seuls les effets Joule et de peau rentrent en jeu. Ceci signifie notamment que les isolants (faible conductivité) ne chauffent pas par effet Joule et

sont transparents aux micro-ondes, tandis que les métaux (forte conductivité) chauffent fortement mais essentiellement en surface et sont quasiment opaques aux micro-ondes. Le cas intermédiaire des semi-conducteurs est alors un compromis entre l'absorption des micro-ondes et leur transformation en énergie thermique. Aussi, une particularité du recuit micro-ondes est que celui-ci se fait de manière volumique, au contraire des recuits four ou laser qui sont surfaciques. Le recuit micro-ondes a donc également été démontré pour l'activation de dopants, par exemple pour l'arsenic et le bore (Figure 1.27) [Lee14].

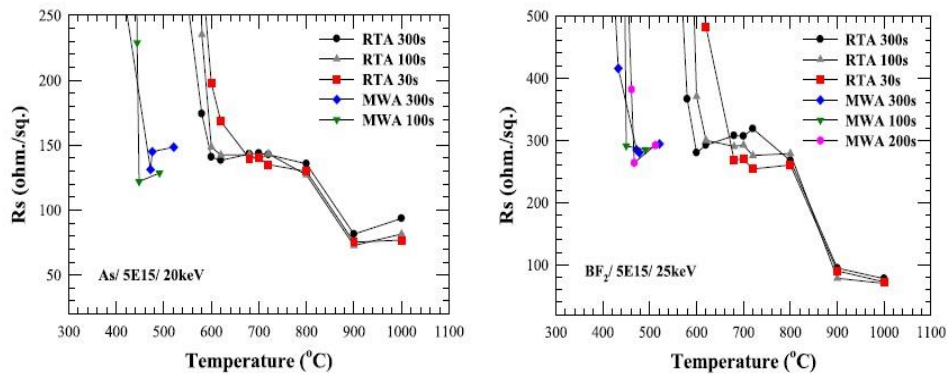


Figure 1.27 : Résistance carrée d'une couche de silicium amorphisée par implantation d'arsenic (à gauche) ou de BF₂ (à droite) en fonction de la température pour des recuits four ou micro-ondes (pris de [Lee14]).

Sur ces résultats, on remarque tout d'abord la présence de trois régions pour le recuit four (RTA – *Rapid Thermal Anneal*) contre deux pour le recuit micro-ondes (MWA – *MicroWave Anneal*). La troisième région apparaissant pour les températures élevées ($T > 800^\circ\text{C}$) dans le premier cas est due à la diffusion des dopants, augmentant de fait le niveau d'activation total ce qui diminue donc la résistance carrée. Le recuit micro-ondes étant sur une gamme de température beaucoup plus faible ($T < 600^\circ\text{C}$ ici), aucune diffusion ne se produit permettant d'obtenir des jonctions plus abruptes. Une deuxième observation est qu'une résistance carrée équivalente est obtenue avec le recuit micro-ondes pour des températures plus faibles d'environ 150°C , indiquant que des effets athermiques (i.e. non liés à la température) peuvent se produire avec les micro-ondes. Une première hypothèse est que l'énergie électromagnétique induit une force cinétique sur les dopants qui peut leur permettre de se déplacer en position substitutionnelle (effet de *ion hopping*) et donc être activés électriquement [Lee14]. Une autre hypothèse avancée par DSG Technologies (équipementier fabricant des fours micro-ondes pour l'électronique) est un effet d'échauffement sélectif et local des défauts. Pour faire simple, lorsque les électrons se déplacent dans le réseau d'un matériau sous l'effet du champ électrique, ceux-ci peuvent s'accumuler préférentiellement au niveau de défauts (dopants, dislocations, liaisons pendantes...). La densité plus élevée de porteurs mobiles sur ces zones amplifie alors l'effet Joule et chauffe donc de manière locale et sélective les défauts à très haute température. En conséquent, les dopants en position interstitiel sont mieux activés grâce à cette énergie thermique supplémentaire. Une dernière hypothèse possible et plus simple serait un défaut de mesure de la température, les ondes électromagnétiques pouvant interférer avec la mesure des pyromètres. L'écart de température (150°C) similaire obtenu quel que soit le dopant utilisé pourrait aller dans ce sens.

Plusieurs solutions existent donc afin d'activer les dopants à bas budget thermique, raison pour laquelle cette problématique ne sera pas abordée plus en détails ici. Dans ce manuscrit, nous parlerons donc des briques d'espaceurs de grille, de surélévation des sources et drains par épitaxie et de l'empilement de grille. Nous verrons notamment que le développement de ces briques à bas budget thermique passe par la recherche de nouveaux matériaux, du développement de nouveaux procédés et de la (tentative de) compréhension des phénomènes physiques mis en jeu.

Dans le Chapitre 2, nous verrons que le développement d'une brique espaceurs de grille à basse température passe par une étude de différents diélectriques alternatifs au SiN de référence à 630°C. En particulier, l'intérêt des diélectriques à faible permittivité (dits low-k) sera soulevé, ceux-ci permettant de réduire les capacités parasites et de fait d'améliorer les performances dynamiques des circuits. Une caractérisation des propriétés matériaux, morphologiques et électriques de chaque diélectrique candidat sera effectuée, ceux-ci devant répondre à des critères précis afin d'être intégrables.

La surélévation des sources et drains sera ensuite étudiée dans le Chapitre 3. La question de la préparation de surface avant épitaxie sera abordée et l'utilisation d'un nettoyage par plasma démontré afin d'obtenir une surface idéale pour l'épitaxie à basse température. Nous verrons ensuite que la croissance épitaxiale d'une couche cristalline à basse température passe d'une part par un changement de stratégie pour le procédé, et d'autre part par une optimisation de plusieurs paramètres de l'épitaxie.

En outre, les propriétés de la grille sont également directement impactées par le budget thermique (brique espaceurs, épitaxie et activation de dopants) vu par l'empilement. **Nous verrons donc dans le Chapitre 4 l'impact des budgets thermiques subséquents sur les propriétés de l'empilement de grille.** Nous présenterons également différents leviers technologiques à bas budget thermique pour modifier les propriétés de la grille selon l'application recherchée. Enfin, nous étudierons plus particulièrement la fiabilité de la grille qui, nous le verrons, est une difficulté majeure pour une intégration à bas budget thermique.

Enfin, **nous concluons le manuscrit dans le Chapitre 5** en résumant les résultats importants obtenus dans cette thèse. Nous verrons également les applications potentielles de l'intégration 3D Séquentielle (en particulier pour les mémoires), mais également des avantages possibles au développement de briques technologiques à bas budget thermique.

Chapitre 2 – Matériaux pour les espaceurs de grille

Introduction au Chapitre 2

Les espaceurs sont des couches diélectriques minces servant à isoler électriquement deux électrodes ou à protéger certaines zones ou matériaux lors de la fabrication du transistor. Comme nous l’avons vu dans le Chapitre 1, trois jeux d’espaceurs sont utilisés dans l’intégration planaire sur FDSOI :

- Le *spacer0* entre la grille et les sources et drains (aussi appelés espaceurs de grille).
- Le *spacer1* qui sert à protéger les espaceurs de grille lors du retrait masque dur et qui est ensuite totalement retiré.
- Le *spacer2* qui permet de contrôler la zone de siliciuration et d’éviter qu’elle ne se forme trop près du canal.

Dans ce chapitre, nous nous concentrerons uniquement sur les espaceurs de grille, les seuls qui ne soient pas compatibles avec la 3D Séquentielle du fait de la température de dépôt.

Les espaceurs de grilles ont plusieurs rôles à tenir dans la fabrication du transistor ainsi que sur ses propriétés électriques. Tout d’abord, la présence des espaceurs est essentielle afin de protéger l’empilement de grille lors des étapes subséquentes, par exemple durant les nettoyages chimiques ou lors de l’épithaxie des sources et drains. L’absence ou la mauvaise étanchéité des espaceurs peut alors mener à la consommation d’une ou plusieurs couches de la grille par les chimies humides (Figure 2.1), ou bien à une épithaxie parasite sur les flancs qui causeraient un court-circuit entre les sources et drains et la grille comme nous le verrons plus tard.

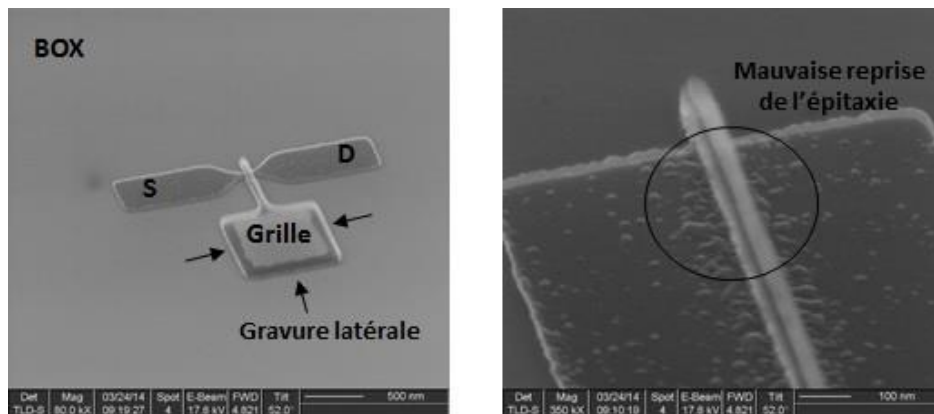


Figure 2.1 : Observation SEM après l’étape de surélévation des sources et drains par épithaxie avec des espaceurs $\text{SiN}_{500^\circ\text{C}}$. La mauvaise étanchéité des espaceurs a mené à une gravure latérale de l’empilement de grille lors des nettoyages humides. Par ailleurs, on observe également une mauvaise reprise d’épithaxie sur les zones actives avec l’apparition d’îlots.

En outre, cette couche permet également d’encapsuler les matériaux de l’empilement de grille, empêchant notamment une contamination des équipements par le métal TiN ou bien un phénomène de « *oxygen ingress* », c’est-à-dire une suroxydation indésirée du diélectrique de grille par les flancs de l’empilement. Aussi, les espaceurs influencent le profil des jonctions lors de l’implantation ionique des dopants en masquant une partie du canal. Cet effet est particulièrement important dans une intégration froide où la température ne permet pas la diffusion des dopants et où la position de la jonction est donc directement déterminée par l’épaisseur des espaceurs [Pasini15]. Comme nous l’avons vu dans le Chapitre 1, une intégration dite « extension-first » où les dopants sont implantés

avant le dépôt des espaceurs permet néanmoins de contourner ce problème. Enfin, combiné avec le *spacer2*, les espaceurs de grille permettent de séparer la siliciuration de la grille et des sources et drains ainsi que leurs contacts.

Le choix du matériau diélectrique en tant qu'espaceurs est donc crucial et doit alors répondre à plusieurs critères.

Propriétés matériau :

- Afin de tenir son rôle d'isolant, le matériau doit être peu fuitueux et doit avoir une tension de claquage (i.e. tension nécessaire pour rendre le diélectrique conducteur) élevée pour éviter des pertes de courant entre la grille et les sources et drains.
- Le diélectrique doit également conserver ses caractéristiques tout au long de l'intégration, notamment lorsqu'il est soumis aux budgets thermiques des étapes suivantes. Dans notre travail, un budget thermique de 600°C 15min a été retenu (correspondant approximativement aux budgets thermiques des étapes de surélévation des sources et drains et d'activation des dopants [Fenouillet14a]), ce dernier étant celui le plus faible de l'intégration lors du démarrage de ces études.

Dépôt :

- L'épaisseur déposée doit être uniforme sur toute la plaque (d'un diamètre de 300mm dans notre cas), ce afin d'éviter une variabilité sur les propriétés électriques des dispositifs finaux. A noter que l'uniformité du dépôt est surtout une considération industrielle voire une problématique liée à l'équipement ou la technique de dépôt utilisée, et donc à l'échelle de notre étude, ce critère n'est pas forcément primordial. L'amélioration de l'uniformité passe alors par un bon contrôle de l'injection des précurseurs et de l'optimisation de leurs pressions partielles, de la disposition de la chambre (e.g. distance entre la plaque et l'injection de gaz), etc.
- Le dépôt doit également être conforme, c'est-à-dire que l'épaisseur déposée sur les flancs de la grille (verticalement) doit être la même que celle déposée sur le haut de la grille et à la surface de la plaque (horizontalement). Par ailleurs, la composition chimique du matériau doit également être la même quelle que soit la topographie, sans quoi des zones peuvent être consommées plus ou moins rapidement, menant à une mauvaise protection de la grille à certains endroits.
- Un dernier paramètre à prendre en compte est l'effet de densité des structures (*pattern loading*). Un dépôt sur un dispositif isolé peut en effet ne pas être identique à un dépôt sur une zone dense (i.e. sur des grilles proches les unes des autres). Cependant, le masque utilisé pour cette étude (aux règles de dessin 65nm) ne dispose pas de motifs assez denses pour étudier ce paramètre.

Intégrabilité :

- La couche doit pouvoir être gravée de manière anisotropique avec un profil droit pour pouvoir contrôler correctement la position des dopants. La consommation latérale doit également être minimale pour la même raison.
- Le film de silicium après gravure doit permettre une reprise d'épitaxie pour l'étape de surélévation des sources et drains. La Figure 2.1 illustre par exemple les problèmes d'épitaxie liés à une mauvaise surface initiale (possible présence de résidus de gravure).
- Enfin, la couche diélectrique doit résister à toutes les chimies humides utilisées lors de l'intégration, notamment lors des étapes de retrait polymères après gravure, nettoyage avant épitaxie ou bien stripping de la résine après photolithographie.

Un dernier critère propre à la 3D Séquentielle est la température de dépôt du diélectrique qui doit être inférieure à 500°C. Le matériau standard qui sera notre référence dans cette étude est un nitrure de silicium déposé à 630°C (**SiN_{630°C}**) par ALD (*Atomic Layer Deposition*). Deux autres nitrures de silicium déposés à plus basse température ont été étudiés afin de le remplacer :

- **SiN_{500°C}** : déposé à 500°C par ALD dans le même équipement que le diélectrique de référence.
- **SiN_{400°C}** : déposé à 400°C par PEALD (*Plasma Enhanced ALD*) spatiale (nommé par la suite sPEALD). Ce dépôt a été fait en démo chez l'équipementier Applied Materials (Santa Clara, USA).

En parallèle, les matériaux à faible permittivité (dits low-k) apparaissent également très prometteurs pour les applications espaceurs. En effet, l'utilisation de diélectriques introduit systématiquement des capacités proportionnelles à leur permittivité, augmentant alors les temps de propagation des signaux dans un circuit et donc réduisant la fréquence maximale d'utilisation des dispositifs (Figure 2.2).

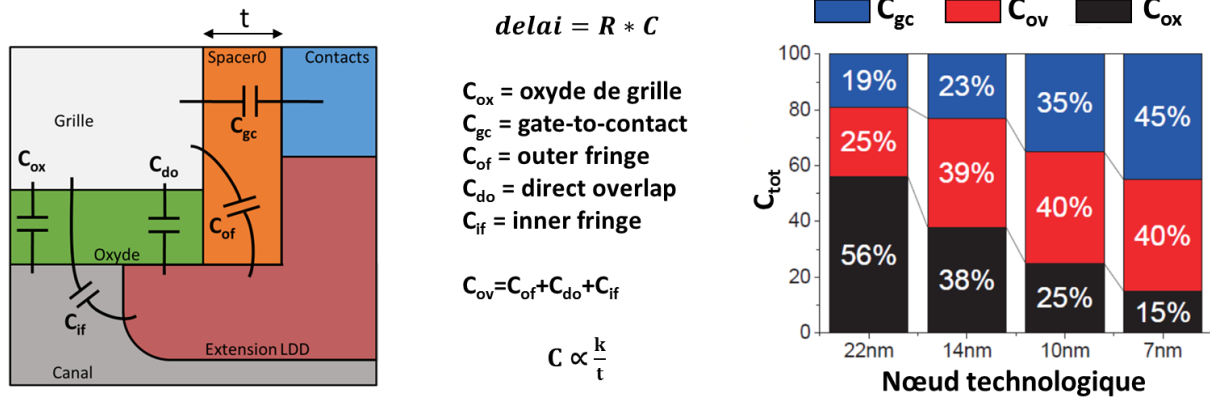


Figure 2.2 : Schéma récapitulatif des différentes capacités liées à l'oxyde de grille et au diélectrique d'espaceurs et projection de leur contribution sur la capacité totale en fonction du nœud technologique (Résultats pris de [Yamashita15]). Les capacités parasites liées aux espaceurs deviennent prépondérantes avec la miniaturisation des transistors, justifiant l'introduction des diélectriques low-k.

Ainsi, les espaceurs de grille introduisent des capacités entre la grille et les extensions LDD (capacité *outer fringe* – C_{of}) et entre la grille et les contacts des sources et drains (capacité *gate-to-contact* C_{gc}). Ces capacités parasites, avec les capacités induites par l'oxyde de grille (C_{ox} , C_{if} et C_{do}), participent alors à la capacité totale (C_{tot}) mesurée sur un dispositif. Cependant, la contribution des capacités parasites liées aux espaceurs (C_{gc} et C_{ov}) tendent à devenir plus importantes avec la miniaturisation des transistors du fait de la réduction de leur épaisseur. De fait, outre la disponibilité de matériaux low-k déposables à basse température, ces diélectriques ont également un intérêt pour les nœuds technologiques avancés où leur utilisation permet d'améliorer les caractéristiques dynamiques des dispositifs. Ainsi, l'utilisation des low-k a déjà été démontrée mais ceux-ci ne sont pas forcément compatibles avec notre problématique, soit à cause de leur température de dépôt trop élevée, soit par des propriétés matériaux non satisfaisantes. Par exemple, les matériaux poreux comme le SiCOH ($k=3$) présentent des permittivités très faibles, mais au détriment d'une sensibilité plus élevée aux budgets thermiques ainsi qu'aux chimies humides [Huang08]. Son utilisation a donc nécessité plusieurs changements dans l'intégration afin qu'il ne soit soumis à aucun recuit ou nettoyage. Ces matériaux sont en revanche largement utilisés dans le back-end où la température ne dépasse pas 400°C afin de diminuer les couplages électromagnétiques entre les lignes métalliques [Hatton06]. Par ailleurs, des SiCBN déposés par CVD [Ko08] ou par ALD [Yamashita15, Weber15b] avec des permittivités de 5.0-5.2 ont également été intégrés pour la diminution des capacités parasites. Néanmoins, leurs températures de dépôt (550°C et 600°C respectivement) ne conviennent pas à notre problématique initiale.

Dans le cadre de la 3D Séquentielle, de nouveaux matériaux low-k déposés à basse température ont donc également été étudiés :

- **SiOCN_{350°C}** et **SiOCN_{480°C}** : déposés à 350°C et 480°C respectivement par PEALD en alternant des couches de SiO₂ et de SiCN [Lu16a]. Ces dépôts ont aussi été faits en démo chez Applied Materials.
- **SiCO_{400°C}** : déposé à 400°C par PECVD (*Plasma Enhanced Chemical Vapor Deposition*) [Benoit15]. Ce dépôt a été fait en démo chez l'équipementier LAM Research (Santa Clara, USA).
- **SiCBN_{CVD}** et **SiCBN_{PECVD}** : déposés à 480°C par CVD [Lu15] et PECVD respectivement. Ces dépôts ont été faits au Leti.

Les différents matériaux étudiés sont résumés dans la Table 2.1.

Table 2.1 : Récapitulatif des diélectriques étudiés

	SiN			SiOCN		SiCO	SiCBN	
Température	630°C	500°C	400°C	350°C	480°C	400°C	480°C	480°C
Technique dépôt	ALD	ALD	sPEALD	PEALD	PEALD	PECVD	CVD	PECVD
Location	STM	STM	AMAT	AMAT	AMAT	LAM	Leti	Leti

(Ref)

Dans ce chapitre, nous nous focaliserons tout d'abord sur les propriétés matériau des différents diélectriques étudiés sur pleine plaque dans la Section 2.1, où nous verrons notamment les effets de la température et de la technique de dépôt, avant d'étudier la morphologie de ces dépôts sur des structures patternées dans la Section 2.2. Enfin, nous verrons dans la Section 2.3 l'influence des espaceurs de grille sur les propriétés électriques statiques et dynamiques de transistors en prenant les cas du $\text{SiOCN}_{480^\circ\text{C}}$ et du $\text{SiCO}_{400^\circ\text{C}}$.

2.1 Étude des matériaux

2.1.1 Caractérisation des matériaux

Dans un premier temps, les propriétés des différents matériaux ont été étudiées sur pleine plaque. Après avoir déposé une couche du diélectrique sur du silicium bulk, différentes caractérisations ont été effectuées.

- Ellipsométrie :

L'ellipsométrie est une technique d'analyse optique consistant à envoyer un faisceau lumineux monochromatique (i.e. une seule longueur d'onde) et polarisé linéairement sur la surface d'un échantillon et d'en récupérer le faisceau réfléchi (Figure 2.3).

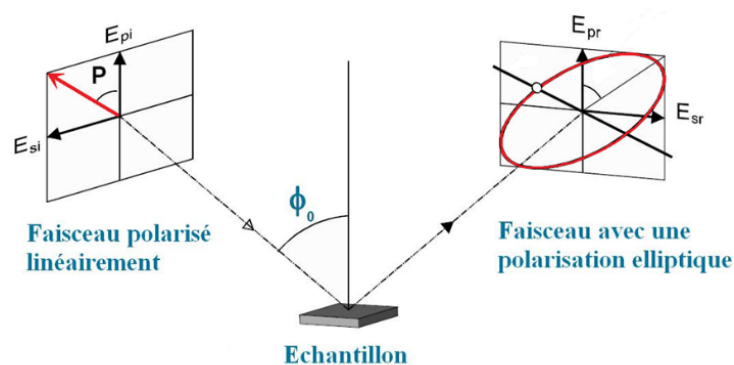


Figure 2.3 : Principe de la mesure ellipsométrique. Un faisceau monochromatique et polarisé linéairement est envoyé sur l'échantillon (angle Φ_0 entre 55° et 75° dans notre cas). Le faisceau réfléchi est alors détecté et comparé à un modèle mathématique afin d'extraire les constantes optiques et l'épaisseur de l'échantillon.

L'analyse du faisceau réfléchi dont la polarisation est modifiée selon les couches de l'empilement (sous forme ellipsoïdale, d'où le nom de la technique) permet alors d'extraire les indices optiques (**indice de réfraction** et coefficient d'extinction) ainsi que l'**épaisseur** du diélectrique grâce à un modèle mathématique. Une variante est l'ellipsométrie spectroscopique qui effectue des mesures sur un

spectre large de longueur d'ondes (~200-1000nm), ce qui permet notamment de mesurer des empilements de plusieurs couches. En outre, la mesure ellipsométrique sur toute la plaque nous permettra d'évaluer l'**uniformité** des dépôts.

- Réflectométrie de rayons X (XRR) :

La mesure XRR est une technique également basée sur l'envoi d'un faisceau monochromatique sur la surface et la collecte du faisceau réfléchi. Cependant, dans ce cas-ci, l'énergie du faisceau est beaucoup plus élevée (rayons X) et envoyée avec un angle rasant par rapport à la surface (entre 0.25° et 3.75°). Lorsqu'un angle critique est dépassé, une partie du faisceau pénètre dans le diélectrique et forme des interférences constructives ou destructives avec l'onde incidente lorsqu'il en ressort qui modifient alors l'intensité du faisceau collecté. L'analyse de l'intensité en fonction de l'angle d'incidence des rayons X permet alors de remonter entre autre à la **densité massique** et à l'**épaisseur** du diélectrique (Figure 2.4).

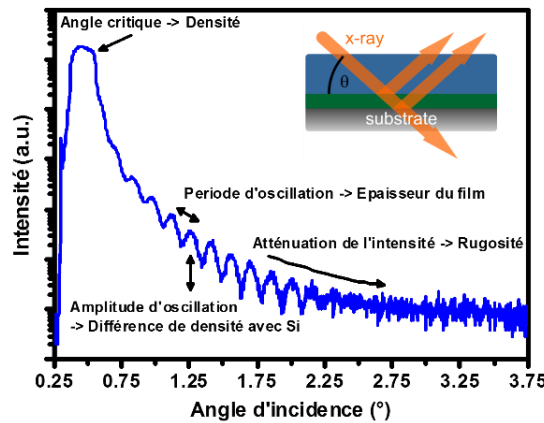


Figure 2.4 : Exemple de spectre XRR d'une couche de 30nm de SiCBN_{PECVD}. L'angle critique, la période et l'amplitude des oscillations et l'atténuation de l'intensité en fonction de l'angle d'incidence des rayons X permet d'extraire la densité et l'épaisseur du film ainsi que la différence de densité et la rugosité avec le silicium bulk respectivement.

- Mesure de flèche

La flèche correspond à la courbure d'une plaque sous l'effet du poids ou des couches présentes sur le silicium lorsqu'elle n'est tenue que par ses bords. A partir de la distance mesurée entre un détecteur et le milieu de la plaque mesurée des deux faces de la plaque (e.g. par interférométrie), on peut remonter à la flèche f de la plaque ainsi qu'à son rayon de courbure R (Figure 2.5).

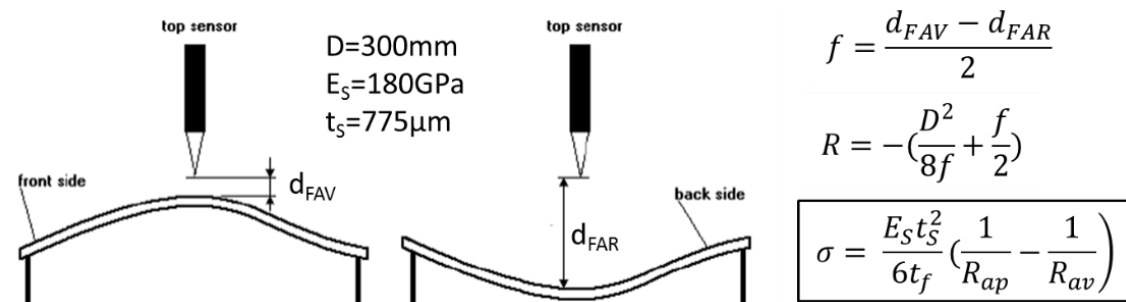


Figure 2.5 : Principe de la mesure de flèche et équations provenant du modèle de Stoney permettant de remonter à la contrainte induite par la couche mince.

Ces valeurs permettent alors de calculer la **contrainte** σ (stress) induite par la couche diélectrique sur le substrat, celle-ci modifiant la courbure de la plaque selon si elle est tensile (contrainte positive) ou compressive (négative). Pour cela, nous utiliserons l'équation :

$$\sigma = \frac{E_s t_s^2}{6 t_f} \left(\frac{1}{R_{ap}} - \frac{1}{R_{av}}\right)$$

Où E_s et t_s sont le module de Young et l'épaisseur du substrat silicium, t_f l'épaisseur du film diélectrique déposé et R_{av} et R_{ap} les rayons de courbures avant et après dépôt de la couche. Cette formule simplifiée du modèle de Stoney est valable pour des films très minces et nous ferons l'hypothèse qu'ils sont également homogènes sur toute la surface [Ardigo14].

- Spectroscopie infrarouge :

La spectroscopie infrarouge (IR) repose aussi sur l'envoi d'un faisceau de lumière en balayant sur un spectre de longueur d'onde dans l'infrarouge (2.5 à 25 μ m). Selon la longueur d'onde (i.e. inverse de l'énergie), les liaisons atomiques ou moléculaires de l'échantillon peuvent absorber l'énergie, être excités et rentrer en résonance. L'analyse de l'énergie absorbée par (ou transmise à travers) l'échantillon en fonction du nombre d'onde (i.e. l'inverse de la longueur d'onde) permet alors d'évaluer les **liaisons chimiques** présentes dans le diélectrique (Figure 2.6). Plusieurs sortes de spectroscopie infrarouge existent, mais nous utiliserons essentiellement la spectroscopie FTIR (*Fourier Transform InfraRed*), technique basée sur l'acquisition simultanée de mesures sur plusieurs longueurs d'onde, mesures qui sont ensuite traitées mathématiquement (par transformée de Fourier). Par ailleurs, dans le cas de couches plus fines (inférieures à 50nm), nous utiliserons la variante ATR-FTIR (*Attenuated Total Reflectance*). Pour faire simple, l'ATR-FTIR est basée sur la réflexion d'un faisceau passant d'un milieu dense à un milieu moins dense (le diélectrique à analyser). Lors de la réflexion, une fraction du faisceau – alors appelée onde évanescente - pénètre dans le diélectrique, réagit ou non avec les liaisons présentes selon la longueur d'onde, puis rejoint le faisceau initial qui est collecté. Comme seule une partie du faisceau est modifiée et analysée, ceci permet d'obtenir une intensité du signal plus élevée dans le cas d'échantillons minces.

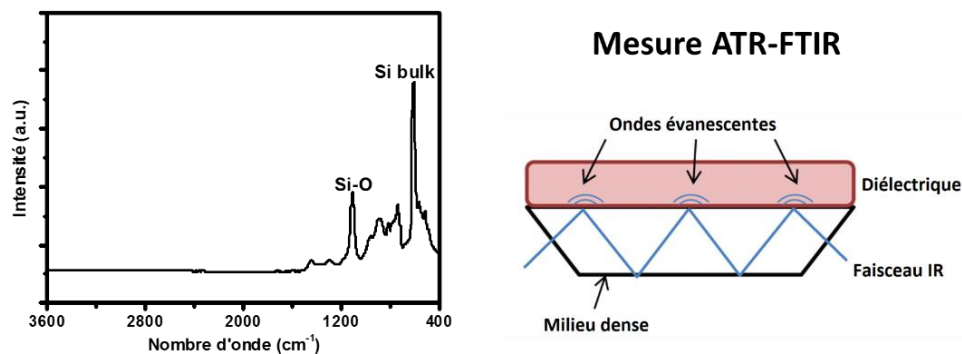


Figure 2.6 : Exemple d'un spectre du silicium bulk acquis par spectrométrie FTIR. Les différents pics en fonction du nombre d'onde du faisceau incident représentent les liaisons contenues dans l'échantillon. Ce spectre sera soustrait aux spectres des diélectriques étudiés afin de retirer la contribution du silicium bulk. A droite, le schéma illustre le principe de la mesure ATR-FTIR.

A noter qu'une même liaison peut vibrer de différentes manières (e.g. étirement ou torsion de la liaison) et donc réagir à différentes énergies incidentes, mais nous ne rentrerons pas dans ces détails dans cette étude.

- RBS/ERDA/NRA

Les mesures RBS (*Rutherford Backscattering Spectrometry*), ERDA (*Elastic Recoil Detection Analysis*) et NRA (*Nuclear Reaction Analysis*) sont trois techniques permettant d'analyser la **composition chimique** d'une couche mince (Figure 2.7). Pour cela, un flux accéléré d'ions ($^4\text{He}^+$ pour la RBS et l'ERDA dans notre cas) est envoyé sur l'échantillon et interagit avec les atomes de la couche, le tout sous vide (de l'ordre de 10^{-6} torr). La technique RBS se porte alors sur les ions incidents $^4\text{He}^+$ qui ont été rétro-diffusés (*backscattering*) par les atomes de la couche à analyser tandis que l'ERDA se concentre sur les atomes plus petits que l'hélium et diffusés vers l'avant (*elastic recoil*) sous l'effet de l'ion. On utilisera donc l'ERDA essentiellement pour analyser la concentration en hydrogène des couches minces, tandis que

la RBS nous permettra de remonter à la concentration en silicium. D'autres atomes sont accessibles par la technique RBS mais se superposent au signal du Si, d'où l'utilisation d'une troisième technique. La méthode NRA, elle, est basée sur l'envoi de protons ($^1\text{H}^+$) ou de deutons ($^2\text{H}^+$) sur la surface qui peuvent déclencher des réactions nucléaires selon l'énergie de l'ion incident, son angle ainsi que son isotope. L'analyse des particules ainsi créées (dans notre cas, $^4\text{He}^+$ ou $^1\text{H}^+$) permet alors également de remonter à la concentration de l'atome analysé dans la couche.

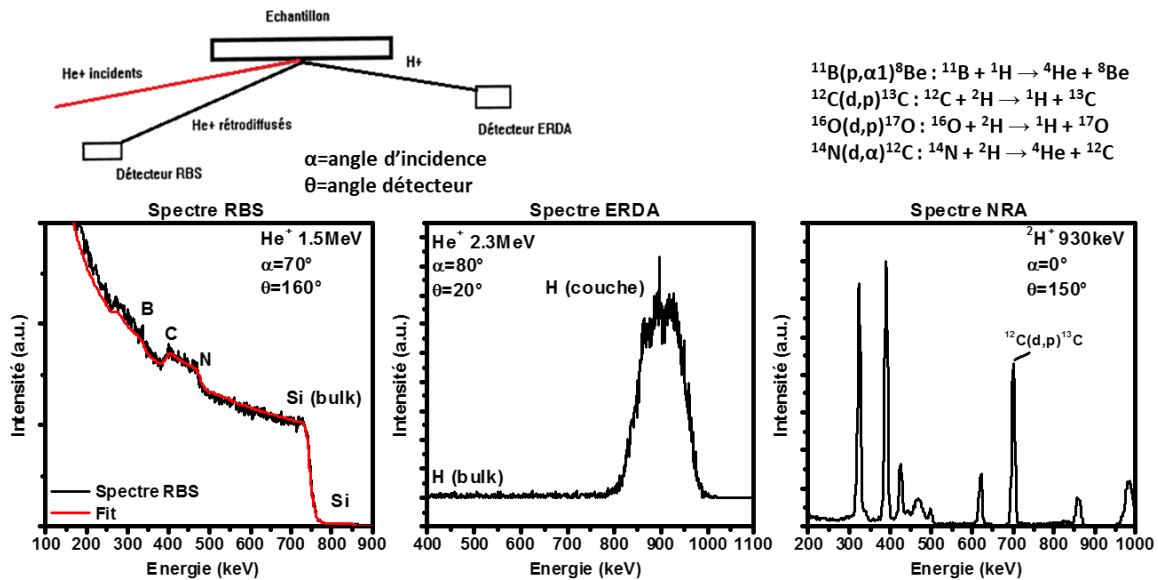


Figure 2.7 : Exemples de mesures RBS, ERDA et NRA pour une couche de 30nm de SiCN_{CVD} . La comparaison avec des spectres simulés (RBS, ERDA) ou des étalons (NRA) permet de remonter à la concentration atomique des éléments présents dans la couche diélectrique et donc sa composition.

- Goutte de mercure

Afin d'évaluer les propriétés électriques du matériau, une structure « MOS » peut être obtenue par l'apposition d'une goutte de mercure sur la surface, celle-ci constituant alors le contact métallique. L'utilisation du mercure comme contact permet non seulement de bénéficier de sa forme liquide à température et pression ambiantes, mais également de sa capacité à adhérer facilement à tout type de surface. Ainsi, une mesure capacité-voltage du MOS permet de remonter à la **permittivité** du diélectrique, tandis qu'une mesure courant-voltage permet d'extraire les **fuites de courant à travers le diélectrique** (correspondant à l'intensité mesurée à un champ électrique donné) ainsi qu'à sa **tension de claquage** (prise lorsque la densité de courant atteint $5\text{E-}4 \text{ A/cm}^2$).

Afin d'évaluer la stabilité thermique des diélectriques candidats, ces caractérisations ont été faites après dépôt ainsi qu'après un recuit de 600°C 15min. Les valeurs de permittivité (k), d'indice de réfraction (n), de densité massique (rho), de contrainte (sigma), de fuite J (donnée pour un champ électrique de 2MV/cm), de tension de claquage (Vbd) ainsi que la variation d'épaisseur (Delta t) des différents matériaux et leurs variations avant et après recuit sont résumées dans la Table 2.2.

Tout d'abord, nous voyons que les diélectriques low-k SiOCN , SiCO et SiCN ont effectivement tous une permittivité inférieure à celle du SiN quel que soit la température ou la technique de dépôt. Ces trois matériaux ont notamment en commun la présence d'atomes de carbone qui permet de réduire la constante diélectrique. Pour comprendre cela, la relation de Clausius-Mossotti permet de relier la permittivité k (grandeur macroscopique) aux propriétés microscopiques (densité d'atomes N et leur polarisabilité alpha) du matériau :

$$\frac{k - 1}{k + 2} = \frac{N \cdot \alpha}{3\epsilon_0}$$

Avec ϵ_0 la permittivité du vide.

A partir de cette expression, la permittivité peut être réduite de deux manières :

- En incorporant des atomes faiblement polarisables (i.e. dont les molécules ne réagissent pas ou peu à un champ électrique), ce qui est le cas du carbone ($\alpha_C^{4+} = 1E-27cm^3$) ou du bore ($\alpha_B^{3+} = 3E-27cm^3$) comparé au silicium ($\alpha_{Si}^{4+} = 1.7E-26cm^3$).
- En diminuant la densité atomique du matériau, par exemple par l'incorporation de pores dans le matériau mais au détriment de son intégrabilité (e.g. SiCOH poreux). On remarque notamment que le SiN_{400°C} et le SiCBN_{PECVD} qui sont respectivement plus denses que le SiN_{630°C} et le SiCBN_{CVD} ont effectivement une permittivité plus élevée.

Table 2.2 : Propriétés matériau et électrique des différents diélectriques directement après dépôt ou après un recuit 600°C 15min.

Température	SiN			SiOCN		SiCO	SiCBN	
	630°C	500°C	400°C	350°C	480°C	400°C	480°C	480°C
Technique dépôt	ALD	ALD	sPEALD	PEALD	PEALD	PECVD	CVD	PECVD
k	7	NM	7.4	4.5	4.7	4.5	3.8	5.2
Après recuit	7**	NM	7.3	4.9**	5.1**	4.6	4.2	NM
n@633nm	1.98	NM	1.98	1.73	1.76	1.66	1.91	1.88
Après recuit	NM	NM	=	=	=	=	=	=
ρ (g/cm ⁻³)	2.9	NM	3.01	2.23	2.34	2.19	1.78	2.05
Après recuit	NM	NM	+1%	+1.8%	+0.4%	+3.2%	=	+0.5%
σ (MPa)	-1000	NM	-880	-20	-30	55	-145	-370
Après recuit	NM	NM	+80MPa	+50MPa	+30MPa	+115MPa	=	+335MPa
Δt après recuit	NM	NM	+0.2%	-2.0%	-0.8%	-0.2%	0%	-1.0%
J (A/cm ²) @2MV/cm	~1E-8*	NM	1E-8	NM	1.4E-9	1.4E-9	2.8E-5	7.5E-9
Après recuit	NM	NM	NM	NM	NM	NM	5.5E-5	NM
Vbd (MV/cm)	~5.4*	NM	7.2	NM	11	8.8	5.6	6.3
Après recuit	NM	NM	NM	NM	NM	NM	5.4	NM

NM = Non Mesuré

*Pris de [Yamashita15] **Après 1050°C 2s (Résultats ST)

Ensuite, la diminution « simple » de la température de dépôt a un impact direct sur les propriétés du diélectrique. En effet, comparant SiOCN_{350°C} et SiOCN_{480°C}, une température plus faible de dépôt mène à une densité massique plus faible, et par conséquent, à une permittivité inférieure. En contrepartie, le matériau déposé à basse température semble plus sensible à un budget thermique subséquent, où ses propriétés varient légèrement plus après un recuit 600°C 15min que le SiOCN_{480°C}. Notons tout de même que les variations après le recuit restent assez faibles quel que soit le diélectrique et ne sont donc pas prohibitives pour l'intégration, mais nous verrons dans les prochaines sections que la diminution de la température a également d'autres impacts négatifs sur le diélectrique. Similairement, un dépôt assisté par plasma permet d'obtenir un matériau déposé plus dense, que cela soit dans le cas du SiCBN ou du SiN. L'utilisation d'un plasma permet en effet d'apporter une énergie supplémentaire à l'énergie thermique. Dans le cas du SiN_{400°C}, la densification du matériau obtenue par l'utilisation du plasma a probablement aussi été contrebalancée par la diminution de la température de 630°C à 400°C. Concernant les propriétés électriques, les diélectriques étudiés exhibent des fuites à travers le matériau ainsi que des tensions de claquage équivalentes voire meilleures que celles de la référence, exception faite du SiCBN_{CVD} où les fuites sont plus élevées de plusieurs décades.

Afin de mieux comprendre ces résultats, une étude des liaisons chimiques présentes dans les diélectriques a été réalisée au moyen de mesures FTIR. Les spectres du SiN, SiOCN, SiCO et SiCBN sont données dans la Figure 2.8.

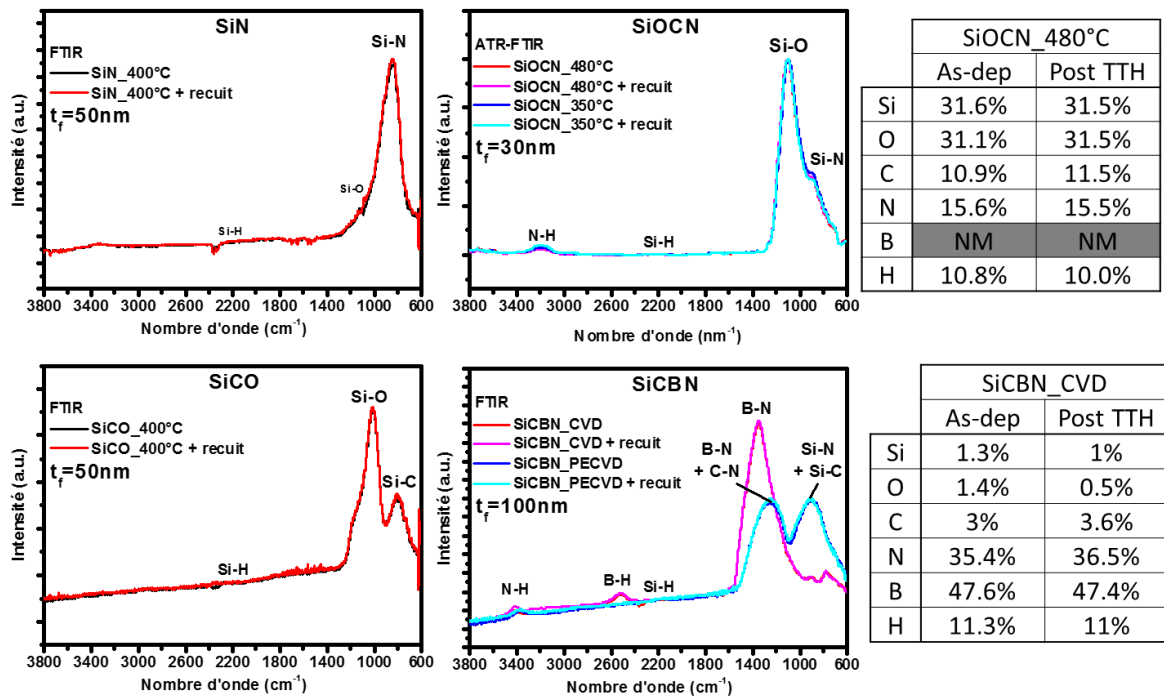


Figure 2.8 : Spectres (ATR-)FTIR acquis pour le SiN_{400°C}, SiOCN_{350°C} et SiOCN_{480°C}, SiCO_{400°C} et SiCBN_{CVD} et SiCBN_{PECVD} avant et après recuit 600°C 15min. Les compositions chimiques obtenues par analyses RBS/ERDA/NRA du SiOCN_{480°C} et du SiCBN_{CVD} sont également données.

Tout d'abord, ces spectres FTIR confirment la bonne stabilité des matériaux après un recuit 600°C 15min, les différentes liaisons dans les couches étant relativement peu impactées. Ceci est notamment expliqué par la faible quantité d'hydrogène (qui peut par exemple provenir des précurseurs utilisés) et de liaisons associées (e.g. Si-H vers 2200cm⁻¹) qui a tendance à dégazer sous l'effet de la température. On remarque toutefois la présence de pics de faible intensité à 3200cm⁻¹ et 2600cm⁻¹ dans le cas du SiOCN et du SiCBN_{CVD} qui peuvent être attribués à des liaisons N-H et B-H respectivement. Ces liaisons sont néanmoins plus stables thermiquement que les liaisons Si-H. Les mesures de compositions atomiques par RBS/ERDA/NRA avant et après recuit confirment par ailleurs une faible concentration d'hydrogène (~10%) dans le cas du SiOCN_{480°C} et du SiCBN_{CVD}.

L'analyse des pics entre 600cm⁻¹ et 1600cm⁻¹ est en revanche plus délicate, les contributions des différentes liaisons se mélangeant dans cette région. Le spectre du SiN nous permet d'abord d'identifier le pic à ~800-900cm⁻¹ correspondant aux liaisons entre le silicium et l'azote, liaisons que l'on retrouve également dans le SiCBN_{PECVD} et dans une moindre mesure dans les deux SiOCN. En outre, le pic à ~1000-1100cm⁻¹ dans le spectre du SiCO correspond aux liaisons entre le silicium et l'oxygène, l'autre pic à ~760-800cm⁻¹ étant alors caractéristique des liaisons Si-C [Benoit15]. Le « coude » aux alentours de 1200cm⁻¹ provient également de vibrations de liaisons Si-O mais suivant un autre mode. Les atomes d'oxygène et de carbone sont donc directement liés au silicium sans former de liaisons C-O (qui se manifesteraient par un pic vers 1700cm⁻¹) et qui sont plus sensibles à la température. On note aussi une intensité plus importante pour les liaisons Si-O que celles de Si-C, signe d'une concentration d'oxygène plus élevée dans la couche. Par ailleurs, le SiOCN présente également un pic intense à 1100cm⁻¹ correspondant aux liaisons Si-O. En revanche, les spectres du SiOCN ne mettent pas en évidence de liaisons avec le carbone. Ces liaisons sont probablement « noyées » dans les autres contributions, l'alternance de couches SiO₂ et SiCN lors du dépôt devant mener à la formation de liaisons Si-C (masquées par les liaisons Si-N) et/ou de liaisons C-N (masquées par les liaisons Si-O). Les analyses NRA confirment de plus la présence du carbone dans la matrice du SiOCN. Enfin, nous remarquons que les spectres du SiCBN déposé par CVD ou PECVD sont très différents, le premier

présentant un seul pic intense à $\sim 1250\text{-}1350\text{cm}^{-1}$ correspondant à des liaisons B-N et aucun pics associés à des liaisons avec le silicium ou le carbone. Ceci est confirmé par les analyses de composition chimique qui montrent qu'en réalité, le dépôt par CVD n'a pas ou peu incorporé de Si ni de C dans le film (il s'agit donc en fait d'un nitrure de bore). Au contraire, le procédé par PECVD présente deux pics, un à 900cm^{-1} et un à 1250cm^{-1} . Le premier provient probablement de liaisons Si-N, mais la largeur du pic indique également la contribution d'une autre liaison, possiblement des liaisons Si-C. De la même manière, le pic à 1250cm^{-1} peut être une combinaison de liaisons B-N et de liaisons C-N (normalement trouvées à 1220cm^{-1}). Ceci peut alors indiquer que la faible température (480°C) lors du dépôt CVD n'a pas permis de décomposer correctement les précurseurs de dépôt du SiCBN, tandis que l'utilisation d'un plasma permet de fournir une énergie supplémentaire.

Ces mesures FTIR nous apportent donc plusieurs informations. La faible présence d'hydrogène (et confirmée par les mesures de composition chimique pour le $\text{SiOCN}_{480^\circ\text{C}}$ et le $\text{SiCBN}_{\text{CVD}}$) explique la bonne stabilité thermique des couches étudiées. Elles nous ont également permis d'identifier les atomes incorporés dans la couche et, dans une certaine mesure, de les quantifier. Ceci explique en partie la permittivité des matériaux, le SiOCN et le SiCO se rapprochant de la permittivité du SiO_2 ($k=3.9$) et le $\text{SiCBN}_{\text{PECVD}}$ plutôt du SiN ($k=7$). Pour le $\text{SiCBN}_{\text{CVD}}$, la très faible permittivité est surtout expliquée par sa faible densité. Dans la prochaine section, nous verrons que les compositions des couches impactent également la stabilité chimique des matériaux.

2.1.2 Résistance aux chimies humides

Afin de garder son intégrité physique tout au long de l'intégration, les espaceurs doivent résister aux différentes chimies humides utilisées lors de la fabrication du transistor. Ces chimies diffèrent notamment selon le but (retrait polymères, nettoyage, stripping...) mais également selon les technologies et les équipements disponibles. Nous détaillons ci-dessous les différentes chimies étudiées dans cette section et leur utilité dans notre intégration :

- HF(0.5%)/HCl(1%) :

Le mélange HF/HCl est utilisé comme nettoyage chimique, l'acide fluorhydrique étant efficace pour la gravure des oxydes tandis que l'ajout d'acide chlorhydrique permet de retirer les éventuels éléments métalliques (e.g. Hf ou Ti de la grille). En particulier, ce nettoyage est utilisé juste après la gravure des espaceurs afin de retirer les polymères qui peuvent être formés ainsi qu'avant épitaxie (voir Section 3.1).

- UV-O₃ :

Le mélange HF/HCl crée une couche de passivation assez instable à la surface et qui a tendance à se recontaminer. Ainsi, une réoxydation chimique de la surface du silicium est généralement réalisée après le retrait des polymères par de l'ozone obtenu en irradiant de l'oxygène avec des rayons UV. Cet oxyde chimique est de nouveau retiré par un nettoyage HF/HCl enchaîné rapidement avec l'épitaxie pour préserver la surface du silicium de tout oxyde.

- HotSC1 (mélange $\text{NH}_4\text{OH}/\text{H}_2\text{O}_2$) et HotSPM (mélange $\text{H}_2\text{SO}_4/\text{H}_2\text{O}_2$) :

Le SC1 (pour *Standard Clean 1*) et le SPM (pour *Sulfuric Peroxide Mixture*) sont des chimies qui peuvent être utilisées pour retirer des couches organiques telles que les résines utilisées en photolithographie, et sont donc « vues » par les espaceurs après les étapes d'implantations LDD. Le choix de la chimie du stripping dépend notamment du ou des éléments implantés qui peuvent modifier la nature de la résine, mais également de la nature des sources et drains (e.g. Si, SiGe, SiC) afin d'en limiter la consommation. La dénomination « Hot » vient du fait que le SC1 est généralement chauffé entre 60°C et 80°C (température assez élevée pour activer le mélange sans toutefois décomposer l' H_2O_2), tandis que le mélange d'acide sulfurique et de peroxyde d'hydrogène est naturellement exothermique, portant le mélange à plus de 200°C [Kern90].

- H_3PO_4 :

L'acide orthophosphorique porté à 100°C est utilisé pour le retrait du masque dur SiN afin de libérer l'accès à la grille. A cette étape, les espaceurs de grille sont normalement protégés par des espaceurs SiO₂ très résistants au H₃PO₄, mais une faible vitesse de gravure dans cette chimie permettrait alors de s'affranchir des étapes de dépôt et de gravure du *spacer1* (cf. Chapitre 1).

Une bonne résistance à toutes ces chimies assurerait donc l'intégrité du matériau, en plus de permettre un meilleur contrôle de l'épaisseur des espaceurs tout au long de l'intégration. La Table 2.3 résume les vitesses d'attaque des différents matériaux étudiés dans les chimies citées.

Table 2.3 : Vitesse de gravure des différents diélectriques étudiés dans les chimies utilisées lors de la fabrication d'un transistor.

Vitesse en Å/min	SiN			SiOCN		SiCO	SiCBN	
Température	630°C	500°C	400°C	350°C	480°C	400°C	480°C	480°C
Technique dépôt	ALD	ALD	PEALD	PEALD	PEALD	PECVD	CVD	PECVD
HF 0.5%/HCl 1% @RT	3.9	9.6	3.2	NM	2.4	0.3	3	1.4
O ₃ @RT	NM	NM	0.03	NM	0.8	0.06	10	0.1
SC1 @60°C	NM	NM	0.1	NM	0	0.06	1.5	0.9
HotSPM @200°C	NM	NM	5.2	NM	3.3	3.2	2000	35
H ₃ PO ₄ @100°C	5	NM	NM	NM	0.06	0	5.5	2.7

Tout d'abord, nous voyons que la réduction de la température du même dépôt SiN de 630°C à 500°C dégrade la résistance aux chimies humides du matériau, la consommation dans un mélange HF/HCl étant triplée. En revanche, l'utilisation d'un plasma lors du dépôt ALD du SiN_{400°C} mène à une vitesse de gravure équivalente malgré la température de dépôt bien plus faible. Par ailleurs, le SiOCN_{480°C} et le SiCO_{400°C} exhibent des vitesses de gravure assez faibles quelle que soit la chimie. En particulier, le mélange HF/HCl consomme peu les deux low-k comparé au SiN malgré une présence importante de liaisons Si-O sensibles au HF, indiquant que l'incorporation de carbone est bénéfique pour la résistance chimique de ces matériaux. En outre, leurs consommations nulles ou presque dans l'H₃PO₄ pourrait éventuellement simplifier l'étape de retrait masque dur où le dépôt et la gravure du *spacer1* ne serait plus nécessaire. Enfin, si le SiCBN_{CVD} résiste bien aux chimies de nettoyage (HF/HCl, SC1), celui-ci est fortement sensible aux chimies oxydantes, en particulier au HotSPM, où il est consommé quasi instantanément. L'utilisation d'un plasma lors du dépôt permet d'améliorer la résistance chimique en densifiant le diélectrique (de la même manière que pour le SiN), avec de plus l'effet bénéfique d'une incorporation plus importante de carbone dans la couche. La consommation dans le HotSPM reste toutefois assez élevée.

Ces études sur pleines plaques donnent une première idée sur l'intégrabilité des différents matériaux, une attention particulière étant alors nécessaire pour le SiN_{500°C} lors du retrait polymère et du nettoyage avant épitaxie, ou bien pour le SiCBN_{CVD} où l'utilisation du HotSPM est prohibée. Par ailleurs, l'utilisation d'un plasma pour le dépôt du SiN_{400°C} améliore la tenue du matériau malgré sa faible température de dépôt, tandis que le SiOCN_{480°C} et le SiCO_{480°C} sont très peu consommés dans toutes les chimies étudiées. La température et la technique de dépôt du diélectrique ainsi que la composition chimique (incorporation de C) sont donc des paramètres importants pour la résistance chimique de ces matériaux. Dans la prochaine section, nous nous intéresserons à des structures patternées, où le comportement des diélectriques après dépôt et gravure des espaceurs puis épitaxie des sources et drains sera étudié.

2.2 Étude morphologique

2.1.3 Uniformité et conformité du dépôt

L'uniformité, la conformité et l'effet de densité sont des caractéristiques importantes d'un dépôt d'une couche mince. En effet, une variation d'épaisseur de la couche déposée sur la plaque ou au niveau d'un motif mène alors à des variabilités sur les propriétés électriques, liées par exemple à une position des jonctions ou à des capacités parasites différentes. La Figure 2.9 montre les observations SEM et TEM (lorsque disponibles) après dépôt des différents diélectriques basses températures. L'écart type σ sur l'épaisseur ainsi que la conformité (i.e. rapport épaisseur latérale sur épaisseur verticale) sont également reportés sur les images.

Tout d'abord, on remarque dès le dépôt l'apparition de trous aux pieds des grilles dans le cas du $\text{SiN}_{500^\circ\text{C}}$ et du $\text{SiCBN}_{\text{PECVD}}$. Dans le premier cas, ceci peut être dû à une température trop faible menant à un dépôt moins conforme (et qui sera confirmé dans la section suivante), tandis que l'observation TEM du $\text{SiCBN}_{\text{PECVD}}$ montre que l'épaisseur sur les flancs de la grille est très fine et deux fois inférieure à celle au-dessus. Si ce dernier problème peut être contourné en déposant simplement une couche plus épaisse, une optimisation des pressions partielles des précurseurs ou des conditions de dépôt peuvent améliorer la conformité. Par ailleurs, dans le cas du $\text{SiN}_{400^\circ\text{C}}$, on observe une épaisseur d'espaceur plus élevée d'un côté de la grille que de l'autre (8nm à gauche et 11nm à droite) qui peut être liée à la méthode de dépôt (à noter que l'indentation observée sur le flanc droit est liée à un problème de gravure de la grille et non au dépôt). En effet, la PEALD dite spatiale est basée sur le déplacement de la plaque à travers différentes chambres contenant les précurseurs de dépôt séparés par des rideaux de gaz inertes (N_2), contrairement au procédé classique où les précurseurs sont alternativement injectés et purgés dans une seule et même chambre. Dans le cas du $\text{SiN}_{400^\circ\text{C}}$, la plaque est placée sur un « carrousel » tournant entre les chambres, ce qui peut expliquer une vitesse de dépôt plus rapide d'un côté de la grille si la vitesse de rotation n'est pas optimisée.

Ainsi, ces premières observations permettent de mettre en lumière l'apparition de défauts dès le dépôt, en particulier dans le cas du $\text{SiN}_{500^\circ\text{C}}$. Nous verrons dans la section suivante que les défauts sont révélés ou amplifiés lors de la gravure des espaceurs. Par ailleurs, la faible conformité du $\text{SiCBN}_{\text{PECVD}}$ a été compensée en déposant une couche deux fois plus épaisse dans la suite de notre étude.

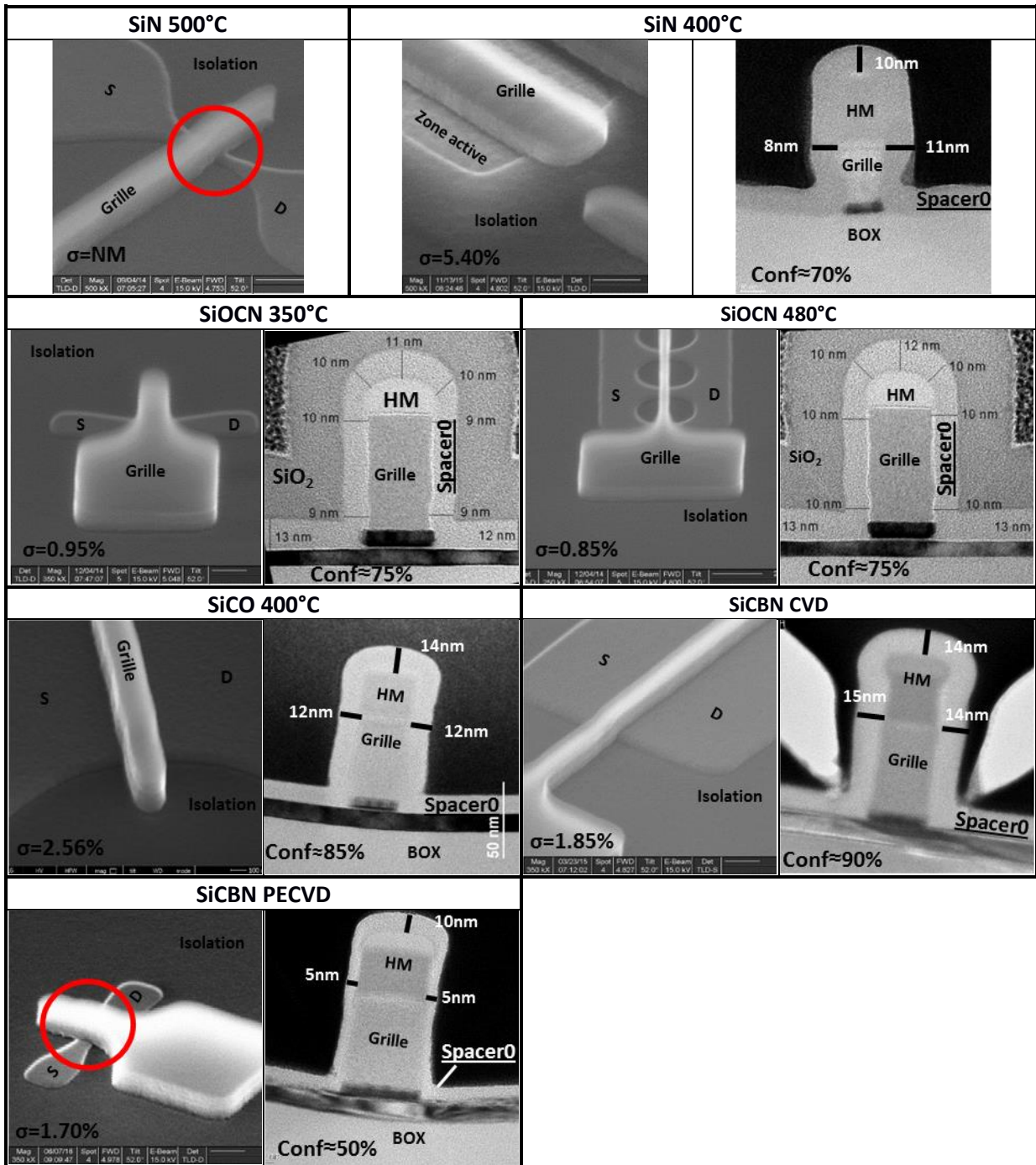


Figure 2.9 : Observations SEM et TEM des différents diélectriques étudiés après dépôt des espaceurs.

2.1.4 Gravure anisotropique des espaceurs

Après avoir déposé les espaceurs de grille, ceux-ci doivent être gravés afin de découvrir les zones d'accès sources, drains et grilles tout en conservant le diélectrique sur les flancs de la grille. Pour cela, une gravure anisotropique est réalisée, où des atomes réactants sont créés dans un plasma et bombardés unidirectionnellement sur la surface, permettant de retirer le diélectrique. Typiquement, des chimies fluorocarbonées sont utilisées, celles-ci permettant une bonne vitesse de gravure du SiN. La gravure se fait alors généralement en deux étapes, la première (appelée *main etch*) permettant de graver la majorité de la couche d'espaceur (e.g. avec du CHF₃ pour notre procédé standard), suivie par une deuxième étape (*over etch*) où des gaz plus sélectifs (e.g. CH₃F) sont utilisés pour limiter la

consommation du silicium de la zone active. Plusieurs leviers sont alors disponibles pour optimiser le profil et la sélectivité de la gravure, notamment par l'ajout d'autres gaz (O_2 , He, $SiCl_4$...) [Regis97].

Les observations SEM et TEM après gravure anisotropique standard sont données dans la Figure 2.10. On remarque alors que dans le cas des deux SiN déposés à basse température, des trous apparaissent tout au long du bas de la grille après gravure. Ces défauts, déjà observés après dépôt pour le $SiN_{500^\circ C}$, sont révélés lors de la gravure et peuvent notamment être dus à une composition non uniforme de ces couches minces, avec par exemple une concentration d'azote plus élevée au niveau des arêtes, causant une consommation plus importante de la couche dans ces zones. Ces trous dans les espaceurs posent problème lors de l'épitaxie qui peut alors se faire à travers les espaceurs et causer un court-circuit entre les sources et drains et la grille. Ce même problème est observé sur le $SiOCN_{350^\circ C}$, tandis qu'un dépôt à plus haute température permet une meilleure tenue lors de la gravure pour le $SiOCN_{480^\circ C}$. Par ailleurs, une forte consommation latérale des espaceurs $SiCBN_{CVD}$ est observée, passant d'une épaisseur de 12nm après dépôt à 2nm après gravure. Cette observation indique que sous l'effet du plasma, la nature du diélectrique a été modifiée, menant à une sensibilité plus importante aux chimies standards de gravure. Dans une moindre mesure, le $SiCBN_{PECVD}$ (où l'épaisseur déposée a été augmentée pour compenser la faible conformité) semble également plus rugueux sur les flancs de la grille après la gravure, avec cette fois une consommation latérale de 3-4nm. Enfin, aucun défaut n'est visible après la gravure du $SiCO_{400^\circ C}$, avec ici une consommation latérale nulle.

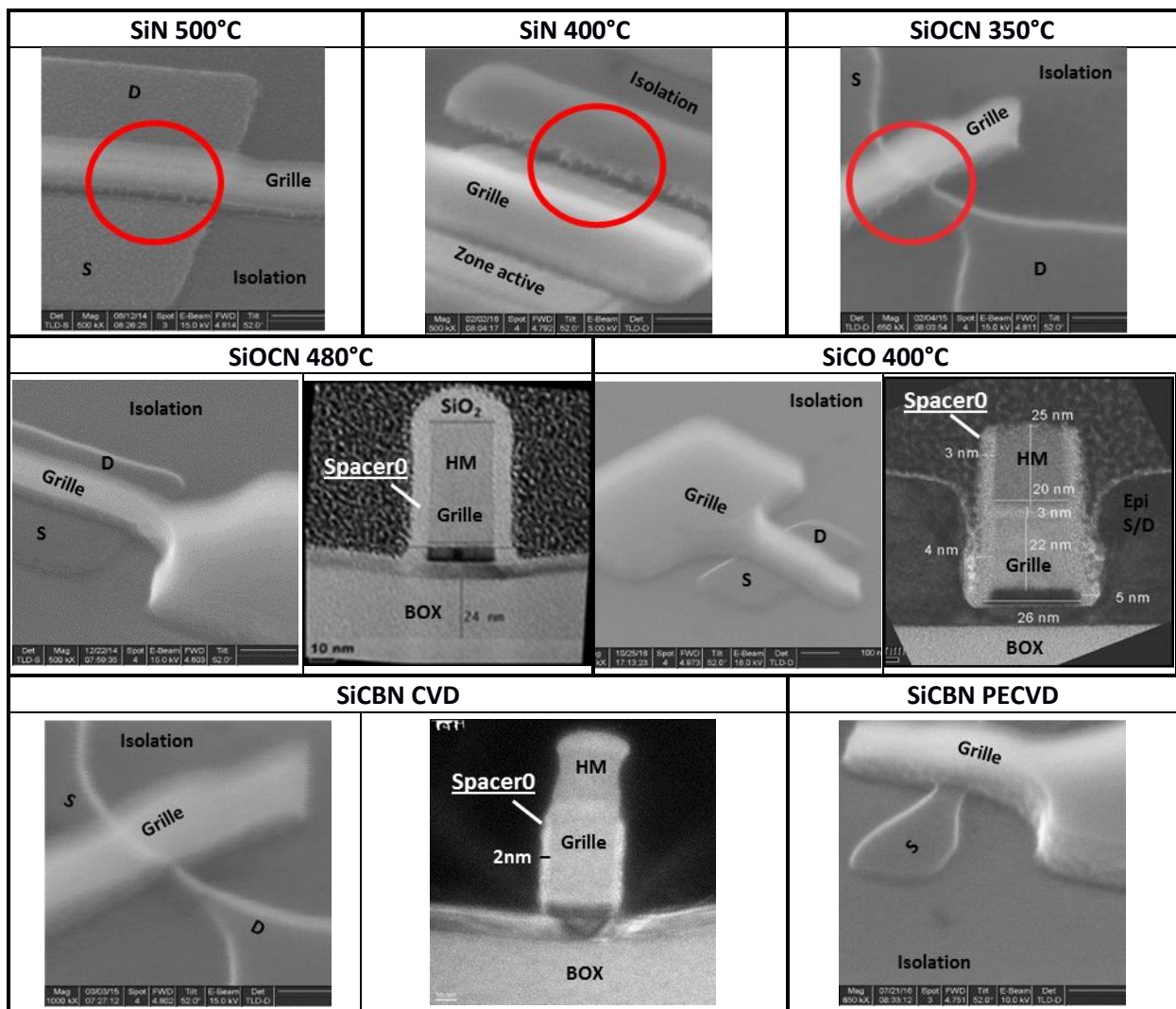


Figure 2.10 : Observations SEM et TEM des différents diélectriques étudiés après gravure anisotropique.

Ces résultats montrent que si le dépôt peut sembler correct, la gravure des espaceurs vient empirer les défauts déjà présents, révéler les problèmes de conformité ou même modifier le matériau en volume. Des optimisations de la gravure anisotropique peuvent néanmoins être faites afin de préserver l'intégrité des espaceurs. Par exemple, le pré-dépôt d'une couche protectrice (e.g. d'oxyde ou de polymères) à la surface des espaceurs permet de protéger les flancs des espaceurs lors de la gravure (couche retirée ensuite par un nettoyage HF), méthode utilisée pour le $\text{SiCBN}_{\text{CVD}}$ par exemple (Figure 2.11). Cependant, si l'épaisseur latérale est effectivement préservée après la gravure, la couche de protection mène à l'apparition de « pieds » d'espaceurs. Cette épaisseur plus importante aux pieds des espaceurs pose d'une part le problème de la position des jonctions lors de l'implantation des dopants qui se trouvent alors plus éloignés du canal, et d'autre part cette zone peut également faciliter la pénétration de chimies humides lors des étapes suivantes. Par exemple, un phénomène de « lift-off » de la grille est apparu lors de l'étape de retrait masque dur indiquant que les espaceurs de grille ont totalement disparus malgré la présence du *spacer1* en SiO_2 censé les protéger.

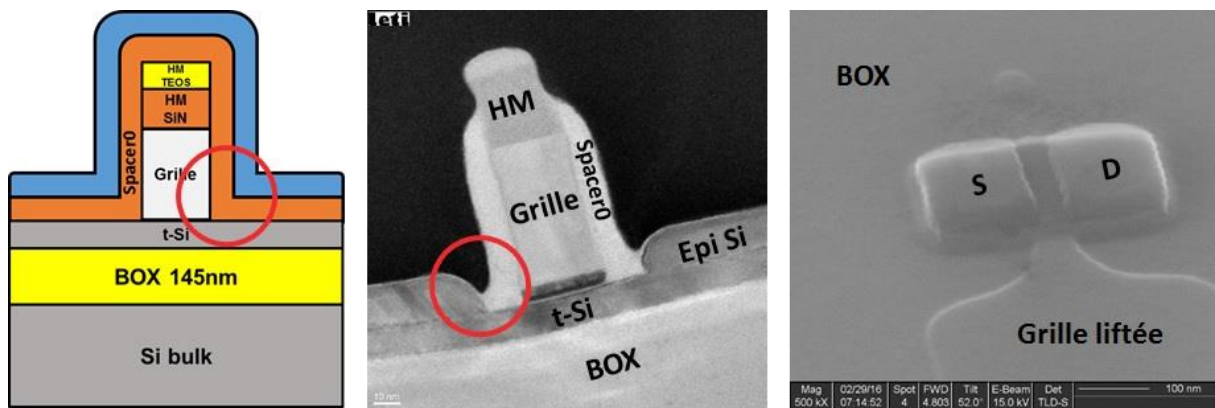


Figure 2.11 : Principe de la protection des flancs de grille par le dépôt d'une couche d'oxyde ou de polymères. Cette couche préserve effectivement les flancs des espaceurs de grille mais mène en contrepartie à la formation de pied d'espaceurs. Ces pieds sont une zone d'infiltration potentielle des chimies humides qui ont par exemple mené au « lift-off » de la grille lors du retrait masque dur ($\text{HF}/\text{HCl}+\text{H}_3\text{PO}_4$).

Nous avons donc vu que l'utilisation d'un plasma peut endommager les flancs des espaceurs. D'autres techniques de gravure des espaceurs ont également été récemment développées, telles que la gravure par implantation/gravure chimique [Posseme14]. L'idée est d'implanter de l'hélium ou de l'hydrogène unidirectionnellement dans les espaceurs de grille, ce qui a pour effet de modifier le diélectrique sans impacter les flancs. Une gravure humide (e.g. avec du HF) vient ensuite retirer les zones implantées tout en limitant la consommation latérale du matériau. L'énergie utilisée lors de l'implantation ionique étant faible, les flancs du diélectrique devraient être préservés. Cette technique reste toutefois en développement et n'a pas été implémenté dans notre étude.

Ainsi, si la plupart des matériaux semblaient prometteurs au regard de leurs propriétés matériau, leur stabilité thermique et leur résistance chimique, seuls le $\text{SiCO}_{400^\circ\text{C}}$, le $\text{SiCBN}_{\text{PECVD}}$ et le $\text{SiOCN}_{480^\circ\text{C}}$ ont pu être gravés correctement avec une chimie standard, les autres matériaux ayant été abandonnés.

2.1.5 Reprise de l'épitaxie source et drain

Un dernier critère important pour valider l'intégrabilité des matériaux étudiés est la bonne reprise de l'épitaxie des sources et drains. Comme observé sur la Figure 2.1, la présence de résidus sur les zones actives après la gravure peut par exemple empêcher une bonne croissance. La Figure 2.12 montre les observations SEM après épitaxie standard de silicium à 750°C pour le SiOCN_{480°C}, SiCO_{400°C} et le SiCBN_{PECVD}.

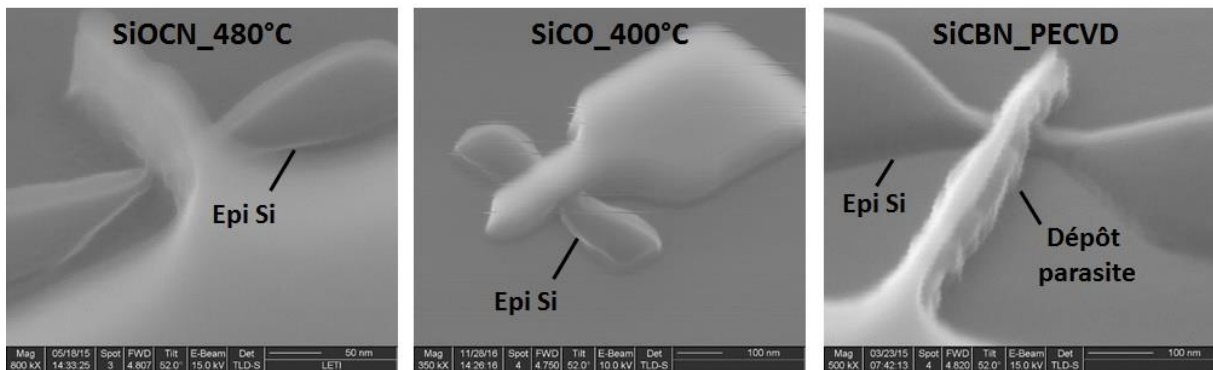


Figure 2.12 : Observations SEM après épitaxie sources et drains standard (Si à 750°C) avec différents espaceurs de grille. Dans le cas du SiCBN_{PECVD}, un dépôt parasite tout le long de la grille est apparu.

On remarque que si dans le cas des deux premiers low-k, l'épitaxie est lisse et faite uniquement sur les zones sources et drains, signe que l'épitaxie a bien reprise et qu'aucun dépôt n'a été fait sur ou à travers les espaceurs, un dépôt parasite s'est formé tout autour de la grille censée être protégée par le SiCBN_{PECVD}. Ceci peut indiquer soit une mauvaise étanchéité des espaceurs laissant passer les précurseurs, soit une disparition de la couche. Le suivi de l'épaisseur latérale des espaceurs de chaque côté de la grille (Figure 2.13) montre qu'en réalité, le diélectrique a été totalement consommé lors de l'étape de retrait des polymères qui est faite par la combinaison d'un nettoyage HF/HCl suivi d'une oxydation chimique UV-O₃. Ainsi, si les vitesses de gravure mesurées après dépôt sur pleine plaque dans ces chimies humides étaient faibles, le passage sur des structures patternées ainsi que la modification du matériau sous l'effet du plasma lors de la gravure ont rendu le diélectrique plus sensible.

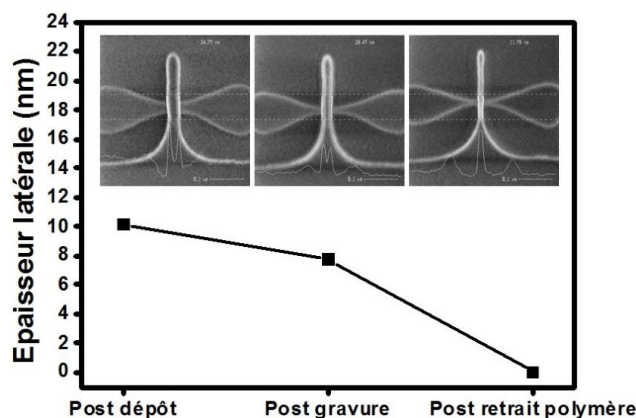


Figure 2.13 : Evolution de l'épaisseur latérale d'espaceurs SiCBN_{PECVD} déposés à 480°C au cours de l'intégration. L'étape de retrait polymère (HF/HCl+UV-O₃) a totalement consommé les espaceurs après sa probable modification lors de la gravure anisotropique.

Au final, seuls les low-k SiCO et SiOCN ont montré une intégrabilité satisfaisante dans notre étude, le SiCBN_{PECVD} étant donc également abandonné. Des optimisations de la séquence

dépôt/gravure/épitaxie sont possibles mais sortent alors du cadre de ce travail. Dans la section suivante, nous allons évaluer l'influence de ces deux low-k sur les performances électrique statique et dynamique de transistors. Les travaux ayant été menés séparément, nous traiterons tout d'abord du cas du SiOCN dont l'intégration a été menée au Leti avec des règles de dessin 65nm, puis celui du SiCO sur une technologie 14FDSOI de STMicroelectronics.

2.3 Étude électrique

2.1.6 Intégration du SiOCN

Afin d'étudier l'influence électrique d'un low-k, des transistors CMOS ont été fabriqués intégrant du SiOCN_{480°C} comme espaceurs de grille (Figure 2.14). L'influence des trous observés lors de la gravure du SiOCN_{350°C} sera également présentée.

Le dépôt et la gravure des espaceurs ont été suivis d'une épitaxie sources et drains standard de Si à 750°C, des implantations LDD (BF₂ côté pMOS et As côté nMOS), du retrait masque dur, des implantations HDD (co-implantation B+BF₂ côté pMOS et P+As+P côté nMOS) et enfin d'un recuit d'activation à haute température. Dans un premier temps, seule la brique espaceurs de grille a donc été modifiée par rapport au process flow standard afin de valider leur intégration.

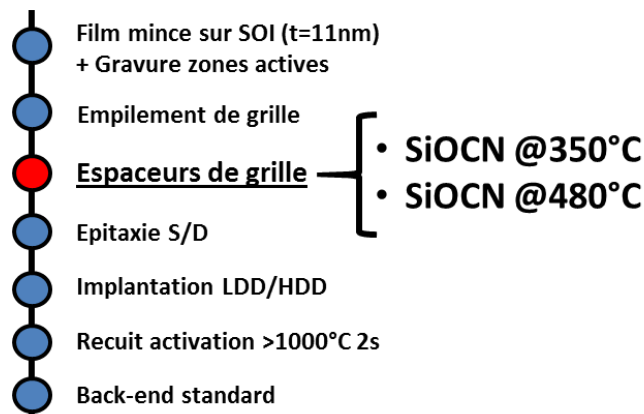


Figure 2.14 : Etapes de fabrication de transistors CMOS intégrant des espaceurs de grille SiOCN déposés à 350°C ou 480°C par PEALD. Seuls les espaceurs ont été modifiés par rapport au process flow à haute température.

La Figure 2.15 montre tout d'abord le compromis Ion/Ioff pour les pMOS (à gauche sur le graphe) et les nMOS (à droite) des deux variantes SiOCN_{350°C} et SiOCN_{480°C} comparées à une référence avec des espaceurs SiN_{630°C}.

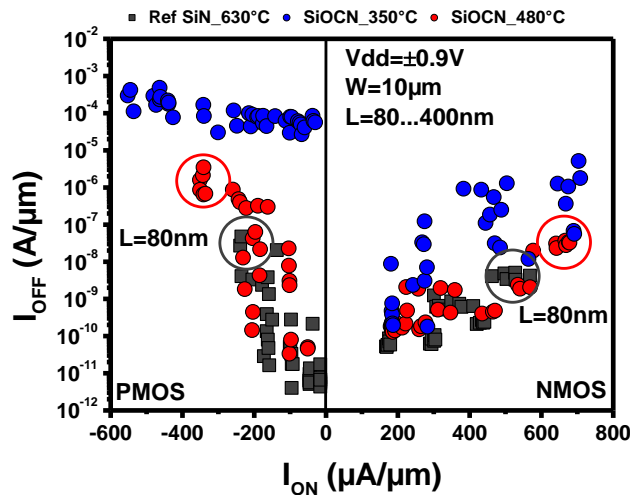


Figure 2.15 : Compromis Ion/Ioff pour les pMOS et nMOS intégrant différents espaceurs de grille sur des transistors larges ($W=10\mu\text{m}$).

Nous observons que les transistors intégrant le $\text{SiOCN}_{480^\circ\text{C}}$ exhibent des courants débités I_{on} ainsi que des courants de fuite I_{off} équivalents à la référence avec des espaceurs $\text{SiN}_{630^\circ\text{C}}$. Ceci indique notamment que le remplacement du $\text{SiN}_{630^\circ\text{C}}$ par le low-k n'a pas eu d'influence sur les accès sources et drains, preuve de la bonne reprise d'épitaxie après la gravure des espaceurs. En revanche, l'utilisation d'un $\text{SiOCN}_{350^\circ\text{C}}$ mène à plus de variabilité, avec des dispositifs plus fuitieux (I_{off} plus élevé), en particulier sur les transistors pMOS. Ces performances dégradées sont notamment expliquées par des fuites de grilles importantes dues aux trous observés dans la Section 2.2.2 et à une probable épitaxie parasite sur la grille.

Par ailleurs, afin d'évaluer les performances dynamiques des transistors, des oscillateurs en anneau (*ring oscillator*) ont été mesurés. Un oscillateur en anneau est un circuit composé d'une succession de N inverseurs en boucle fermée, inverseurs qui sont eux-mêmes composés d'un transistor de type n et un de type p dont les grilles sont interconnectées (Figure 2.16). Une tension d'entrée d'un inverseur V_{in} faible (0) rend passant le transistor pMOS, mettant le potentiel de V_{out} à V_{DD} (1). Inversement, une tension V_{in} forte (1) rend passant le transistor nMOS, changeant le potentiel de V_{out} à la masse (0).

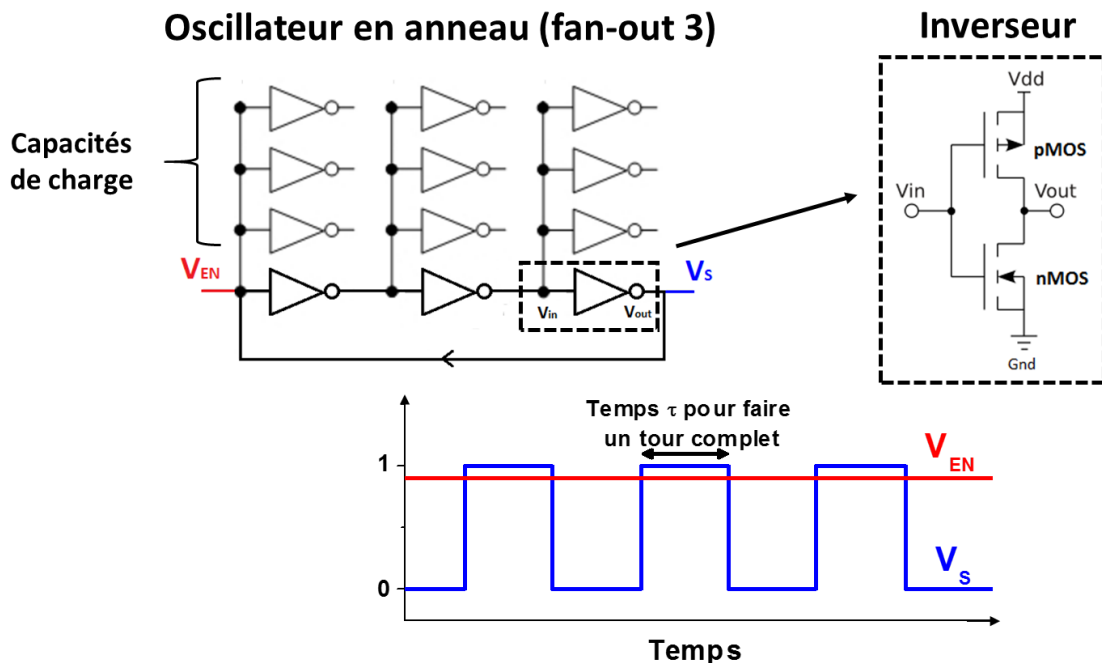


Figure 2.16 : Schémas électriques d'un ring oscillator fan-out 3 et d'un inverseur CMOS.

Ainsi, l'application d'une tension continue V_{EN} en entrée de l'oscillateur en anneau fait commuter successivement chaque inverseur du circuit, faisant alors osciller la tension de sortie entre V_{DD} et $0V$. Un nombre impair N d'inverseurs est donc nécessaire afin que le circuit puisse effectivement osciller. Des inverseurs peuvent également être ajoutés en parallèle à la sortie de chaque étage (trois dans le cas d'un « fan-out » FO3), ce qui permet d'introduire des capacités parasites (appelés capacités de charge) supplémentaires afin de ralentir l'oscillation et faciliter les mesures. Dans la pratique, des diviseurs de fréquence en sortie du circuit sont aussi rajoutés pour encore ralentir le signal. A partir de la fréquence d'oscillation f de la tension de sortie, on peut alors extraire un délai τ_p correspondant au temps de commutation d'un seul inverseur :

$$\tau_p = RC = \frac{\tau}{N} = \frac{1}{2 \cdot f \cdot N}$$

Une figure de mérite permettant de comparer le SiOCN et la référence SiN est le délai en fonction du courant statique, celui-ci correspondant aux courants de fuites des transistors du circuit à l'état OFF, c'est-à-dire quand $V_{EN}=0V$ (Figure 2.17).

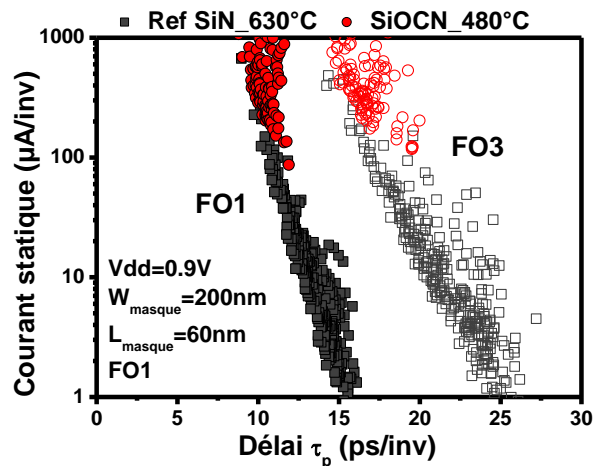


Figure 2.17 : Figure de mérite du courant statique en fonction du délai de basculement d'un inverseur pour des oscillateurs en anneau FO1 et FO3. Le rajout des capacités parasites (FO3) ralentit effectivement les délais par inverseur.

Si l'on ne regarde qu'une seule des deux variantes, on remarque une décroissance linéaire du délai τ_p lorsque le courant statique augmente. Cette tendance renvoie directement au compromis I_{on}/I_{off} , où la commutation plus rapide d'un inverseur est associée à une tension de seuil plus faible, et par conséquent à des courants I_{on} et I_{off} plus élevés. Par ailleurs, l'ajout de capacités de charge (FO3) ralentit effectivement les délais. Ceci dit, on remarque alors que malgré une permittivité plus faible qui devrait permettre de réduire les capacités parasites de chaque transistor, le SiOCN_{480°C} présente la même tendance linéaire que la référence, mais avec des courants de fuites statiques plus élevés. Ce résultat peut être expliqué par le fait que si sa valeur k est plus faible que celle du SiN, l'épaisseur des espaceurs SiOCN_{480°C} après gravure est aussi plus faible, compensant le gain en permittivité dans l'expression d'une capacité (Figure 2.18). Ceci est dû à un dépôt de départ trop fin (9nm pour le SiOCN, 12nm pour le SiN) et qui ne permet donc pas une comparaison rigoureuse des performances dynamiques.

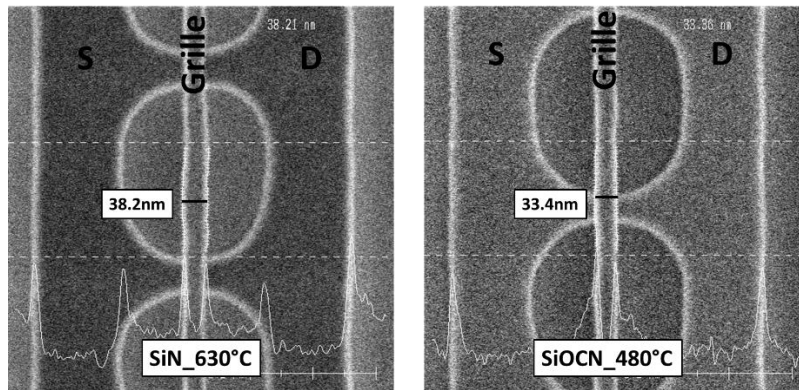


Figure 2.18 : Observations et mesures CD-SEM d'un dispositif en réseau avec des espaceurs SiN (gauche) ou SiOCN (droite) sur des grilles dessinées à $L=15\text{nm}$. Les espaceurs SiOCN sont moins épais dû à un dépôt ciblé de départ trop fin (9nm contre 12nm pour le SiN).

En conséquence, les capacités parasites sont probablement identiques dans les deux cas, avec un courant statique plus élevée pour le SiOCN_{480°C} du fait de son épaisseur plus fine. Ce résultat est corroboré par les paramètres de DIBL et de V_T des transistors donnés en Figure 2.19.

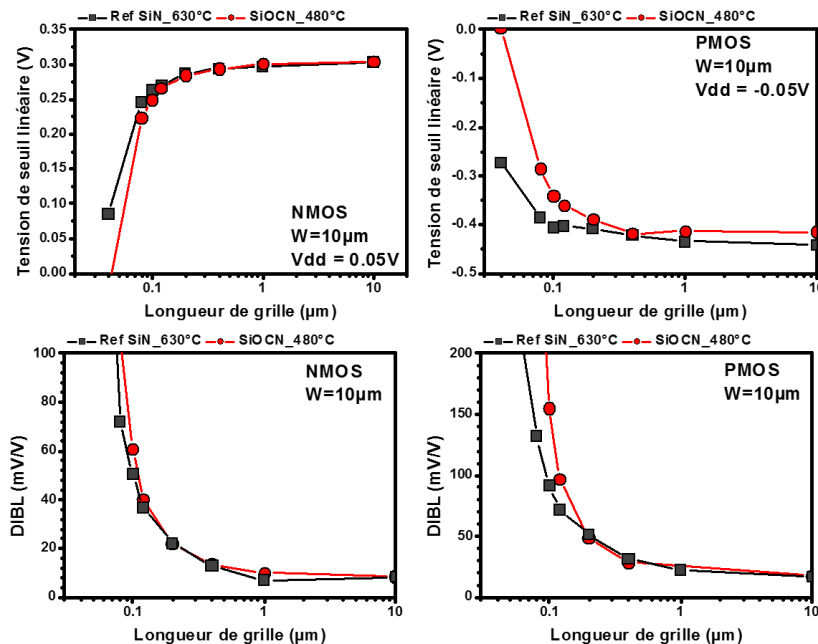


Figure 2.19 : Tension de seuil linéaire et DIBL en fonction de la longueur de grille pour les pMOS et nMOS sur des transistors larges ($W=10\mu\text{m}$). Les effets de canaux courts sont plus importants pour les transistors intégrant le SiOCN_{480°C} du fait de son épaisseur plus faible et d'un overlap plus grand.

On voit en effet que la dégradation du DIBL apparaît plus tôt (i.e. pour des longueurs de grilles plus élevées) dans le cas du SiOCN_{480°C}, signe que les dopants implantés et diffusés par le recuit d'activation à haute température sont plus proches du canal. Par conséquent, les effets de canaux courts sur la tension de seuil sont plus importants pour le SiOCN. Ceci est également visible sur le compromis Ion/Ioff, où des transistors de même dimension (e.g. $W/L=10\mu\text{m}/80\text{nm}$) se retrouvent sur la même tendance linéaire mais décalés, avec des dispositifs moins fuites mais débitant également moins de courant pour le SiN_{630°C}.

Ainsi, nous avons intégré pour la première fois du SiOCN déposé à 480°C sur des transistors et démontré des performances statiques équivalentes à la référence avec des espaceurs SiN déposé à plus haute température. Cette étude électrique ne nous a par contre pas permis de démontrer le gain apporté par un diélectrique low-k du fait de la différence d'épaisseur avec la référence. C'est pourquoi

nous allons également nous intéresser aux résultats électriques de dispositifs intégrant le $\text{SiCO}_{400^\circ\text{C}}$ obtenus dans [Benoit15].

2.1.7 Intégration du SiCO

Le low-k $\text{SiCO}_{400^\circ\text{C}}$ a été intégré sur une technologie 14FDSOI (i.e. aux règles de dessin 14nm) de STMicroelectronics. Une des particularités du 14FDSOI est qu'elle intègre une double couche d'espaceurs $\text{SiN}_{630^\circ\text{C}}$ (*dual spacer*) pour une problématique de surélévation des sources et drains duale, chose sur laquelle nous reviendrons plus en détails dans le Chapitre 3. Ainsi, dans un premier temps, une première couche d'espaceurs est déposée et gravée uniquement du côté des nMOS. Suit une épitaxie de type N (SiC:P) puis le dépôt d'une seconde couche d'espaceurs qui est cette fois gravée uniquement du côté des pMOS, suivi d'une épitaxie de type P (SiGe:B). Enfin, les espaceurs sont également gravés côté nMOS. Au final, l'espacement entre la grille et les sources et drains surélevés diffèrent entre les deux types de transistors, celui-ci étant plus élevé du côté des pMOS (Figure 2.20).

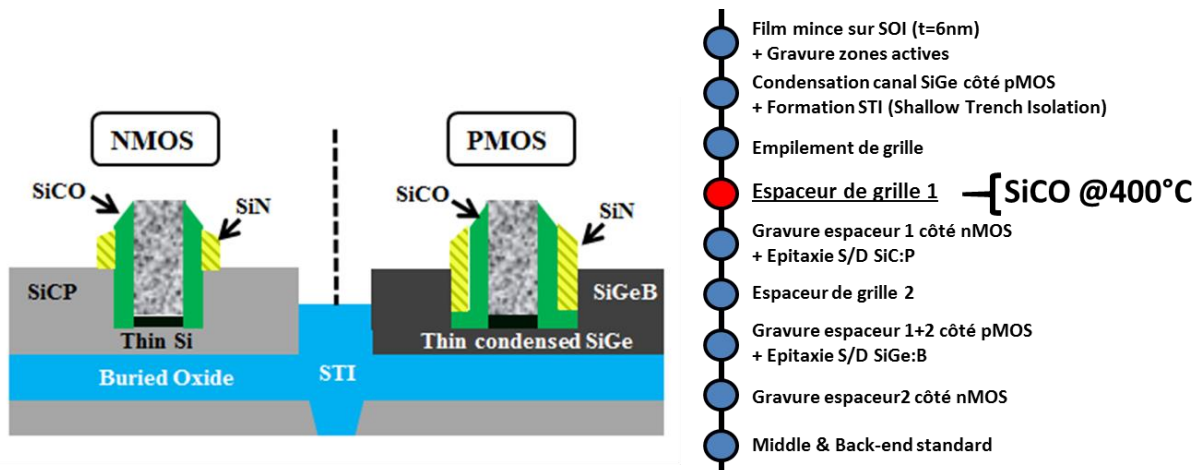


Figure 2.20 : Etapes de fabrication simplifiées de la technologie 14FDSOI de STMicroelectronics. Les espaceurs sont composés d'une double couche de diélectrique nécessaires à l'intégration d'une épitaxie duale selon le type de transistor. Les épaisseurs des espaceurs de grille ne sont donc pas les mêmes selon le type de transistor, ce qu'il faut alors prendre en compte lors du design des circuits.

Electriquement, comme le $\text{SiOCN}_{480^\circ\text{C}}$, les transistors intégrant le $\text{SiCO}_{400^\circ\text{C}}$ exhibent un compromis I_{on}/I_{off} équivalent avec la référence, signe que là aussi le changement de matériau n'a pas impacté les accès sources et drains (Figure 2.21a et b). On notera également que les transistors de même dimension ($W/L=170\text{nm}/20\text{-}30\text{nm}$) ont cette fois les mêmes tensions de seuil, indiquant que les espaceurs low-k et SiN ont bien la même épaisseur.

En terme dynamique, l'intégration du low-k permet la réduction des délais par inverseur τ_p de 5% pour un courant statique équivalent (Figure 2.21c). Ceci est donc lié à la réduction des capacités parasites, la capacité totale (que l'on ne pouvait pas mesurer sur notre masque par manque des structures nécessaires) étant également réduite de 5% (Figure 2.21d). Enfin, les bonnes propriétés électriques du $\text{SiCO}_{400^\circ\text{C}}$ observées dans la Section 2.1.1 sont également retenues après intégration complète, avec par exemple un gain de 3.5V sur la tension de claquage des transistors nMOS (Figure 2.21e). Côté pMOS, la tension de claquage plus faible peut être due à la bicouche avec le SiN .

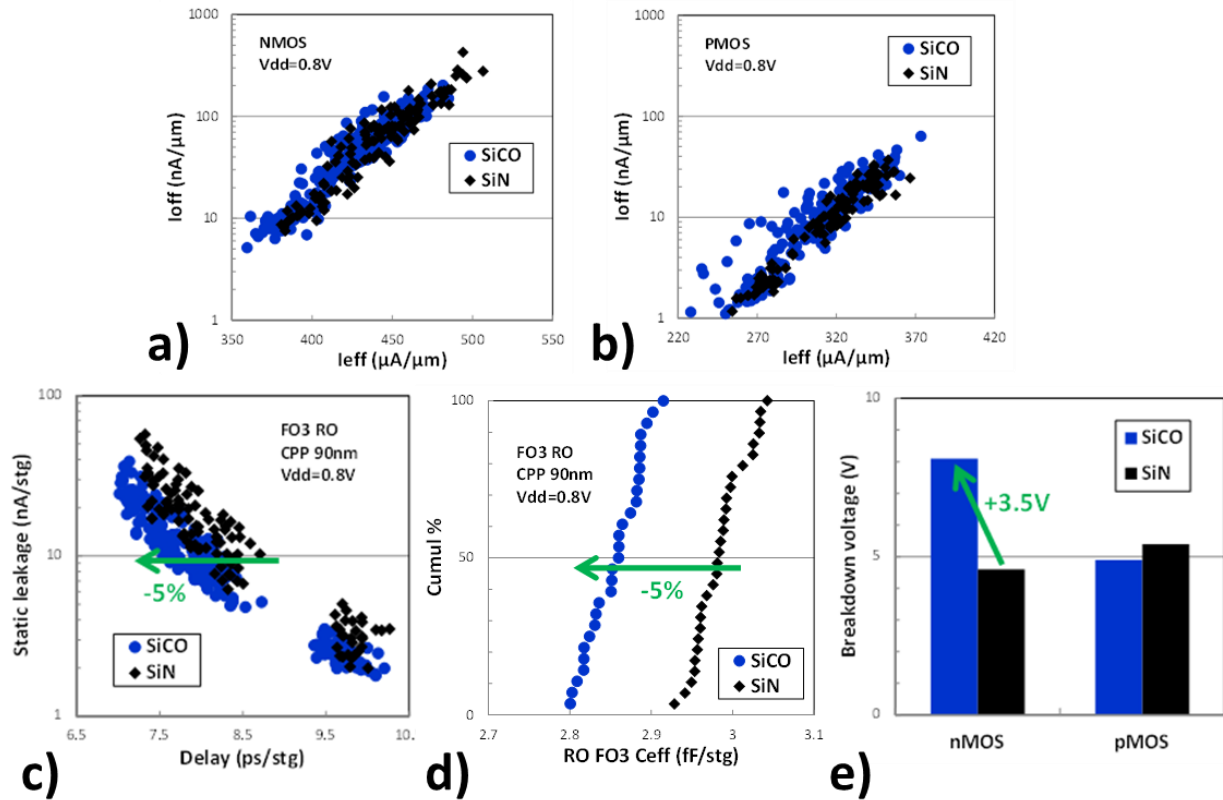


Figure 2.21 : Résultats électriques sur des transistors n et pMOS intégrant des espaceurs SiCO déposés à 400°C. Compromis I_{on}/I_{off} mesuré sur des transistors étroits et courts ($W/L=170\text{nm}/20\text{-}30\text{nm}$) ; c) Figure de mérite courant statique et délai par inverseur d'un oscillateur en anneau FO3 ; d) Capacité totale mesurée par inverseur; e) Tension de claquage mesuré sur des transistors GO2 ($EOT=3\text{nm}$) [Benoit15].

Le SiCO_{400°C} apparaît donc également comme un excellent candidat pour le remplacement du SiN_{630°C} comme espaceurs de grille. Sa stabilité thermique ainsi que sa résistance chimique permettent une intégration simple sans modification des autres procédés (e.g. gravure anisotropique) et ses bonnes propriétés matériaux mènent à des gains dynamiques démontrés dans cette étude.

Conclusions du Chapitre 2

Dans ce chapitre, nous avons évalué différents diélectriques déposés à basse température afin de remplacer le nitrure de silicium comme espaceur de grille dont la température de dépôt (630°C) est trop élevée pour le niveau haut de l'intégration 3D Séquentielle.

Nous avons vu que **la diminution de la température d'un dépôt dégrade l'intégrabilité du matériau** (cas du SiN_{500°C}, SiN_{400°C} et du SiOCN_{350°C}). En effet, un matériau déposé à une température faible est plus sensible aux budgets thermiques subséquents ainsi qu'aux différentes chimies humides utilisées pendant l'intégration (retrait polymères, nettoyages, stripping). De plus, la composition de ces matériaux semble moins conforme lorsque déposés sur des structures patternées. Ceci mène à **la formation de défauts lors de la gravure anisotrope des espaceurs qui sont rédhibitoires pour leur intégration**. L'apparition de trous aux pieds des espaceurs pose par exemple le problème de la reprise d'épitaxie avec un dépôt parasite sur la grille qui court-circuite les accès, ou bien de contamination potentielle des équipements par le TiN qui n'est plus encapsulé correctement. Néanmoins, **l'utilisation d'un plasma lors du dépôt permet d'apporter une énergie supplémentaire à l'énergie thermique**, ce qui permet d'avoir des matériaux plus stables et plus résistants (cas du SiCBN_{PECVD} par exemple).

Par ailleurs, l'étude des matériaux low-k montre l'intérêt de ces diélectriques comme espaceurs ou comme diélectrique de back-end, ce même pour des technologies avancées qui ne sont pas limitées en budget thermique. En effet, **la faible permittivité des low-k permet une réduction des délais de propagation dans les transistors** grâce à une diminution des capacités parasites liée à leur utilisation. Ainsi, l'intégration du SiOCN_{480°C} dans un process flow standard du Leti n'a pas impacté les performances statiques du transistor, signe d'une bonne intégrabilité tandis que les performances dynamiques n'ont pas pu être comparées à la référence SiN_{630°C} de manière rigoureuse. Egalement, un compromis Ion/Ioff similaire à la référence a été obtenu avec l'intégration de SiCO_{400°C}, avec en plus une augmentation de la tension de claquage des transistors nMOS. De plus, une réduction de 5% de la capacité totale ainsi que du délai de commutation d'un inverseur a été démontrée. Enfin, si le SiCBN_{CVD} exhibe la permittivité la plus faible des matériaux étudiés (k=3.8), une stabilité thermique satisfaisante et des consommations faibles dans les chimies humides (hors HotSPM), la modification de son volume par le plasma lors de la gravure anisotrope des espaceurs n'a pas permis son intégration. Ce résultat montre notamment **l'importance de l'enchaînement dépôt/gravure/épitaxie lors de l'étude des espaceurs de grille**. La Figure 2.22 résume les résultats obtenus.

	SiN			SiOCN		SiCO	SiCBN	
Température	630°C	500°C	400°C	350°C	480°C	400°C	480°C	480°C
Technique dépôt	ALD	ALD	PEALD	PEALD	PEALD	PECVD	CVD	PECVD
Propriétés matériaux	k élevé	k élevé	k élevé	++	++	++	+++	+
Stabilité thermique	Ref	NM	OK	OK	OK	OK	OK	OK
Résistance chimique	Ref	HF/HCl	=Ref	NM	++	+++	HotSPM	HotSPM
Gravure anisotrope	Ref	Trous	Trous	Trous	OK	OK	Conso latérale	Modif volume
Reprise épitaxie	Ref	/	/	/	OK	OK	/	Dépôt parasite

Figure 2.22 : Résumé des résultats obtenus dans ce chapitre. Vert = Bon, Orange = Mauvais, Rouge = Rédhibitoire.

Chapitre 3 – Epitaxie à basse température pour la surélévation des sources et drains

Introduction au Chapitre 3

La surélévation des sources et drains (S/D) est une étape essentielle dans les technologies basées sur des films minces comme l'est l'intégration sur substrat FDSOI. En effet, cette étape permet d'épaissir le film de silicium initial (6 à 10nm), évitant ainsi une siliciuration des accès jusqu'à l'oxyde enterré et/ou jusque sous les espaceurs de grille (effet d'*encroachment* latéral), formant alors une jonction Schottky avec le canal (i.e. une interface métal/semi-conducteur non ou faiblement dopé) ou bien encore une perforation des zones sources et drains lors de la gravure des contacts, deux effets non désirables qui dégraderaient les performances électriques (Figure 3.1).

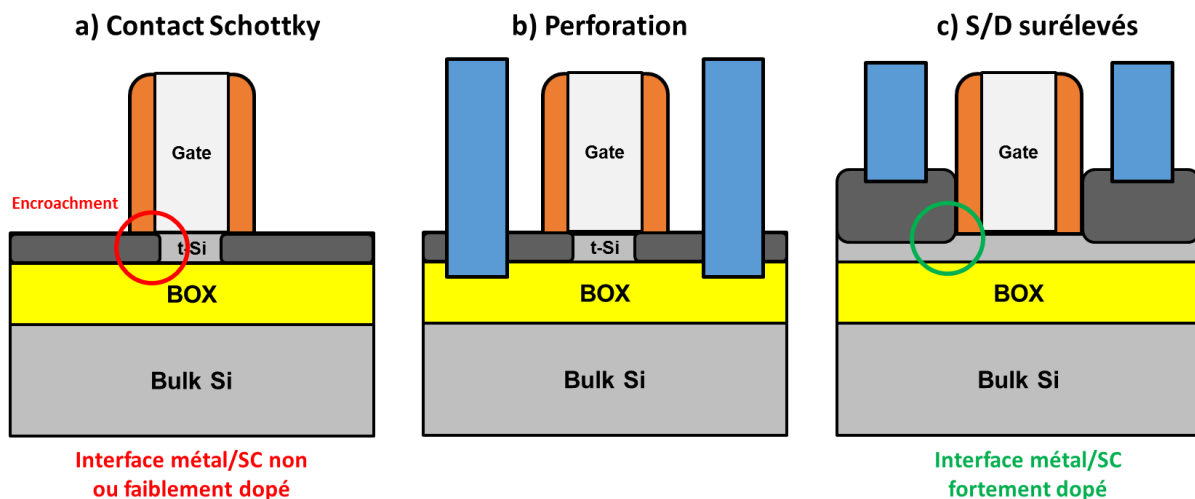


Figure 3.1 : Schémas des problèmes potentiels liés à la fabrication d'un transistor sur film mince : a) Siliciuration jusqu'au BOX et sous les espaceurs de grille créant une jonction Schottky (métal/semiconducteur); b) Perforation du film lors de la gravure des contacts; c) Cas idéal lorsque les sources et drains sont surélevés.

Aujourd'hui, cette surélévation est généralement effectuée par épitaxie, technique permettant de faire croître une couche tout en conservant la maille cristalline de la couche de départ (appelée *seed*). L'épitaxie permet alors *a priori* d'obtenir des couches sans défauts cristallins qui dégraderaient également les caractéristiques électriques des accès. De manière plus générale, l'épitaxie des zones sources et drains peut également être utilisée dans les technologies bulk afin d'introduire des contraintes sur le canal. Ainsi, précédée d'une gravure des zones d'accès, une épitaxie dite enterrée (*embedded*) de SiGe (imposant une contrainte compressive) ou de SiC (une contrainte tensile) par exemple permet d'optimiser le transfert de la contrainte au canal, améliorant alors la mobilité des trous dans les pMOS [Tamura08] ou des électrons dans les nMOS [Verheyen08] respectivement.

Cette épitaxie peut se faire de différentes manières, mais les deux techniques majoritairement utilisés sont l'épitaxie par jet moléculaire (MBE – *Molecular Beam Epitaxy*) et l'épitaxie en phase vapeur (CVD – *Chemical Vapor Deposition*). La technique MBE est basée sur la sublimation d'une source solide des éléments à épitaxier, puis du bombardement des atomes créés sur la surface du film. Cette technique requiert alors un vide très poussé ($\sim 10^{-10}$ Torr) afin d'éviter l'incorporation de tout contaminant dans la couche épitaxiée. L'épitaxie CVD elle est basée sur l'utilisation de précurseurs

gazeux ultra-purifiés injectés dans une chambre contenant la plaque. Cette technique peut être utilisée à différentes pressions, mais dans cette thèse, nous traiterons uniquement de la RPCVD (*Reduced Pressure CVD*) avec des pressions de dépôt aux alentours de 20 Torr. Le contrôle de la contamination est alors assuré par un pré-nettoyage adéquat de la surface à épitaxier ainsi que par l'utilisation d'une grande quantité de dihydrogène (H_2) pur dans lequel les précurseurs sont dilués pour être acheminés jusqu'à la chambre.

Encore une fois, dans le cadre de l'intégration 3D Séquentielle, un développement à basse température de procédés de surélévation des sources et drains est nécessaire afin de remplacer l'épitaxie sélective de Si intrinsèque standard faite à 750°C au Leti [Hartmann13a]. Cependant, ce développement est alors confronté à deux difficultés majeures : la préparation de surface et la croissance épitaxiale.

Préparation de surface :

Comme abordée dans la partie précédente, la reprise de l'épitaxie après gravure des espaceurs nécessite une surface initiale exempte de tout contaminant ou imperfections cristallines qui pourraient mener à des défauts morphologiques (e.g. apparition de dislocations) dans la couche épitaxiée, une dégradation des performances électriques (i.e. résistances des S/D plus élevées) voire entraver la croissance épitaxiale. De manière générale, l'état de surface avant épitaxie dépend de toutes les étapes précédentes, pouvant par exemple laisser des traces d'hafnium ou de titane (lors de la gravure grille), des résidus de diélectrique ou bien des polymères (après la gravure des espaceurs). Par ailleurs, la contamination par l'oxygène est également problématique due à la tendance du film de silicium à s'oxyder naturellement à l'air. Cet oxyde natif peut être désorbé grâce à un recuit sous atmosphère réductrice tel que le dihydrogène, créant alors des liaisons Si-H à la surface qui forment une couche de passivation du silicium. Dans ce cas, une température supérieure à 1000°C pendant quelques minutes est néanmoins nécessaire afin d'obtenir une surface parfaitement passivée [Hartmann16]. Cette température, outre l'incompatibilité avec la 3D Séquentielle, pose également problème dans les technologies à film mince, où des phénomènes de démouillage ou bien d'*islanding* (i.e. une formation d'îlots) peuvent apparaître lorsque le film est porté à trop haute température (Figure 3.2) [Jahan05].

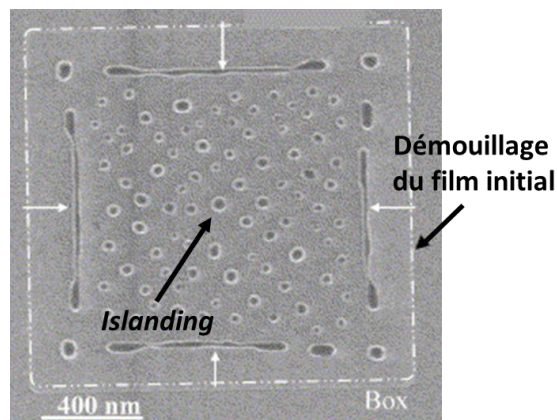


Figure 3.2 : Observation SEM montrant des phénomènes de démouillage (aux bords du film) et d'*islanding* d'un film de silicium de 8nm à la suite d'un recuit H_2 950°C pendant 2min (image pris de [Jahan05]).

Dans la pratique, l'oxyde natif est retiré par voie humide, notamment par un nettoyage à base d'acide hydrofluorique dilué (appelé « HF-last » lorsqu'il s'agit du dernier nettoyage humide avant croissance épitaxiale). Néanmoins, la passivation créée étant limitée dans le temps (une à deux heures maximum), un enchaînement rapide entre le nettoyage humide et l'épitaxie est alors nécessaire. De plus, nous verrons dans la Section 3.1.1 que la préparation de surface humide et le stockage dans un environnement inerte (e.g. les chambres de chargement/déchargement des plaques du bâti d'épitaxie

qui sont sous N₂ ultra-pure) ne sont pas suffisants à eux seuls et qu'un recuit H₂ à relativement haute température (T>650°C-750°C) reste nécessaire.

Croissance épitaxiale :

La croissance épitaxiale par RPCVD est également fortement impactée par la température de procédé. L'énergie thermique apportée par la température permet en effet non seulement de dissocier les précurseurs chimiques (des dérivés silanes dans le cas du silicium), mais aussi de casser les liaisons Si-H et de désorber l'hydrogène à la surface du film, libérant alors des sites pour l'adsorption des précurseurs. Une première conséquence est donc que la vitesse de croissance décroît exponentiellement lorsque la température diminue [Hartmann12]. En outre, l'énergie thermique permet également aux atomes adsorbés à la surface du film de se déplacer localement et de se réorganiser afin d'assurer une structure cristalline parfaite. La diminution de la température peut alors aussi mener à la formation de défauts cristallins voire à une perte de l'information cristalline du *seed* comme nous le verrons dans ce chapitre. Enfin la sélectivité de l'épitaxie est également un critère primordial. En effet, l'épitaxie doit être faite uniquement sur les zones actives cristallines afin d'éviter tout court-circuit, par exemple entre les accès et la grille si un dépôt se fait sur les espaceurs (cf. Chapitre 2). Ainsi, des chimies chlorées sont généralement utilisées afin de graver les couches amorphes ou polycristallines qui se déposent simultanément sur les zones diélectriques, tout en préservant les couches monocristallines sur les zones d'intérêt. En contrepartie, le rajout de ces chimies chlorées diminue également fortement la vitesse de croissance de la couche cristalline, complexifiant donc encore la faisabilité d'une épitaxie à la fois sélective et à basse température.

Ainsi, le développement d'une surélévation des S/D totalement compatible avec l'intégration 3D Séquentielle requiert une préparation de surface ainsi qu'une croissance épitaxiale à basse température. Nous verrons donc tout d'abord les problématiques d'une mauvaise préparation de surface sur la qualité de l'épitaxie dans la Section 3.1, avant de présenter une technique de nettoyage par plasma à basse température récemment développée. Ensuite, nous aborderons la question de la croissance épitaxiale à basse température de couches Si non dopé, Si dopé phosphore et SiGe dopé bore dans la Section 3.2, en introduisant notamment une stratégie d'épitaxie en deux temps avec un dépôt non sélectif suivie d'une gravure.

3.1 Préparation de surface

3.1.1 Préparation de surface classique (HF-last+Recuit H₂)

Aujourd'hui, la préparation de surface avant épitaxie se fait généralement par une combinaison de nettoyages humides terminée par une chimie HF (« HF-last ») et d'un recuit hydrogéné fait dans le bâti d'épitaxie juste avant la croissance épitaxiale. Les nettoyages humides sont régulièrement utilisés tout au long de l'intégration afin d'éliminer les particules ou les contaminants qui peuvent apparaître, notamment après gravure sèche ou après photolithographie. Ainsi, les gravures de la grille ou des espaceurs sont suivies d'une étape de retrait polymères. Les résines utilisées lors de la photolithographie avant implantation ionique sont quant à elles retirées par des étapes de stripping. Le choix des chimies humides est alors un compromis entre l'intégrité et la consommation de la zone active à nettoyer ainsi que des espaceurs de grille (cf. Chapitre 2) et la propreté de la surface nécessaire à une épitaxie de qualité. Par exemple, la Figure 3.3 montre les cartographies d'épaisseur de Si sur les

zones actives après épitaxie précédée de deux variantes de nettoyages humides. Dans le premier cas, l'étape de retrait polymères après gravure des espaceurs a été faite par un nettoyage HF/HCl seul, tandis que dans le deuxième cas, une ré-oxydation chimique (UV-O₃) de quelques monocouches de silicium a été réalisée en plus. Le même nettoyage HF-last a ensuite été effectué juste avant épitaxie pour les deux variantes. On voit alors qu'une consommation de quelques Angströms supplémentaires de la zone active avec l'oxydation chimique permet d'obtenir une surface plus propre menant à une épitaxie plus uniforme (i.e. 3σ plus faible), tandis que dans le premier cas, un retard au démarrage de la croissance épitaxiale apparaît à certains endroits de la plaque (en particulier vers le bord) dû à une contamination plus importante en surface, résultant en une épaisseur plus faible. Par la suite, l'étape de retrait polymères sera donc faite par une combinaison HF/HCl + UV-O₃ malgré une consommation légèrement plus importante des espaceurs de grille.

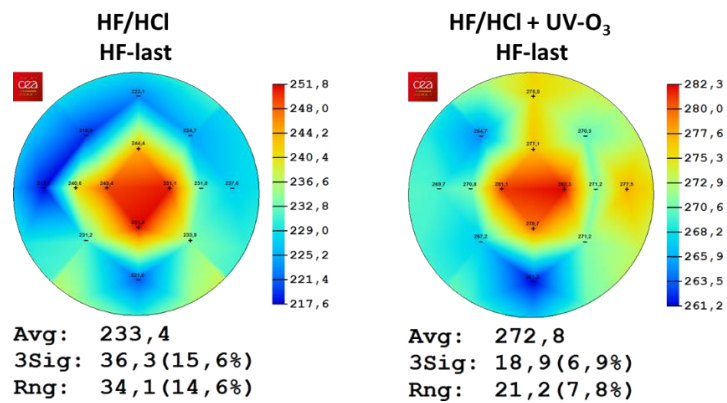


Figure 3.3 : Cartographies d'épaisseur de Si sur les zones actives après épitaxie S/D à 750°C précédée de deux variantes de nettoyages humides et du même recuit H₂ (650°C 2min+750°C 30s). Une sous-consommation du film initial peut dans certains cas entraver l'épitaxie, menant à une épaisseur épitaxiée plus faible et une non-uniformité plus importante (malgré le recuit hydrogéné subséquent à relativement haute température).

Après le nettoyage HF-last, un recuit H₂ reste néanmoins nécessaire afin d'éliminer les traces de contaminants restants. En effet, il a été montré dans [Meyerson90] qu'un nettoyage HF permet au mieux de passiver (i.e. création de liaisons Si-H) environ 85 % de la surface de silicium, des liaisons Si-F et Si-O constituant les 15 % restants. De plus, la remise à l'air entre l'équipement de nettoyage humide et le bâti d'épitaxie constitue également une source de réadsorption d'oxygène à la surface. Dans [Wostyn14], les auteurs ont ainsi montré qu'une température de recuit H₂ supérieure à 800°C (à une pression non spécifiée) était nécessaire afin de désorber quasiment tout l'oxygène présent à la surface (Figure 3.4a). Par ailleurs, la diminution de la pression du recuit hydrogéné à une température fixée permet également d'améliorer la préparation de surface (Figure 3.4b)

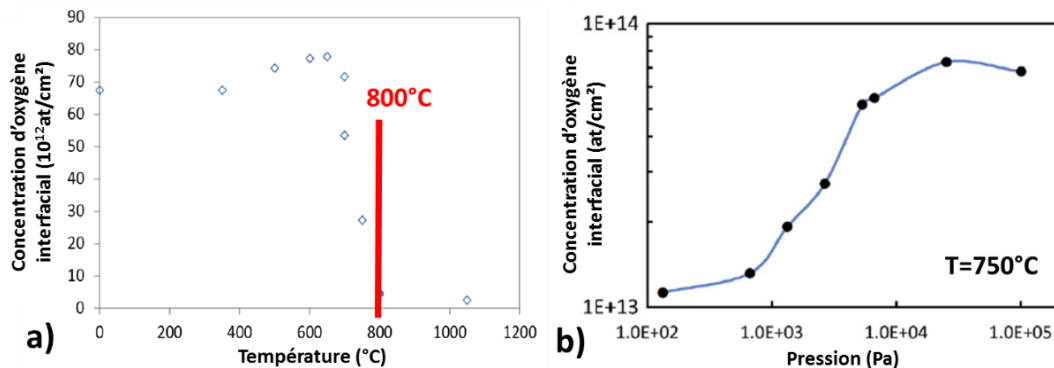


Figure 3.4 : Concentration d'oxygène à l'interface entre le film initial et la couche épitaxiée en fonction a) de la température (pris de [Wostyn14]) ou b) de la pression ([Hykavy16]). Une température élevée ou une pression faible permettent de diminuer la concentration d'oxygène contaminant en surface du silicium.

De plus, une préparation de surface parfaite est d'autant plus importante que la température de la croissance épitaxiale est faible. La Figure 3.5 montre par exemple des mesures AFM (*Atomic Force Microscopy* – technique donnant une image de la topographie de la surface de l'échantillon) après épitaxie de Si intrinsèque à 500°C précédée de recuits H₂ à différentes températures. Ainsi, diminuer simplement la température du recuit hydrogéné peut mener à une agglomération de l'épitaxie à basse température. En effet, l'énergie thermique lors de la croissance épitaxiale étant trop faible pour réorganiser les atomes, tous les défauts initialement présents à la surface sont révélés. Dans la pratique, le recuit hydrogéné généralement effectué au Leti est un recuit dit « hybride » à 650°C pendant 2min suivi d'une montée en température jusqu'à 750°C et d'une stabilisation à cette température pendant 30s, le tout à une pression réduite de 20 Torr.

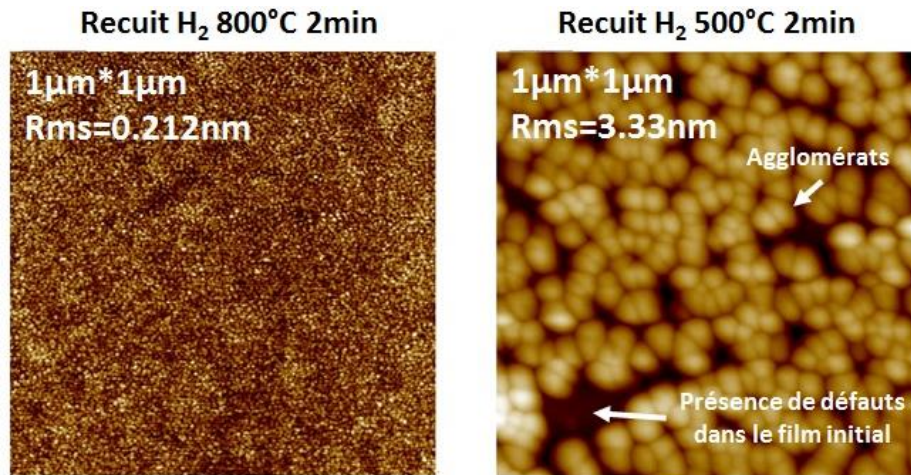


Figure 3.5 : Images AFM de couches de Si déposées à 500°C après nettoyage HF-last et un recuit H₂ haute (gauche) ou basse température (droite). Un recuit H₂ insuffisant mène à la formation d'îlots, les contaminants non désorbés occupant les sites d'adsorption à la surface.

Les techniques de préparation de surface classiques (i.e. nettoyage humide et recuit H₂) sont donc insuffisantes lorsque le budget thermique est limité, en particulier lorsqu'elles sont suivies d'une croissance épitaxiale à basse température. Une nouvelle méthode de préparation de surface par voie sèche a donc été étudiée.

3.1.2 Nettoyage par plasma : Siconi

Afin de nettoyer la surface de silicium efficacement à basse température, une technique de nettoyage par plasma a été développée par Applied Materials : la Siconi [Lei06]. Initialement mis au point pour le nettoyage du silicium avant l'étape de siliciuration, un procédé Siconi comporte principalement trois étapes (Figure 3.6). Un plasma de NH₃/NF₃ est d'abord créé dans une chambre séparée afin de dissocier ces molécules (1). Les produits sont ensuite transportés dans la chambre contenant la plaque et réagissent avec l'oxyde de silicium afin de former des sels solides à la surface (2). Ces sels sont ensuite sublimés à des températures qui peuvent être inférieures à 200°C afin de former des composants volatils, laissant alors le silicium exempt de tout oxyde natif et propre à une croissance épitaxiale (3).

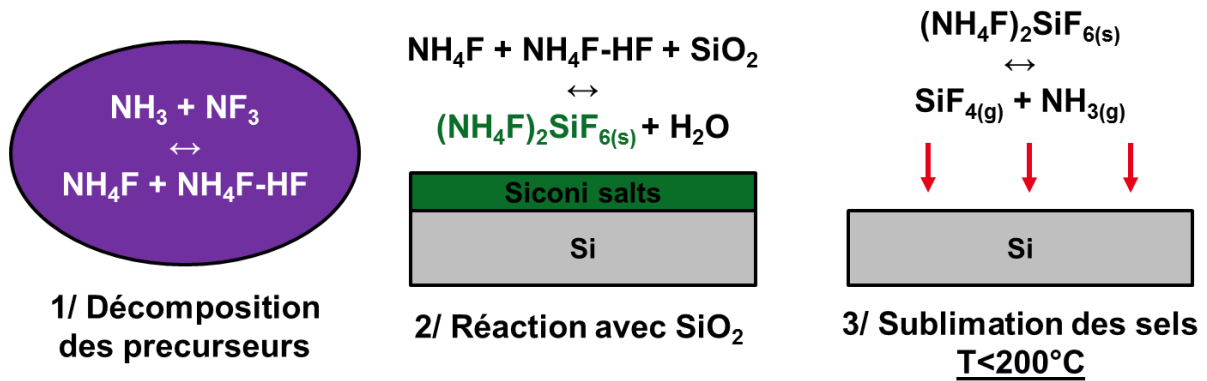


Figure 3.6 : Principes de la préparation de surface Siconi. 1) Les précurseurs NH_3 et NF_3 sont décomposés dans un plasma déporté; 2) Les produits réagissent avec le SiO_2 à la surface pour former un complexe solide et 3) Les sels Siconi sont désorbés à des températures pouvant être inférieures à $200^\circ C$ [Labrot16].

Un premier travail a tout d'abord été réalisé afin d'évaluer l'influence de la préparation de surface sur la qualité de l'épitaxie qui suit. Pour cela, une croissance de sources et drains SiGe dopé bore (SiGe:B) à $650^\circ C$ (température standard au Leti pour cette couche) a été réalisée et précédée de différentes combinaisons de nettoyages humide et/ou sec :

- Nettoyage humide seul (mélange HF/HCl).
- Nettoyage sec seul : Formation *ex-situ* des sels Siconi (réactions 1 et 2) puis sublimation des sels dans le bâti d'épitaxie (réaction 3).
- Nettoyage humide (HF/HCl + oxydation chimique) + Nettoyage sec (formation *ex-situ* des sels et sublimation dans le bâti d'épitaxie).
- Nettoyage sec « complet » (i.e. réactions 1 à 3 dans l'équipement de Siconi) + Nettoyage humide (HF/HCl).

Dans les variantes (b) et (c), un nettoyage Siconi dit « incomplet » a été réalisé, c'est-à-dire que la formation des sels a été réalisée dans un équipement *ex-situ* dédié, puis leur sublimation a été faite dans le bâti d'épitaxie. Au contraire, dans la variante (d), les trois étapes du nettoyage Siconi ont été réalisées *ex-situ*. En outre, dans la variante (c), une oxydation chimique (UV- O_3) de la surface de silicium a été rajoutée afin que la réaction (2) puisse avoir lieu.

Dans un premier temps, tous ces nettoyages ont été suivis du recuit hybride H_2 standard $650^\circ C$ 2min + $750^\circ C$ 30s (référéncé par la suite par $750^\circ C$ 30s par simplification). Dans les variantes (b) et (c), ce recuit H_2 sert alors également à sublimer les sels précédemment formés. La Figure 3.7 montre les observations SEM de dispositifs isolés après épitaxie pour chacune des variantes de nettoyages humide et sec.

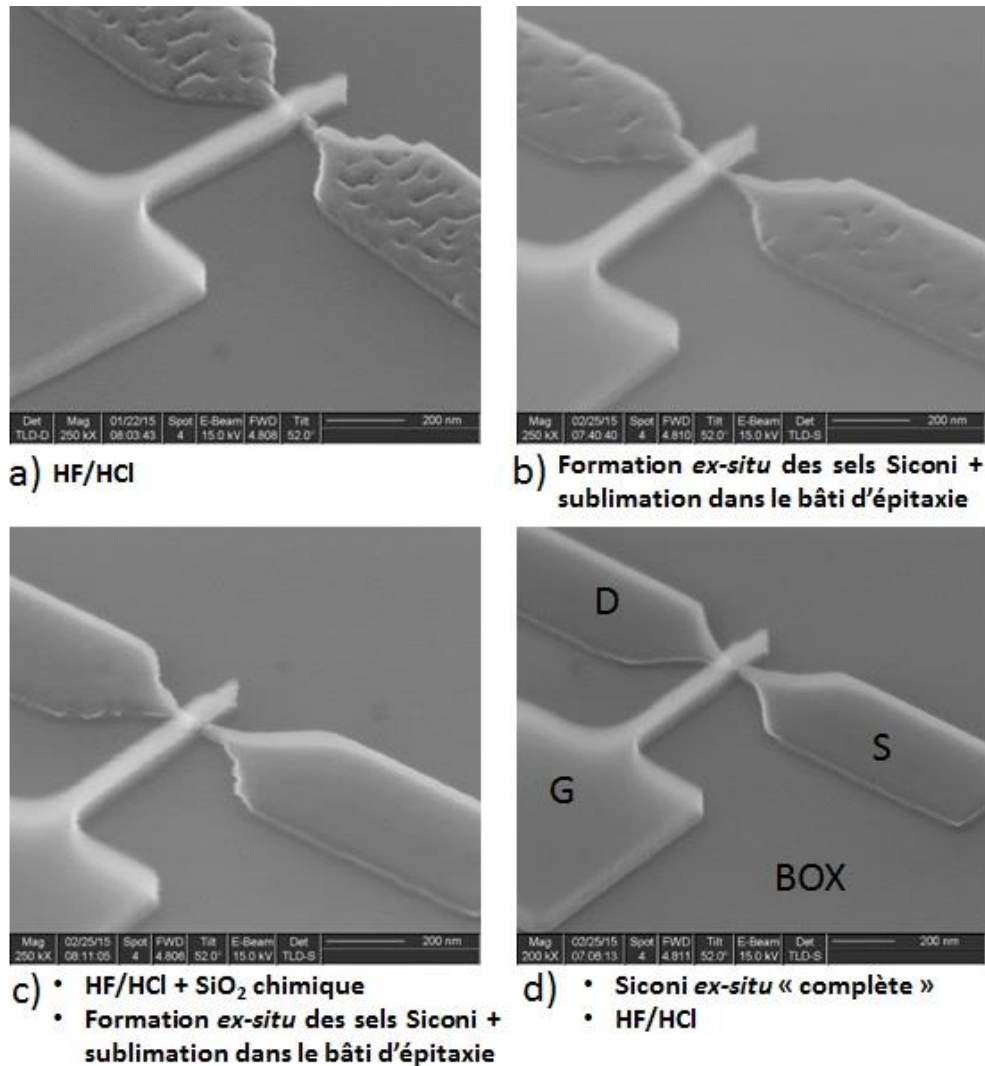


Figure 3.7 : Observations SEM après épitaxie sélective de S/D SiGe:B à 650°C précédée de différentes variantes de combinaisons de nettoyages humide/sec et d'un recuit H₂ hybride 650°C 2min+750°C 30s. La combinaison d'une Siconi et d'un nettoyage HF/HCl donne la meilleure préparation de surface (variante d).

Tout d'abord, nous voyons que malgré un recuit hydrogéné (750°C 30s) et une épitaxie SiGe:B (650°C) à relativement haute température, le nettoyage standard (HF-last + recuit H₂) peut s'avérer insuffisant dans certains cas (suivant l'état de surface laissé par les étapes précédentes) et mener à un effet de rainurage et donc une perte de la parfaite nature 2D de la couche épitaxiée (Figure 3.7a). Ces « trous » dans les accès peuvent alors poser problèmes lors de l'étape de siliciuration qui peut atteindre le BOX. De la même manière, l'utilisation d'un nettoyage Siconi « incomplet » seul mène également à la présence de défauts morphologiques (bien que moins prononcés que la variante précédente) à la surface des accès (Figure 3.7b). Au contraire, la combinaison des deux nettoyages HF puis Siconi « incomplet » permet effectivement d'améliorer la qualité de la préparation de surface, avec néanmoins quelques défauts subsistant aux bords des zones actives (Figure 3.7c). La préparation de surface optimale trouvée dans cette étude est donc une séquence nettoyage Siconi « complète » *ex-situ* et HF-last suivis du recuit hydrogéné (Figure 3.7d).

Cette étude révèle notamment la problématique de la remise à l'air de la surface de silicium entre les différents équipements. Si cette question a déjà été soulevée dans la section précédente entre un nettoyage HF-last et le passage dans le bâti d'épitaxie, nous voyons ici que la remise à l'air des sels de la Siconi diminue l'efficacité du nettoyage sec (variante c). Ce résultat est également obtenu dans [Labrot16], où il est montré que la stabilité du sel (NH₄)₂SiF₆ augmente avec le temps. Par conséquent,

un recuit de 800°C a été nécessaire pour désorber les sels après un délai de 48h sous atmosphère contrôlée (i.e. salle blanche), et de 1050°C après une exposition de 10min à l’atmosphère ambiante. Ainsi, afin d’avoir une préparation de surface parfaite, une séquence de nettoyages humides suivis d’une Siconi *in-situ* est clairement préférable, auquel cas le budget thermique du recuit hydrogéné peut être réduit afin d’être adaptée aux contraintes de la 3D Séquentielle (cf. Section 3.2.2).

Dans un second temps, le budget thermique du recuit H₂ a été diminué d’un recuit « hybride » (650°C 2min + 750°C 30s) à un recuit 650°C 2min seul en appliquant la séquence de nettoyage sec/humide optimale (variante d) trouvée précédemment. La Figure 3.8 montre que la diminution du budget thermique du recuit H₂ n’impacte pas visuellement la morphologie de la couche épitaxiée observée en SEM, avec toutefois une rugosité mesurée par AFM légèrement plus élevée dans le cas du recuit à 650°C.

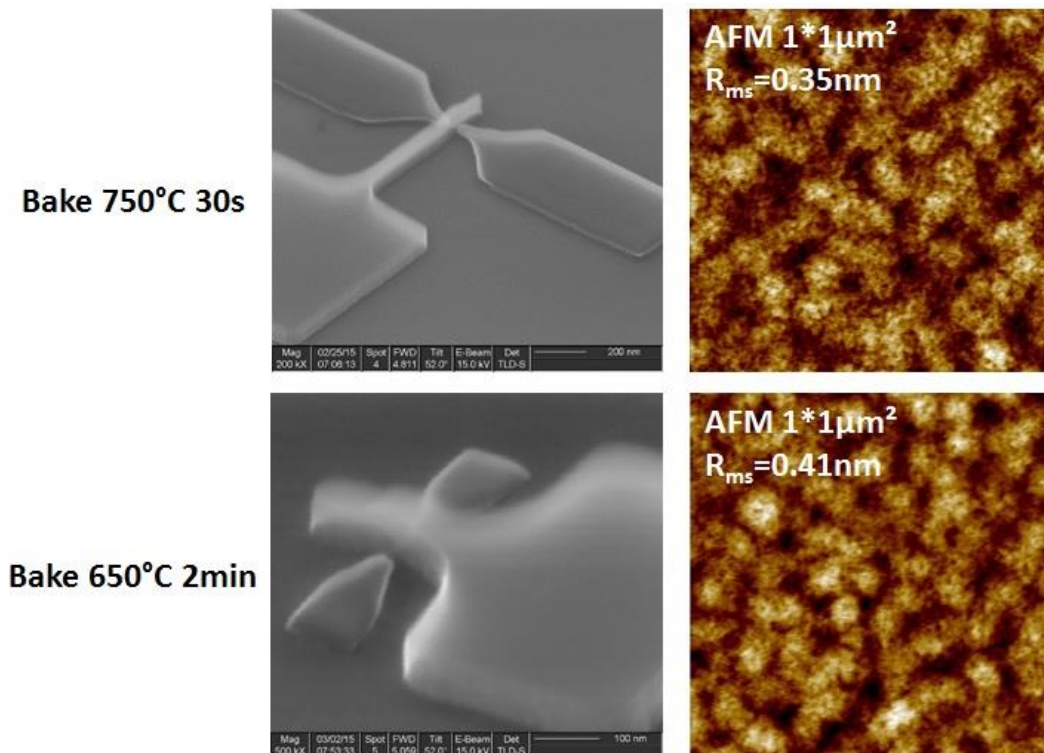


Figure 3.8 : Observations SEM et mesures AFM après épitaxie sélective de S/D SiGe:B à 650°C précédée de la combinaison optimale de préparation de surface humide/sec et d’un recuit hydrogéné à haute ou relativement basse température.

Ce résultat est confirmé électriquement, où les résistances d’accès et les mobilités des trous à champ faible extraites sur des transistors pMOS FDSOI sont équivalentes quel que soit le recuit H₂ (Figure 3.9). En particulier, le gain de mobilité des trous apporté par la contrainte compressive du SiGe:B et mis en évidence par la présence d’un pic pour les canaux courts (Lg≈100nm) est conservé. En conséquence, les performances électriques des transistors pMOS sont identiques en terme de compromis Ion/Ioff. Ainsi, la probable présence plus élevée de contaminants (e.g. oxygène, carbone, fluor) à l’interface entre le film initial et la couche épitaxiée dû à un budget thermique plus faible de recuit H₂ n’impacte pas forcément les caractéristiques électriques des dispositifs finaux.

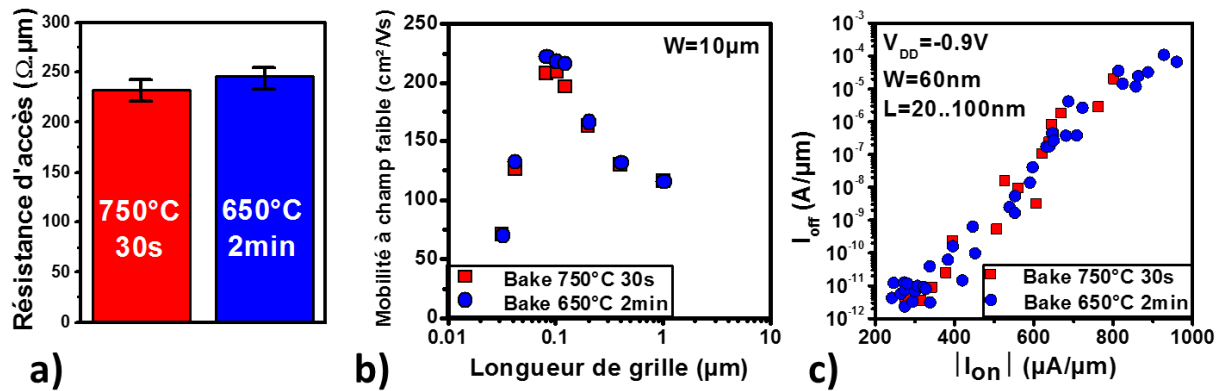


Figure 3.9 : Propriétés électriques associées aux transistors pMOS FDSOI fabriqués avec un recuit H_2 haute ou relativement basse température. a) Résistance d'accès et b) mobilité des trous à champ faible extraites par la méthode Y; c) Compromis $I_{\text{on}}/I_{\text{off}}$ mesuré sur des transistors étroits ($W=60\text{nm}$) [Lu16b].

Néanmoins, ce budget thermique réduit (650°C 2min) reste trop élevé pour l'intégration 3D Séquentielle. Si cette étude démontre la viabilité d'une Siconi *ex-situ* pour apporter plus de robustesse à la préparation de surface, une chambre *in-situ* intégrée au bâti d'épitaxie (mais donc plus coûteuse) est nécessaire afin de diminuer le budget thermique en dessous de 500°C.

3.2 Croissance épitaxiale

3.1.3 Epitaxie sélective vs dépôt/gravure cyclée

Comme abordé dans l'introduction, un premier problème lié à la diminution de la température est la décroissance exponentielle de la vitesse de croissance de la couche épitaxiée. Par exemple, la Figure 3.10 montre que la diminution de la température de 750°C à 600°C dégrade d'un facteur 100 la vitesse de croissance du silicium avec l'utilisation du silane (SiH_4) comme précurseur (en bleu).

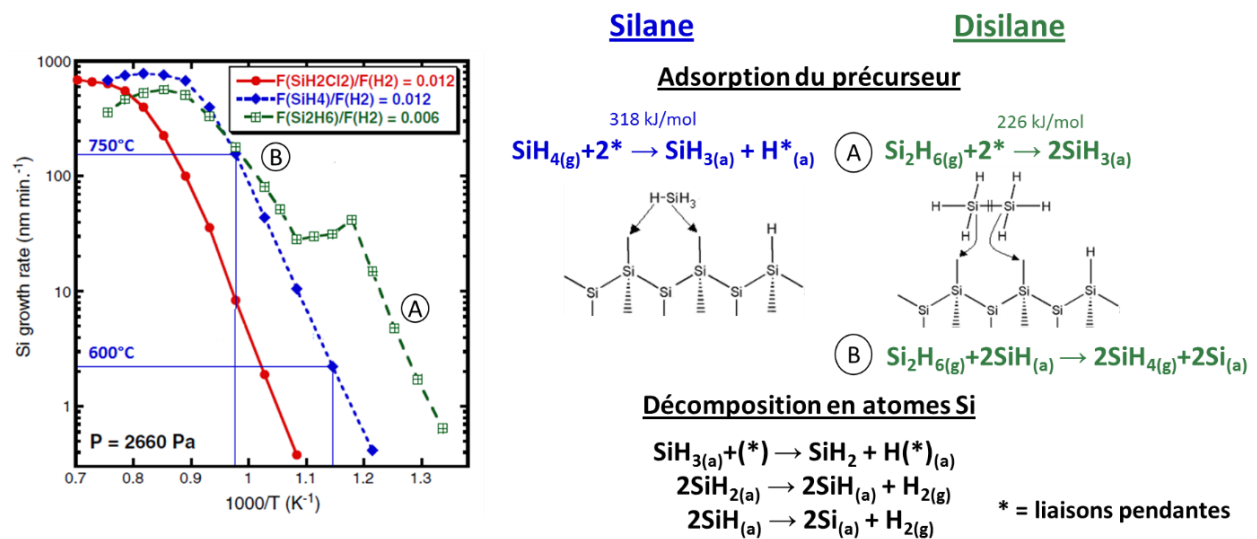


Figure 3.10 : Vitesse de croissance d'une épitaxie Si intrinsèque en fonction de la température et du précurseur (pris de [Hartmann12]). La différence de vitesse entre le silane et le disilane peut être expliquée par les réactions mises en jeu lors de leur décomposition jusqu'à l'obtention d'atomes de Si à la surface du film [Chung10].

Une solution à ce problème est l'utilisation d'un précurseur silicium d'ordre plus élevé, par exemple du disilane (Si_2H_6). En effet, la décomposition du disilane en radicaux silyles ($-\text{SiH}_3$) passe par la cassure d'une liaison Si-Si ($E_{\text{Si-Si}}=226\text{kJ/mol}$), réaction nécessitant moins d'énergie que la cassure d'une liaison Si-H ($E_{\text{Si-H}}=318\text{kJ/mol}$) dans le cas du silane, permettant donc des vitesses de croissance plus élevées (en vert) lorsque la température est inférieure à 550°C (région A). En revanche, lorsque la température est supérieure à 650°C (région B), les vitesses de croissance avec du silane et du disilane tendent à être équivalentes. Ceci est expliqué par le fait qu'à plus haute température, une autre décomposition possible du disilane apparaît, décomposition formant des molécules silanes, d'où le comportement similaire. Le plateau entre les deux régions correspond alors à la compétition entre ces deux réactions. Pour aller plus loin, des précurseurs d'ordre encore plus élevé (trisilane Si_3H_8 [Gouyé09], tetrasilane Si_4H_{10} [Kanoh93] voire neopentasilane Si_5H_{12} [Chung08]) ont été considérés dans la littérature. Néanmoins, si leur utilisation améliore effectivement la vitesse de croissance du silicium, leur coût plus élevé (du fait de la complexité à les produire et à les purifier) les rend moins viables industriellement. De plus, ces précurseurs se trouvent sous forme liquide à pression et température ambiantes, nécessitant alors un équipement dédié pour leur acheminement jusqu'à la chambre contenant la plaque. Ainsi, dans notre cas, l'utilisation du disilane reste suffisante pour bénéficier de la faiblesse des liaisons Si-Si comparé au silane.

Par ailleurs, l'obtention d'une épitaxie intrinsèquement sélective passe par l'utilisation obligatoire de chimies chlorées afin de graver les matériaux non cristallins qui se déposent simultanément sur les surfaces diélectriques (oxyde d'isolation et espaceurs de grille). Cependant, comme le montre également la Figure 3.10, l'utilisation de dichlorosilane (SiH_2Cl_2) par exemple dégrade fortement la vitesse de croissance du silicium (en rouge). Ceci est notamment expliqué par la tendance du chlore à créer des liaisons Si-Cl à la surface qu'il faut désorber afin de libérer des sites d'adsorption et qui nécessitent une énergie plus importante pour être cassées ($E_{\text{Si-Cl}}=376\text{kJ/mol}$) comparé à la désorption d'hydrogène ($E_{\text{Si-H}}=196\text{kJ/mol}$). La même problématique est trouvée avec l'utilisation de HCl nécessaire pour être sélectif vis-à-vis des espaceurs où l'augmentation de son débit diminue la vitesse de croissance d'un facteur environ 2 pour une couche de SiGe épitaxiée à 650°C par exemple (Figure 3.11a). Ainsi, dans la pratique, **le procédé standard d'épitaxie Si sélective est réalisé à 750°C (SiH_2Cl_2 + HCl) au Leti** afin d'obtenir une vitesse de croissance convenable [Hartmann13a].

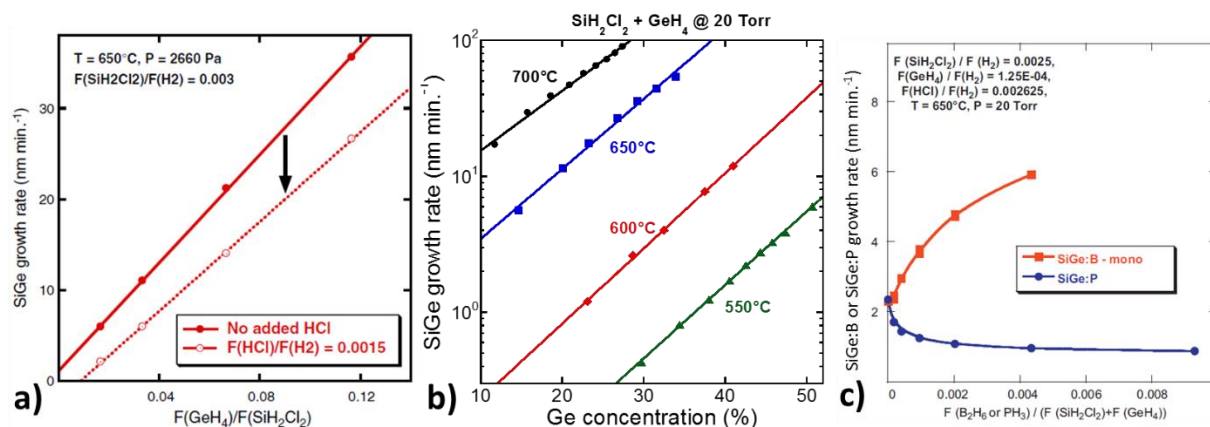


Figure 3.11 : a) Influence de l'ajout de HCl sur la vitesse de croissance du SiGe en chimie $\text{SiH}_2\text{Cl}_2+\text{GeH}_4$; b) Vitesse de croissance du SiGe en fonction de la température et de la concentration de Ge (ou du flow de GeH_4); c) Influence de l'ajout de bore ou de phosphore sur la vitesse de croissance du SiGe:B ou SiGe:P (pris de [Hartmann12, Hartmann11]).

Une autre possibilité afin d'améliorer la vitesse de croissance (en particulier lors de l'utilisation du dichlorosilane) est l'ajout de germane (GeH_4). En effet, les atomes de germanium ont un effet catalytique sur la désorption de l'hydrogène voire du chlore à la surface du film, les liaisons Ge-H/Ge-

Cl ($E_{Ge-H}=155\text{kJ/mol}$ et $E_{Ge-Cl}=213\text{kJ/mol}$) nécessitant une énergie plus faible pour être cassées que les liaisons Si-H/Si-Cl. Par conséquent, l'augmentation du flux de GeH_4 à une température donnée augmente la vitesse de croissance du SiGe, et par la même la concentration en germanium contenue dans la couche, augmentant ainsi la contrainte compressive sur le canal et donc la mobilité des trous pour les pMOS (Figure 3.11b). Les valeurs absolues de la Figure 3.11b sont néanmoins à relativiser dans la mesure où, comme pour le Si intrinsèque, l'ajout de HCl reste nécessaire afin d'obtenir la sélectivité désirée sur plaques patternées, ce qui a donc pour effet de diminuer à nouveau les vitesses de croissance (Figure 3.11a). Ainsi, **le procédé standard d'une épitaxie source et drain SiGe sélective sera faite à 650°C ($\text{SiH}_2\text{Cl}_2 + \text{GeH}_4 + \text{HCl}$) au Leti [Hartmann13a]**. Similairement, l'ajout de diborane (B_2H_6) aux autres précurseurs permet non seulement de doper la couche de SiGe et de diminuer sa résistivité, mais aussi d'augmenter la vitesse de croissance pour les mêmes raisons que pour le germanium (Figure 3.11c). Cependant, un flux trop important de diborane peut mener à une dégradation de la structure cristalline de la couche épitaxiée, ce sur quoi nous reviendrons dans la Section 3.2.3.2. Ainsi, **le procédé standard d'épitaxie SiGe:B sélective sera également réalisée à 650°C ($\text{SiH}_2\text{Cl}_2 + \text{GeH}_4 + \text{B}_2\text{H}_6 + \text{HCl}$) au Leti [Hartmann13a]**. Enfin, l'ajout de phosphine (PH_3) diminue au contraire la vitesse de croissance, ceci pouvant être dû à un phénomène de ségrégation des atomes de phosphore à la surface, empêchant l'adsorption des éléments à épitaxier [Hartmann11].

La faisabilité d'une croissance épitaxiale à la fois sélective et à basse température par une stratégie dite « co-flow » où les gaz de dépôt et de gravure sont injectés en même temps dans la chambre est donc principalement limitée par le compromis entre une bonne vitesse de croissance, la préservation d'une sélectivité entre zones actives et diélectriques et l'obtention d'une bonne qualité cristalline de la couche. Comme nous l'avons vu dans le paragraphe précédent, il est donc difficile de descendre en dessous de 650-750°C pour l'étape de surélévement des S/D avec cette stratégie co-flow tout en restant sélectif. Une solution est donc de séparer les étapes de dépôt et de gravure sélective en deux étapes distinctes. Dans un premier temps, un dépôt est réalisé sur toute la surface de la plaque de manière non sélective. Dans un second temps, les précurseurs de dépôt sont purgés et le gaz de gravure sélective (par exemple du HCl) est injecté afin de retirer les matériaux non cristallins sur les diélectriques (Figure 3.12). Ces deux étapes peuvent alors éventuellement être répétées plusieurs fois afin d'atteindre l'épaisseur ciblée. Ce passage à une épitaxie dite cyclée (CDE – *Cyclic Deposition/Etch*) permet alors l'utilisation de précurseurs plus favorables en termes de vitesse de croissance lors du dépôt (i.e. disilane), ceci même s'ils ne sont pas sélectifs vis-à-vis des matériaux diélectriques.

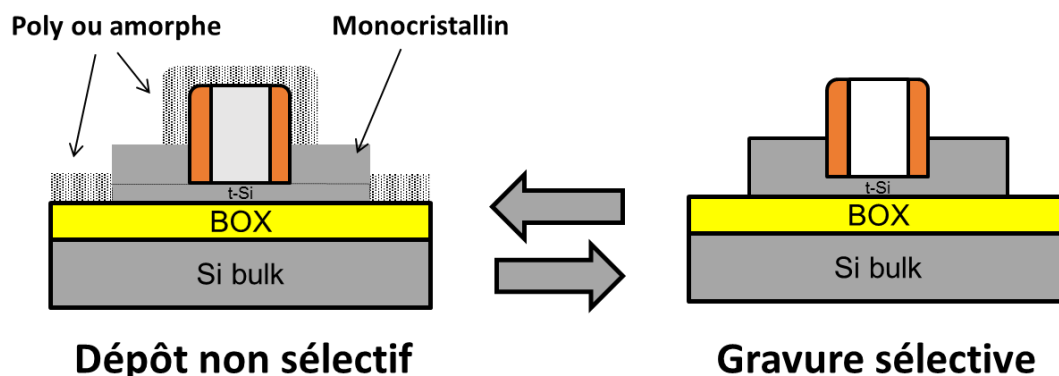


Figure 3.12 : Schéma du principe d'une épitaxie par dépôt non sélectif suivie d'une gravure sélective. Ces deux étapes peuvent être répétées pour obtenir l'épaisseur ciblée (épitaxie Cyclic Deposition/Etch).

Cette stratégie CDE a été initialement utilisée pour l'épitaxie de $\text{Si}_{1-y}\text{C}_y$ intrinsèque voire dopée phosphore (SiC:P) permettant d'apporter une contrainte tensile sur le canal qui améliore la mobilité des électrons dans les nMOS. L'incorporation de carbone en position substitutionnelle est en effet

fortement limitée par sa faible solubilité dans le silicium et sa tendance à former des nano-précipités de SiC [Bauer07]. De plus, la croissance épitaxiale de $\text{Si}_{1-y}\text{C}_y$ est aussi compliquée par la forte différence de paramètre de maille entre le carbone ($a_c=3.567\text{\AA}$) et le silicium ($a_{\text{Si}}=5.431\text{\AA}$). Ainsi, si la croissance d'une fine épaisseur de $\text{Si}_{1-y}\text{C}_y$ est faisable, donnant naissance à une couche en tension (proportionnelle à la concentration de C), l'accumulation de l'énergie élastique peut mener à une relaxation de la contrainte (i.e. à l'apparition d'ondulations en surface de la couche) et/ou à l'apparition de dislocations lorsqu'une épaisseur critique est dépassée [Iyer92, Hartmann13]. Une solution est alors de croître hors équilibre thermodynamique, ce qui permet d'obtenir une couche métastable de $\text{Si}_{1-y}\text{C}_y$ de bonne qualité avec une quantité relativement élevée de carbone (typiquement entre 1 et 2%). Dans la pratique, cela passe par une croissance très rapide de la couche et à basse température pour éviter les différents problèmes cités [Iyer92], ce qui n'est donc pas possible avec une stratégie « co-flow » et des précurseurs chlorés. L'utilisation d'une stratégie CDE permet alors l'utilisation de précurseurs hydrogénés ($\text{Si}_3\text{H}_8@550^\circ\text{C}$ [Bauer06] ou $\text{Si}_2\text{H}_6@550^\circ\text{C}$ [Hartmann13b] par exemple) caractérisés par des vitesses de croissance importantes à basse température pour le dépôt, combinée à une gravure sélective ($\text{Cl}_2@550^\circ\text{C}$ [Bauer06] ou HCl voire $\text{HCl}+\text{GeH}_4@600^\circ\text{C}$ [Hartmann13b]) pour les applications de surélévation des sources et drains.

Ainsi, la stratégie d'épitaxie par dépôt/gravure séparés sera adoptée pour toutes nos études de surélévation des sources et drains à basse température. Ceci étant dit, le développement est alors double. D'un côté, le dépôt sur surface monocristalline (i.e. les zones actives) doit être de bonne qualité (i.e. croissance monocristalline sans défauts dans la couche) tout en gardant une vitesse de croissance correcte. De l'autre, la vitesse de gravure de ce même dépôt mais sur des surfaces diélectriques doit être significativement plus rapide que celle sur du silicium monocristallin, c'est-à-dire que la sélectivité de la gravure doit être élevée. Ceci passera donc notamment par une optimisation des quantités de gaz injectés, de leurs pressions partielles et du temps de chaque étape. D'autres paramètres importants doivent être également pris en compte (mais nous ne nous focaliserons pas forcément dessus dans ce manuscrit) :

Contrôle de la température :

La mesure de la température du wafer en temps réel est effectuée par des pyromètres en différents points de la plaque. Cependant, la présence des gaz dans la chambre peut perturber cette mesure. Par exemple, la Figure 3.13 montre une variation de la température mesurée par un pyromètre sur la face avant de la plaque d'une dizaine de degré lors d'une épitaxie CDE selon si la chambre contient du H_2 (lors des cycles de dépôt) ou du N_2 (lors de la gravure).

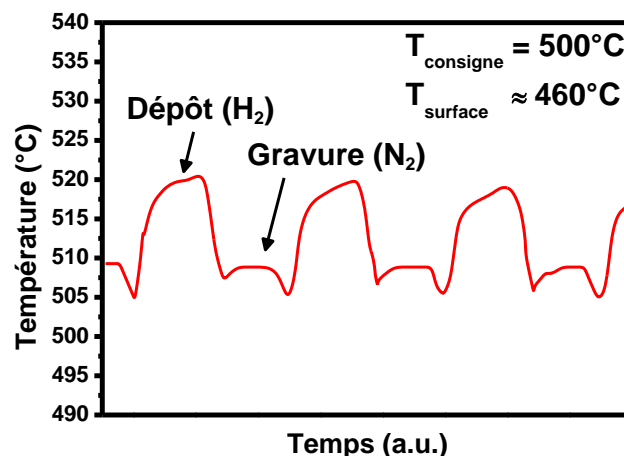


Figure 3.13 : Evolution de la température mesurée par un pyromètre en face avant de la plaque. Le changement de gaz entre du H_2 (lors du dépôt) et du N_2 (lors de la gravure) perturbe la mesure et introduit une variation d'une dizaine de degrés (données AMAT).

Par ailleurs, les pyromètres sont parfois calibrés à des températures supérieures à celle d'utilisation et peuvent donc introduire un décalage lors de la mesure de températures plus faibles. Dans le cas précédent, une température de consigne de 500°C et mesurée à environ 515°C par le pyromètre correspond à une température de surface « réelle » aux alentours des 460°C. Enfin, la présence d'un oxyde enterré peut également modifier l'absorption d'énergie provenant d'une source externe (e.g. des lampes chauffantes) et mener à une différence de température plus importante avec celle de consigne. Ce contrôle de la température est crucial d'une part en termes d'homogénéité dans l'épaisseur de la couche épitaxiale sur une plaque, mais aussi dans le développement de procédés à très basse température comme nous le verrons par la suite.

Orientation cristalline :

La vitesse de croissance peut également varier selon l'orientation cristalline. Si la vitesse de croissance du silicium ne varie que très légèrement avec l'orientation du silicium initial (Figure 3.14a), celle du SiGe sur (100) – correspondant à la direction de croissance des sources et drains – peut être plusieurs fois plus élevée que celle sur (110) – axe perpendiculaire aux flancs des espaceurs (Figure 3.14b). Cet écart peut alors mener à l'apparition de facettes selon l'orientation (111) lors de la croissance du SiGe (Figure 3.14c). Si ces facettes peuvent poser problème lors de l'implantation ionique, une amorphisation totale du film de silicium pouvant se produire à cet endroit, elles peuvent également être bénéfiques dans la réduction des capacités parasites entre les accès et la grille [Cheng09].

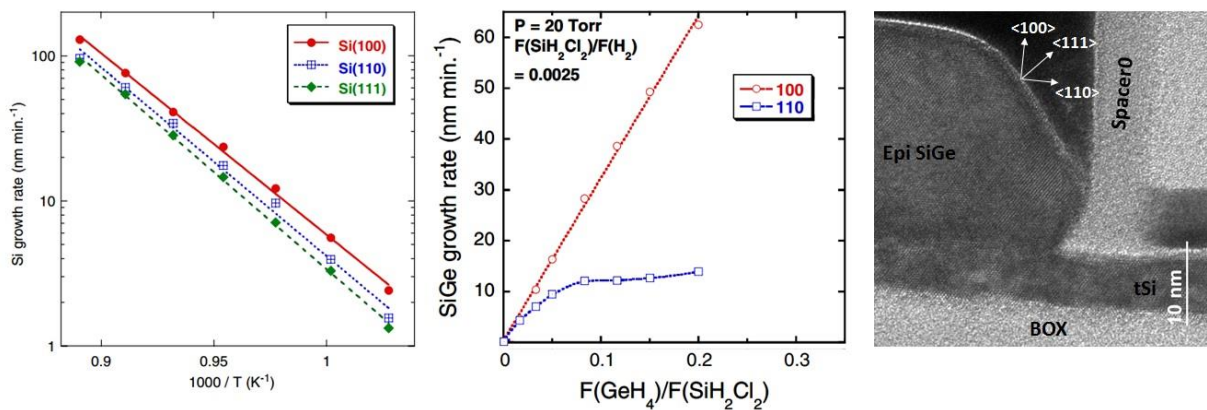


Figure 3.14 : Influence de l'orientation cristalline de la surface du film initial sur la vitesse de croissance de a) Si intrinsèque (pris de [Hartmann06]) et b) SiGe intrinsèque [Destefanis10]); c) Observation TEM montrant l'apparition d'une facette selon l'axe <111> lors d'une épitaxie sélective de S/D SiGe à 650°C.

Taux d'ouverture :

Le taux d'ouverture représente la proportion de zones actives en silicium par rapport aux zones diélectriques sur une plaque masquée. Ce taux change notamment en fonction de la technologie et du masque utilisé, où le nombre de dispositifs et leurs dimensions peuvent largement varier. Une conséquence est alors une différence de vitesse de croissance et/ou de gravure, par exemple entre des études sur pleine plaque (taux=100%) et sur plaques patternées (taux≈30% sur les masques du Leti), nécessitant donc une adaptation des temps de procédés.

3.1.4 Epitaxie Si intrinsèque

Dans un premier temps, intéressons-nous à l'épitaxie de silicium non dopé à très basse température, épitaxie qui permet à terme de réaliser des transistors de type n et p sur la même plaque en dopant les accès par implantation ionique. Les études préliminaires de dépôt et de gravure sélective sur pleines plaques ont été faites par Applied Materials afin de trouver des points de fonctionnement qui ont ensuite été appliqués sur des plaques patternées caractérisées au Leti. Ces développements ont notamment été faits sur un bâti d'épitaxie intégrant une chambre Siconi. Sur toutes les variantes, la préparation de surface a donc été réalisée par un nettoyage humide (HF-last) suivi d'une Siconi *in-situ*, ce qui permet de s'affranchir totalement d'un recuit hydrogéné à haute température – le seul recuit H₂ effectué étant durant le chargement des plaques dans la chambre qui est fait à la température de l'épitaxie. Ce bâti d'épitaxie permet également l'utilisation de chlore (Cl₂) pour la gravure sélective. Si l'utilisation du Cl₂ comporte des contraintes logistiques (nécessité d'une ligne d'acheminement dédiée et utilisation de N₂ au lieu de H₂ comme gaz vecteur afin d'éviter la formation de HCl), il offre des vitesses de gravure à basse température plus élevées. En effet, la vitesse de gravure du silicium avec du HCl décroît également exponentiellement avec la température, atteignant alors des valeurs trop faibles sous 500°C pour être utilisé dans une stratégie CDE (Figure 3.15a). Par ailleurs, l'ajout de GeH₄ au HCl permet d'augmenter la vitesse de gravure du Si d'un rapport 12 à 36 (Figure 3.15b). Ceci peut être expliqué par le fait que le germane conduit à l'incorporation d'atomes germanium dans le silicium, formant quelques mono-couches de SiGe. Or, comme vu dans la Figure 3.15a, la gravure par HCl est d'autant plus rapide que la concentration de germanium est élevée, le Ge agissant là aussi comme un catalyseur. Toujours est-il que même avec l'ajout de GeH₄, les vitesses de gravure du silicium restent trop faibles à très basse température [Hartmann13b]. Au contraire, l'utilisation du Cl₂ permet d'obtenir des vitesses de gravure du silicium monocristallin de l'ordre du nm/min à une température inférieure à 500°C (Figure 3.15c). De plus, la vitesse de gravure du silicium amorphe de l'ordre de la dizaine de nm/min devrait permettre une bonne sélectivité de gravure.

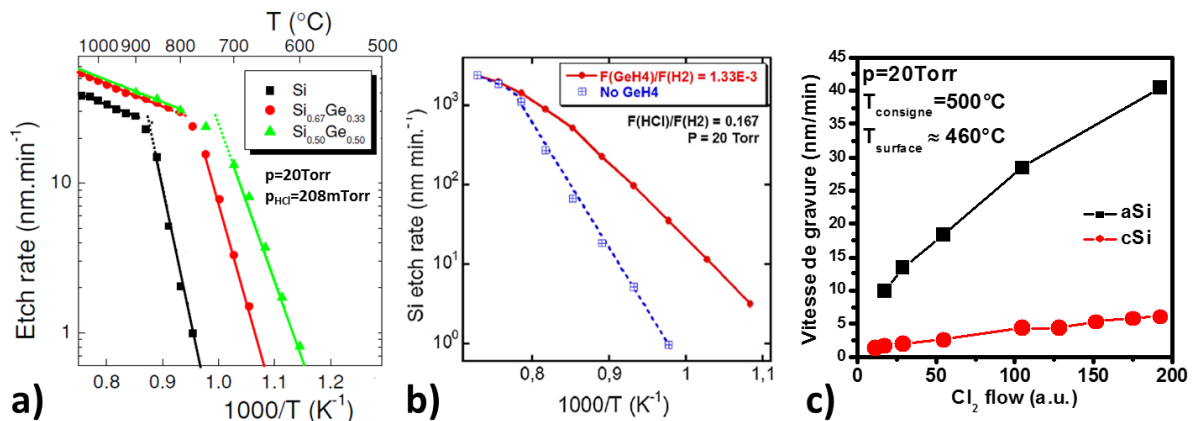


Figure 3.15 : a) Vitesse de gravure HCl de Si et Si_{1-x}Ge_x intrinsèques en fonction de la température (pris de [Bogumilowicz05]); b) Influence de l'ajout de GeH₄ au HCl sur la vitesse de gravure de Si intrinsèque ([Hartmann13b]); c) Vitesse de gravure de Si amorphe (aSi) et monocristallin (cSi) en fonction du flow de Cl₂ à 500°C (Données AMAT).

Le premier essai d'intégration d'une épitaxie Si non dopée à très basse température a été fait à 568°C (température de consigne, la température « réelle » de la surface du wafer étant plutôt de 500°C) à une pression de 20 Torr (Figure 3.16). A noter que ces premières études ont été réalisées sur une technologie 14FDSOI avec un film de silicium de 6nm et un BOX de 20nm (cf. Chapitre 2). Les transistors pMOS sont donc protégés et ne sont pas épitaxiés. Si la sélectivité vis-à-vis des isolations, de la grille

et du nitrure ainsi que la cristallinité des accès sont bonnes, la surface de l'épitaxie est fortement rugueuse.

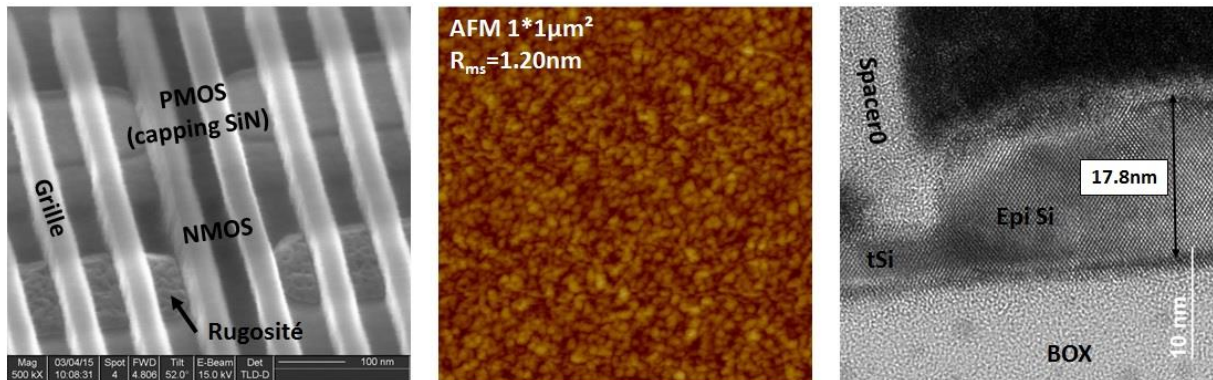


Figure 3.16 : Observations SEM (gauche) et TEM (droite) sur des oscillateurs en anneau et mesure AFM (milieu) après une épitaxie CDE (12 cycles - Dépôt @20Torr/Etch @20Torr) de S/D Si à $T_{consigne}=568^\circ\text{C}$ ($T_{surface}=500^\circ\text{C}$). Aucune croissance ne se produit sur les zones pMOS qui sont recouvertes par du SiN.

Ce résultat met en lumière deux difficultés. La première est l'obtention d'un dépôt parfaitement monocristallin à très basse température. En effet, l'énergie thermique permet entre autres de réorganiser localement les atomes à la surface afin d'obtenir le même arrangement cristallin que celui du silicium initial. Ainsi, le dépôt de silicium initialement monocristallin à basse température peut devenir amorphe au cours de la croissance, en particulier si la vitesse de croissance est trop rapide. Une autre difficulté associée est le contrôle de la gravure sélective avec le Cl_2 , plus agressif que le HCl , et qui tend donc à révéler les défauts cristallins de la couche menant à une rugosité plus importante. Dans le pire des cas, le film de silicium initial peut même être entièrement consommé lors des étapes successives de gravure (Figure 3.17). Une fois le BOX atteint, la couche déposée à chaque cycle de dépôt est alors entièrement amorphe avec une vitesse de croissance importante et n'est ensuite pas totalement gravée, d'où la présence d'une couche amorphe résiduelle après le cycle final de gravure. A noter aussi qu'au bord de la plaque, une consommation plus importante mais non totale du film près des grilles apparaît et qui peut être liée à un effet de densité.

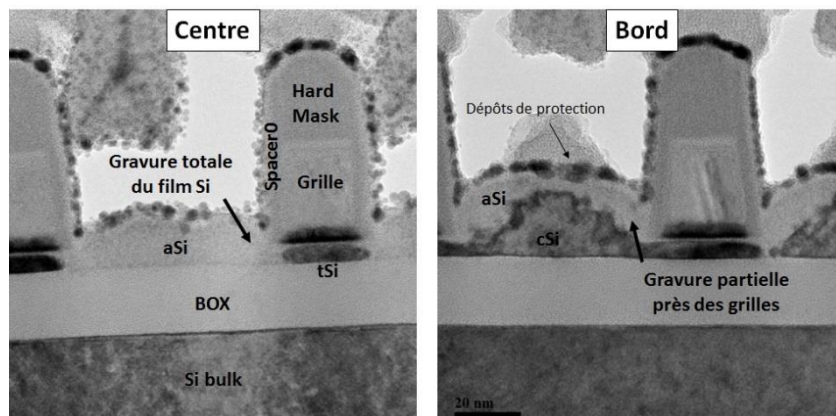


Figure 3.17 : Observations TEM en section transverse après épitaxie CDE (12 cycles - Dépôt @20Torr/Etch @20Torr) de S/D Si à $T_{consigne}=568^\circ\text{C}$. Le film initial de silicium (6nm) a été totalement/partiellement gravé sur les motifs au centre/bord de la plaque.

Une optimisation des débits des différents gaz permet tout de même d'obtenir un premier point de fonctionnement donnant une excellente rugosité de la surface des accès tout en préservant la sélectivité et la qualité cristalline (Figure 3.18). Pour ce point, des temps très faibles (inférieurs à 10s) ont été utilisés pour le dépôt et la gravure menant à une épaisseur de silicium cristallin déposé de 1.68nm par cycle lors des études sur pleine plaque. L'épaisseur plus faible de 1.41nm/cycle (16.9nm

déposé au total au bout des 12 cycles) sur plaque patternée est dû à un taux d'ouverture moins important.

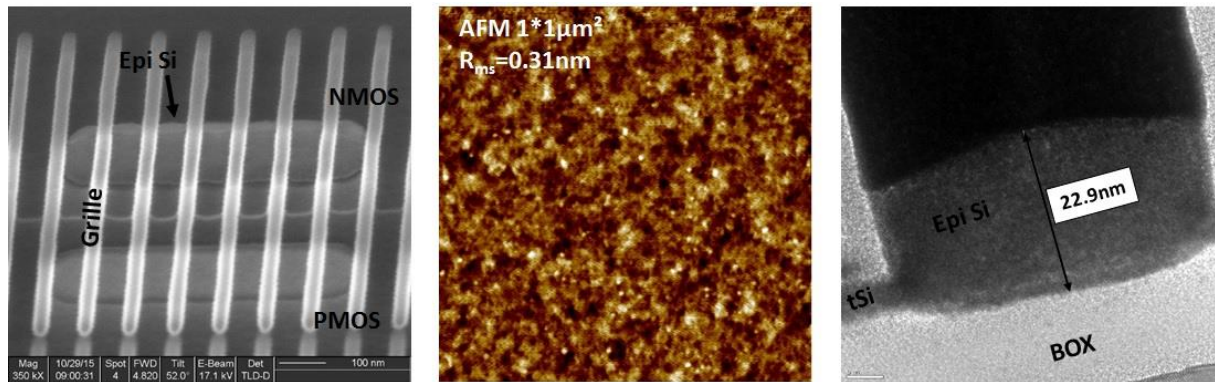


Figure 3.18 : Observations SEM (gauche), TEM (droite) et mesure AFM (milieu) après une épitaxie CDE (12 cycles - Dépôt @20Torr/Etch @20Torr) de S/D Si à $T_{consigne}=568^{\circ}\text{C}$.

Cependant, un problème avec un tel procédé où les temps très courts de dépôt et de gravure sont à la limite de l'équipement, est sa robustesse vis-à-vis des variations des conditions d'épitaxie. En particulier, le conditionnement de la chambre d'épitaxie peut dériver dû à son utilisation pour le dépôt de diverses couches à différentes températures ou différentes pressions au cours du temps. Par exemple, les mêmes conditions d'épitaxie réutilisées 6 mois plus tard ont mené à une épaisseur déposée plus importante de 2.14nm/cycle (contre 1.68nm/cycle précédemment) sur pleine plaque et qui peut être attribué à une température de la chambre plus importante qu'attendu, ce qui nécessite de réadapter le procédé.

Dans le but d'être plus robuste, les développements suivants ont été effectués avec des temps de dépôt et de gravure plus longs. Pour cela, la température a été diminuée à 500°C ($T_{surface}=460^{\circ}\text{C}$) afin de diminuer la vitesse de croissance du silicium et rallonger le temps de dépôt, celle-ci restant de l'ordre du nm/min avec du disilane (Figure 3.19a). Similairement, afin de réduire la consommation de silicium monocristallin, la pression a été diminuée à 5 Torr afin de rallonger significativement le temps de gravure (Figure 3.19b). Ceci permet de plus d'obtenir un meilleur rapport de sélectivité entre le Si cristallin et l'amorphe (Figure 3.19c).

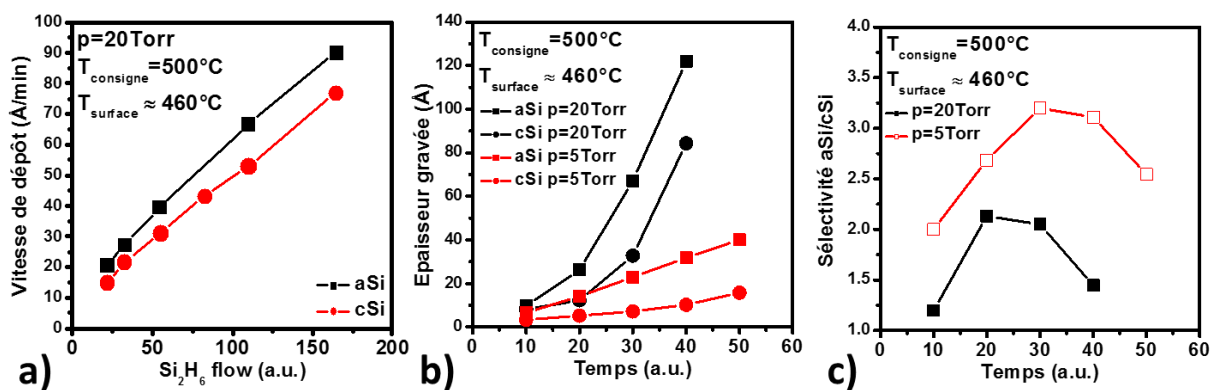


Figure 3.19 : a) Vitesse de dépôt et b) Vitesse de gravure de Si monocristallin ou amorphe à $T_{consigne}=500^{\circ}\text{C}$ ($T_{surface}=460^{\circ}\text{C}$); c) Rapport de sélectivité entre aSi et cSi à 20 ou 5Torr (Données AMAT).

Afin d'avoir un procédé isobare, un essai a été fait avec des cycles de dépôt et de gravure à 5 Torr à $T_{consigne}=500^{\circ}\text{C}$ (Figure 3.20).

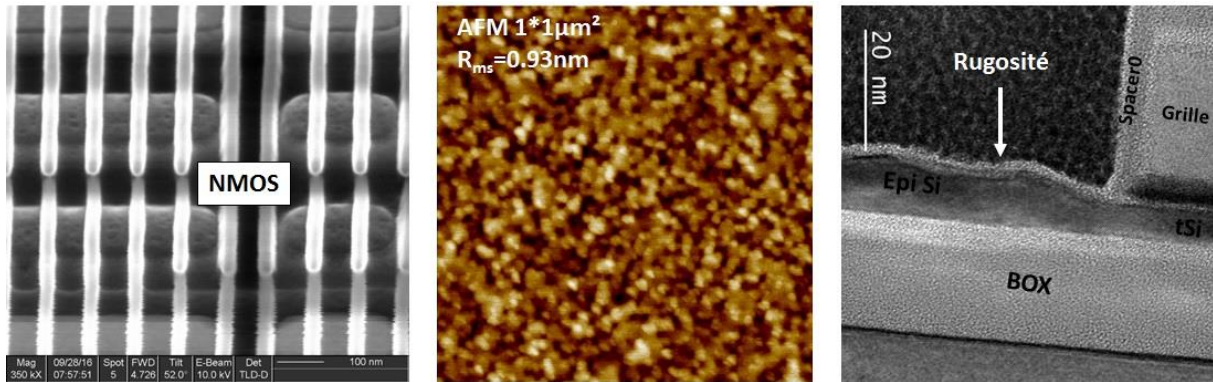


Figure 3.20 : Observations SEM (gauche), TEM (droite) et mesure AFM (milieu) après une épitaxie CDE (15 cycles - Dépôt @5Torr/Etch @5Torr) de S/D Si à $T_{consigne}=500^{\circ}\text{C}$ ($T_{surface}=460^{\circ}\text{C}$).

Outre une épaisseur finale très faible, on observe surtout une très forte rugosité de surface qui est attribuée à une pression de dépôt trop faible. Ainsi, un autre point de fonctionnement a été obtenu en ré-augmentant la température de consigne à 525°C ($T_{surface}=480^{\circ}\text{C}$), permettant une vitesse de dépôt plus rapide pour un meilleur rendement et en repassant à une pression de dépôt de 20 Torr, celle de la gravure restant à 5 Torr (Figure 3.21). La rugosité obtenue est cette fois plus faible que dans le cas précédent (bien que plus rugueuse que les résultats de la Figure 3.18). On note également l'apparition d'une facette qui peut être liée à une gravure plus importante près des grilles comme observé auparavant.

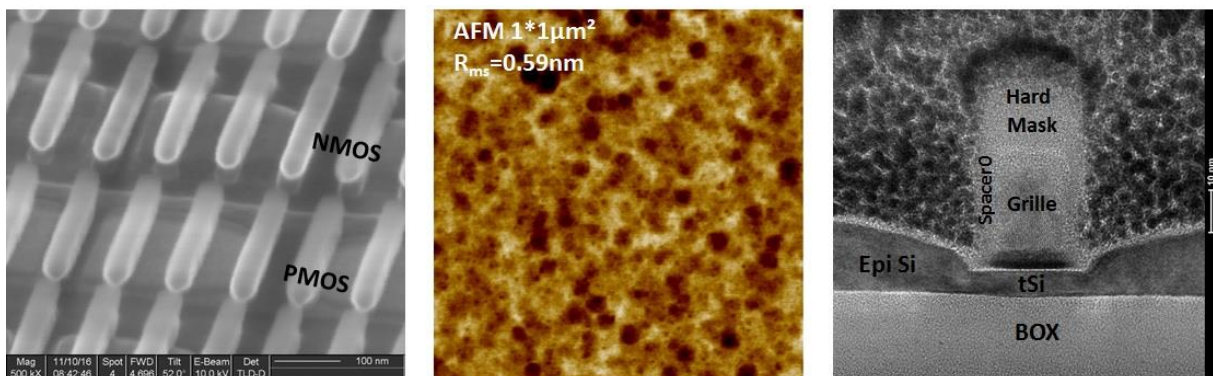


Figure 3.21 : Observations SEM (gauche), TEM (droite) et mesure AFM (milieu) après une épitaxie CDE (9 cycles - Dépôt @20Torr/Etch @5Torr) de S/D Si à $T_{consigne}=525^{\circ}\text{C}$ ($T_{surface}=480^{\circ}\text{C}$).

Deux points de fonctionnement ont donc pu être obtenus sur des plaques type 14FDSOI. Un premier (Dépôt @20Torr/Etch @20Torr) avec des temps par cycle inférieurs à 10s donnant un film d'excellente qualité, et un second (Dépôt @20Torr/Etch @5Torr) moins rapide (cycles ~5fois plus longs) et plus rugueux mais beaucoup plus robuste. Le développement d'une épitaxie Si non dopé à 500°C reste tout de même très délicat à cause de tous les paramètres rentrant en jeu. La pression de la chambre et les pressions partielles des différents gaz et le taux d'ouverture sont entre autres des paramètres importants. Le contrôle de la température est également primordial puisqu'elle joue sur les vitesses de dépôt et de gravure. Par exemple, le premier point de fonctionnement a été transféré sur un masque du Leti dont le taux d'ouverture est relativement similaire mais avec un oxyde enterré de 145nm au lieu de 20nm sur 14FDSOI. La Figure 3.22 montre des sources et drains totalement amorphisés avec également une croissance parasite au-dessus de la grille. Ceci est probablement dû à une température de surface « réelle » beaucoup plus faible qu'attendue et liée à une absorption de l'énergie thermique moins importante avec un BOX de 145nm. Par conséquent, le film Si initialement de 11nm a été totalement gravé dû à un dépôt cristallin moins épais voire amorphe. Tout comme pour

le cas de la Figure 3.17, une fois le BOX atteint, une croissance rapide de Si amorphe se produit qui n'est ensuite pas totalement gravé, menant également à la formation d'un « champignon » sur la grille.

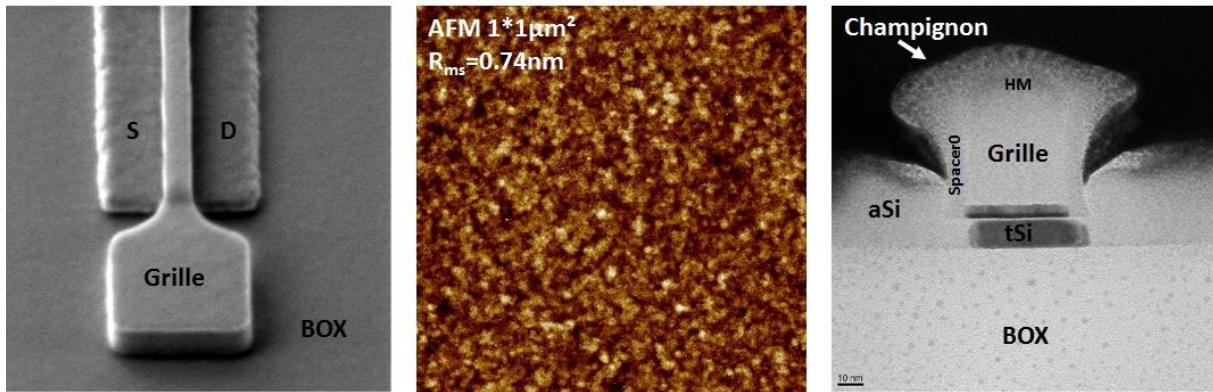


Figure 3.22 : Observations SEM (gauche), TEM (droite) et mesure AFM (milieu) après une épitaxie CDE (10 cycles - Dépôt @20Torr/Etch @20Torr) de S/D Si à $T_{consigne}=568^{\circ}\text{C}$ sur un masque différent avec un BOX de 145m. Le nombre de cycles a été adapté pour une épaisseur visée de 18nm.

Des études sont donc en cours afin de déterminer dans un premier temps la température réelle de la surface de la plaque avec un BOX 145nm, et dans un second temps de réadapter les conditions d'épitaxie. De plus, si la qualité cristalline des deux points de fonctionnement obtenus est en apparence (i.e. sur des observations TEM) bonne, les propriétés électriques de ces couches restent à être validées sur des transistors.

3.1.5 Epitaxie dopée *in-situ*

Les technologies CMOS avancées convergent désormais vers des épitaxies dopées *in-situ*, typiquement de SiC:P pour les nMOS et SiGe:B pour les pMOS. Outre l'intérêt de la contrainte tensile ou compressive apportée par ces sources et drains, l'épitaxie dopée *in-situ* permet également de s'affranchir du ou des étapes d'implantations ioniques de dopants. Comme soulevé dans le Chapitre 1, ces implantations introduisent en effet des défauts dits « end-of-range » qui ne sont pas dissouts dans une intégration basse température. De plus, une implantation trop énergétique risque d'amorphiser totalement le canal, en particulier dans le cas d'une épitaxie facettée (e.g. SiGe) [Cheng09]. A noter toutefois que l'utilisation d'épitaxies dopées nécessite quelques modifications dans l'intégration CMOS, puisqu'il faut alors protéger les transistors pMOS lors de l'épitaxie de type n et inversement. Ceci est notamment fait par un double dépôt d'espaces de grille (*dual spacer*) qui a été présenté dans le Chapitre 2. Dans cette section, nous détaillerons les résultats obtenus sur des épitaxies dopées *in-situ* de type N (Si:P) et de type P (SiGe:B) faites à 500°C.

3.2.1.1 Epitaxie dopée *in-situ* type N

Le phosphore et l'arsenic sont les deux éléments de choix quand il s'agit de doper en atomes donneurs. Cependant, l'utilisation d'arsine (AsH_3 , précurseur standard pour l'As) pose plusieurs difficultés. A relativement haute température ($T > 700^{\circ}\text{C}$), les atomes d'arsenic ont tendance à s'évaporer de la surface de silicium, tandis qu'à basse température ($T < 600^{\circ}\text{C}$), la probabilité d'adsorption de l' AsH_3 sur des liaisons pendantes est réduite [Liu05]. Enfin, dans la région de température intermédiaire ($600^{\circ}\text{C} < T < 700^{\circ}\text{C}$), un phénomène de ségrégation des atomes d'arsenic

tend à « empoisonner » la croissance en occupant les sites d'adsorption, diminuant alors la vitesse de croissance ainsi que la quantité de dopants qui peuvent être incorporés. L'utilisation de phosphine (PH_3), bien que plus stable, pose également le problème de ségrégation des atomes de phosphore comme nous l'avons vu précédemment pour une croissance de SiGe:P à 650°C ($\text{SiH}_2\text{Cl}_2 + \text{GeH}_4 + \text{PH}_3$) sur la Figure 3.11c ou également sur la croissance de Si:P à 650°C avec du silane ($\text{SiH}_4 + \text{PH}_3$) en rouge sur la Figure 3.23a. Au contraire, la croissance d'une couche Si:P à plus basse température (550°C) et avec du disilane n'est pas entravée par ce phénomène (en bleu).

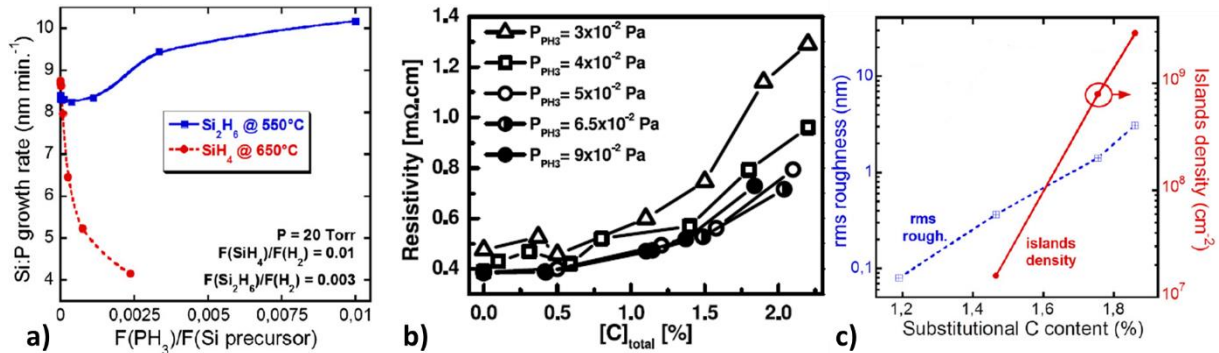


Figure 3.23 : a) Vitesse de croissance d'une couche Si:P en utilisant du SiH_4 ($T=650^\circ\text{C}$) ou du Si_2H_6 ($T=550^\circ\text{C}$); Résistivité d'une couche de SiC:P en fonction de la concentration de C incorporée et du débit de PH_3 ; c) Rugosité et densité de défauts en fonction de la concentration de C incorporée (pris de [Hartmann13b, Loubet12]).

Une autre différence entre As et P est la tendance de ce dernier à diffuser lors d'un recuit à haute température, le phosphore étant un atome plus petit. L'incorporation de carbone dans la matrice par l'ajout de mono-methylsilane (SiH_3CH_3), en plus d'apporter une contrainte tensile, permet de limiter cette diffusion, le contraire se produisant avec l'arsenic dont la diffusivité augmente en présence de carbone [Bauer05]. Ainsi, l'épitaxie de choix aujourd'hui pour les transistors nMOS est une épitaxie de sources et drains SiC:P grâce à une stratégie CDE [Wong08, Loubet12]. Cependant, la quantité de carbone incorporée en position substitutionnelle diminue avec la quantité de PH_3 injectée, une compétition s'installant entre les atomes de C et de P. Inversement, la résistivité d'une couche de SiC:P augmente avec la quantité de carbone (i.e. la diminution de la concentration en P), passant par exemple de $0.39 \text{ m}\Omega\cdot\text{cm}$ (0% de carbone) à $0.73 \text{ m}\Omega\cdot\text{cm}$ (2% C) pour la pression partielle de PH_3 la plus élevée dans la Figure 3.23b. Dernier point et non des moindres, la forte différence de paramètre de maille entre le carbone et le silicium complique la croissance d'une couche de SiC (ou de SiC:P) cristallin. La Figure 3.23c montre par exemple une augmentation de la rugosité de surface mesurée à l'AFM ainsi que l'apparition d'îlots avec l'incorporation de carbone, et ce dès 1.5% de C. Ces défauts sont notamment dus à la relaxation de la couche lorsque la contrainte élastique est trop importante.

Ainsi, dans le cadre de nos études à très basse température, une épitaxie de Si dopée phosphore a été choisie. La croissance d'une couche de SiC:P à 500°C avec du disilane étant probablement amorphe à cette température, cette option n'a pour le moment pas été envisagée. Un essai d'épitaxie sources et drains Si:P ($[\text{P}] = 2.5 \times 10^{20} \text{ at/cm}^3$) à $T_{\text{consigne}} = 568^\circ\text{C}$ ($T_{\text{surface}} = 500^\circ\text{C}$) par CDE (dépôt $\text{Si}_2\text{H}_6 + \text{PH}_3$ /gravure Cl_2) a donc également été réalisé dans le cadre des études avec Applied Materials (Figure 3.24).

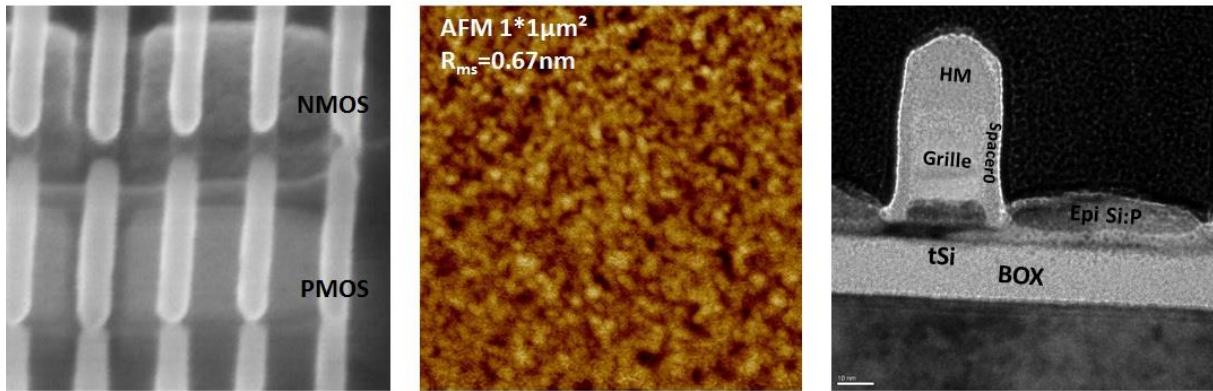


Figure 3.24 : Observations SEM (gauche), TEM (droite) et mesure AFM (milieu) après une épitaxie CDE (Dépôt @20Torr/Etch @20Torr) de S/D Si:P à $T_{consigne}=568^{\circ}\text{C}$ ($T_{surface}=500^{\circ}\text{C}$).

Si la sélectivité vis-à-vis des diélectriques ainsi que la cristallinité sont bonnes, les sources et drains apparaissent rugueuses et plus fines que l'épaisseur épitaxiée visée (7nm contre 18nm visée). Ceci est expliqué par le fait qu'à débit égal de Cl_2 , une couche de Si:P cristallin est gravée beaucoup plus rapidement qu'une couche de Si non dopée (Figure 3.25a). Une réduction du débit de Cl_2 est alors nécessaire afin de diminuer la consommation de cSi:P (Figure 3.25b). On remarque également que pour des couches Si:P, l'emploi d'un flux de Cl_2 réduit conduit à une sélectivité de gravure entre amorphe et cristallin assez faible comparé au Si intrinsèque, ce qui peut être problématique. Des développements sont néanmoins actuellement en cours afin d'améliorer la qualité des sources et drains Si:P.

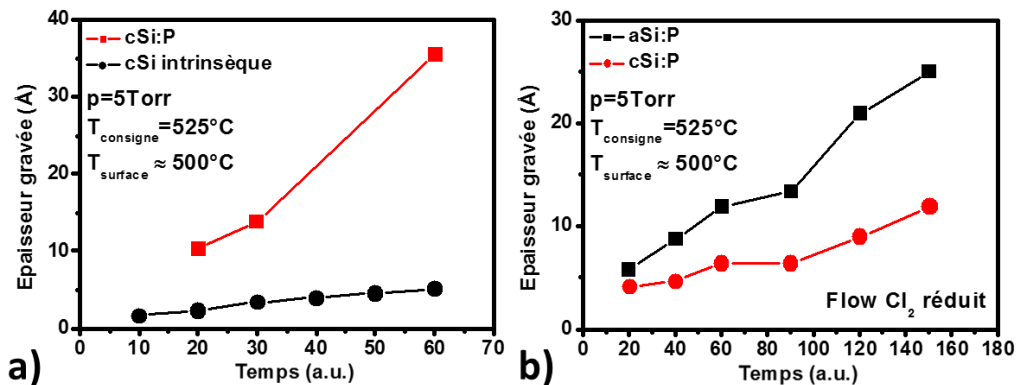


Figure 3.25 : a) Epaisseur gravée de Si intrinsèque ou Si:P monocristallin en fonction du temps; b) Epaisseur gravée de Si:P monocristallin ou amorphe en fonction du temps avec un débit de Cl_2 fortement réduit (Données AMAT).

3.2.1.2 Epitaxie dopée *in-situ* type P

L'épitaxie de choix pour les transistors pMOS est l'épitaxie de sources et drains SiGe:B. Comme nous l'avons vu auparavant, l'ajout de germane et de diborane permet en effet d'améliorer la vitesse de croissance, tandis que le germanium incorporé introduit une contrainte compressive sur le canal, améliorant la mobilité des trous. Contrairement à l'épitaxie de SiC ou de SiC:P, le paramètre de maille du Ge ($a_{\text{Ge}}=5.658\text{\AA}$) a une différence moins importante avec le Si ($a_{\text{Si}}=5.431\text{\AA}$) que le carbone ($a_{\text{C}}=3.567\text{\AA}$), ce qui permet d'obtenir des concentrations de Ge bien plus importantes sans relaxation plastique de la couche (>30%). L'épitaxie standard de sources et drains SiGe:B est faite au Leti à 650°C , 20Torr grâce à une combinaison de gaz fortement chlorés ($\text{SiH}_2\text{Cl}_2+\text{GeH}_4+\text{B}_2\text{H}_6+\text{HCl}$) afin de promouvoir la sélectivité [Hartmann11, Leroyer11].

La diminution de la température est alors confrontée aux mêmes problématiques que pour le Si intrinsèque, c'est-à-dire une vitesse de croissance dégradée avec l'utilisation de dichlorosilane comme gaz précurseur du silicium. Au contraire, l'utilisation de disilane à la place du SiH_2Cl_2 donne les vitesses de croissance voulues, mais au détriment de la sélectivité. La Figure 3.26a montre en effet un rapport de sélectivité entre SiGe poly et monocristallin inférieur à 2 même avec un débit important de HCl. L'épitaxie de couches SiGe:B à 500°C doit donc également s'orienter vers une stratégie de dépôt/gravure sélective [Hartmann14]. La croissance de SiGe:B à très basse température a toutefois un avantage puisqu'elle permet d'augmenter la quantité de dopants incorporés en position substitutionnelle, permettant de réduire la résistivité de la couche (Figure 3.26b). Le procédé développé à basse température mène également à une concentration légèrement supérieure de germanium ce qui devrait permettre d'induire une contrainte plus élevée (Figure 3.26c). Ceci reste toutefois à relativiser dans la mesure où le procédé à 650°C n'est pas forcément optimisé pour incorporer une forte concentration en Ge. Un procédé à plus basse température pourrait tout de même éventuellement permettre d'incorporer plus de germanium sans relaxation plastique de la couche.

Par ailleurs, tout comme le carbone et le phosphore dans une couche de SiC:P, il existe une compétition entre l'incorporation de Ge et de B. On notera également que pour des débits de B_2H_6 importants, la résistivité de la couche de SiGe:B est dégradée bien que la concentration atomique en bore augmente, ce qui peut être dû à une tendance du bore à former des clusters (cf. Chapitre 1) [Reznicek12].

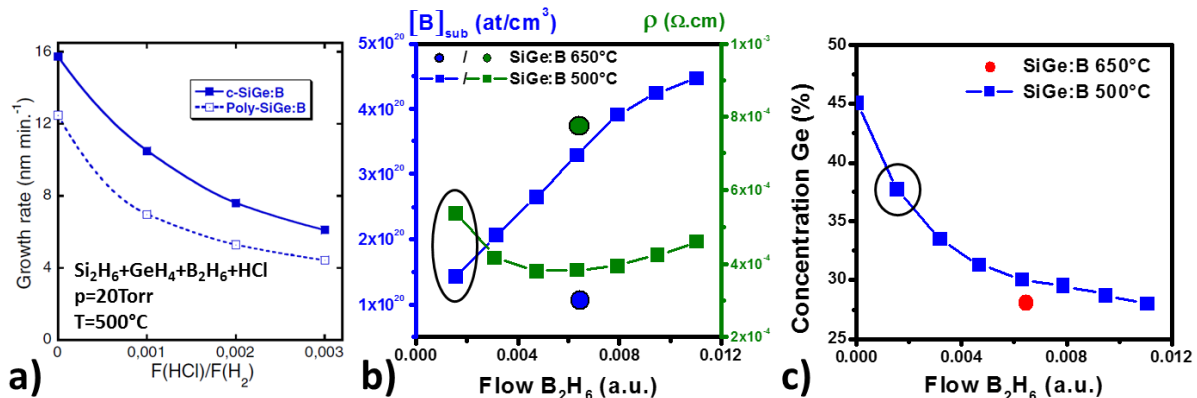


Figure 3.26 : a) Vitesse de croissance d'une couche SiGe:B monocristalline (sur Si) ou amorphe (sur SiO_2) en fonction du débit de HCl; b) Concentration de bore en position substitutionnelle $[B]_{\text{sub}}$, résistivité ρ et c) concentration de Ge incorporé dans une couche de SiGe:B épitaxiée à 650°C (co-flow) ou 500°C (Dep/Etch) à 20 Torr en fonction du débit de B_2H_6 (pris de [Hartmann14]). Les points entourés correspondent au procédé utilisé par la suite.

Un problème néanmoins avec l'épitaxie de SiGe:B par dépôt/gravure cyclée est la tendance de cette couche en contrainte à se rugosifier lors de sa gravure et qui peut être dû à une relaxation élastique près de la surface [Destefanis08]. Ainsi, la stratégie adoptée pour l'épitaxie de sources et drains SiGe:B est d'un « cycle » unique de dépôt non sélectif/gravure sélective. Plusieurs temps et débits de gaz pour les étapes de dépôt/gravure ont été étudiés dans [Hartmann14], le procédé final étant un dépôt durant 530s à 20Torr ($\text{Si}_2\text{H}_6 + \text{GeH}_4 + \text{B}_2\text{H}_6 + \text{HCl}$) avec un débit de B_2H_6 correspondant au point de fonctionnement entouré sur la Figure 3.26 (soit une concentration de 37% Ge et une résistivité de 0.54m $\Omega\cdot\text{cm}$), suivi d'une gravure HCl durant 210s à 740Torr (une pression plus élevée permettant des vitesses de gravure plus rapides). A noter que contrairement aux couches Si et Si:P étudiées précédemment, le Cl_2 n'est pas utilisé ici d'une part car il n'est pas disponible sur le bâti d'épitaxie 300mm utilisé au Leti, et d'autre part car la gravure par HCl reste satisfaisante à 500°C, celle-ci étant d'autant plus rapide que la concentration de Ge est importante.

Ces sources et drains $\text{Si}_{0.63}\text{Ge}_{0.37}:\text{B}$ à 500°C ont été intégrés dans le process flow standard en lieu et place de l'épitaxie Si à 750° (Figure 3.27).

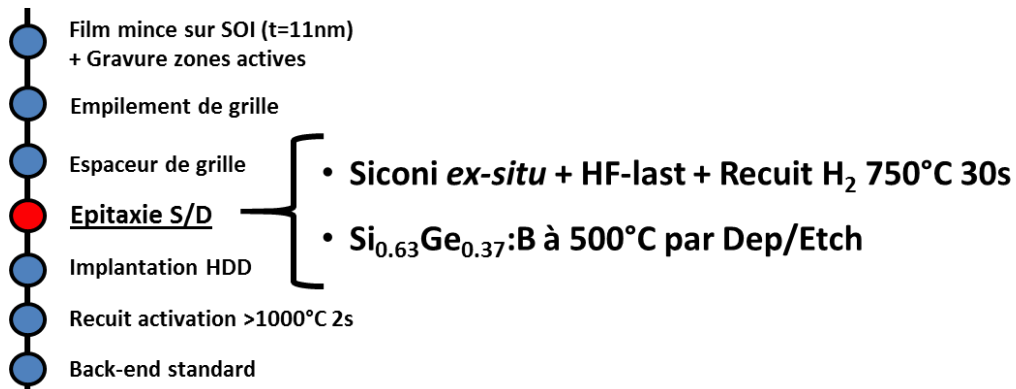


Figure 3.27 : Etapes de fabrication d'un transistor pMOS intégrant une épitaxie de sources et drains SiGe:B à 500°C, précédée d'une combinaison Siconi ex-situ+HF-last+recuit H₂ hybride 650°C 2min+750°C 30s.

A noter qu'une implantation HDD de type P a été réalisée après épitaxie suivie d'un recuit d'activation à haute température (1050°C 2s). De plus, la préparation de surface avant croissance épitaxiale a été faite par une combinaison de nettoyages sec *ex-situ* et humide (séquence optimale trouvée dans la Section 3.1.2) suivie d'un recuit H₂ hybride à relativement haute température (650°C 2min+750°C 30s). L'épitaxie $\text{Si}_{0.63}\text{Ge}_{0.37}:\text{B}$ par dépôt/gravure résultante fût effectivement sélective, bien qu'un peu rugueuse (Figure 3.28).

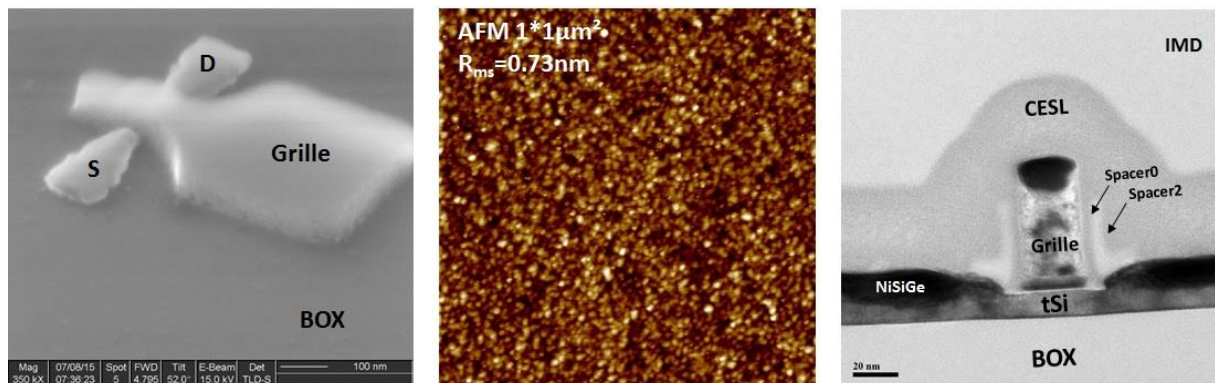


Figure 3.28 : Observations SEM (gauche), TEM (droite) et mesure AFM (milieu) après une épitaxie Dep/Etch (Dépôt 510s@20Torr/Etch 210s@740Torr) de S/D $\text{Si}_{0.63}\text{Ge}_{0.37}:\text{B}$ à 500°C. L'observation TEM a été effectuée à la fin de l'intégration (présence du siliciure et du CESL).

L'observation TEM après siliciuration montre également une facette des S/D moins inclinée qu'une facette <111> tel qu'observé précédemment sur la Figure 3.14c. De manière similaire à l'utilisation de Cl₂ pour le Si intrinsèque à 500°C, ceci pourrait être dû à une gravure plus rapide de la couche près des grilles.

Les résultats électriques montrent que des transistors fonctionnels ont été obtenus, avec néanmoins une dégradation de 25% du compromis Ion/Ioff par rapport à une épitaxie « co-flow » standard $\text{Si}_{0.73}\text{Ge}_{0.27}:\text{B}$ à 650°C (Figure 3.29a). Plusieurs facteurs peuvent expliquer ces résultats qui peuvent sembler surprenants à la vue des concentrations de Ge (37% contre 27%) et de B (1.4E20 contre 1E20at/cm³) qui sont plus élevées à 500°C. Tout d'abord, l'extraction de la mobilité à champ faible montre un pic pour des longueurs de grilles courtes plus faible que dans le cas du $\text{Si}_{0.73}\text{Ge}_{0.27}:\text{B}$ (Figure 3.29b). Ceci peut être dû à l'incorporation plus importante de bore à basse température, atome plus petit et qui peut donc avoir un effet compensateur sur la contrainte. Cet effet a pu être d'autant plus amplifié qu'une étape d'implantation ionique a été rajoutée dans les deux cas après l'épitaxie

dopée *in-situ*. En outre, la concentration trop importante de dopants a également potentiellement mené à la formation de clusters de bore électriquement inactifs, augmentant alors la résistivité des S/D $\text{Si}_{0.63}\text{Ge}_{0.37}\text{:B}$ (comme vu sur la Figure 3.26b). En effet, la résistance d'accès de la variante à très basse température est très largement dégradée par rapport à la référence à 650°C (Figure 3.29c). Une autre contribution à la dégradation du $R_{\text{accès}}$ peut être un problème d'*underlapping* des jonctions. On remarque en effet que le paramètre de DIBL se dégrade à des longueurs de grille plus courtes que pour les S/D à 650°C (Figure 3.29d), ce qui se traduit également par des effets de canaux courts moins importants sur la tension de seuil linéaire pour la variante à 500°C (Figure 3.29e). Une explication possible est la diminution de la diffusivité du bore lorsque la concentration en Ge augmente, phénomène mis en évidence dans [Rajendran00] par exemple. En conséquence, le bore a peut-être moins diffusé lors du recuit d'activation à haute température, résultant en une zone peu dopée sous les espaceurs de grille, donc fortement résistive. Par ailleurs, la présence de clusters de bore pourrait également expliquer la faible diffusion de ceux-ci.

De ces résultats, nous pouvons conclure que :

- Le dopage *in-situ* d'une épitaxie SiGe:B à basse température étant déjà élevée, l'étape d'implantation ionique HDD n'est sans-doute pas nécessaire. Cela permet en outre d'éviter une éventuelle amorphisation du canal lors de l'implantation.
- Le problème de diffusion, qui se retrouve également dans la thématique de la 3D Séquentielle où l'activation des dopants par SPER (600°C 2min) ne permet pas de diffuser les dopants sous les espaceurs de grille, peut être résolu en passant à une stratégie dite « extension-first » permettant de placer les dopants sous les espaceurs avant leur déposition (cf. Chapitre 1).

La morphologie d'une épitaxie dopée *in-situ* à 500°C sur film dopé ainsi que les performances électriques d'une intégration « *implant-free* » restent néanmoins à être validés.

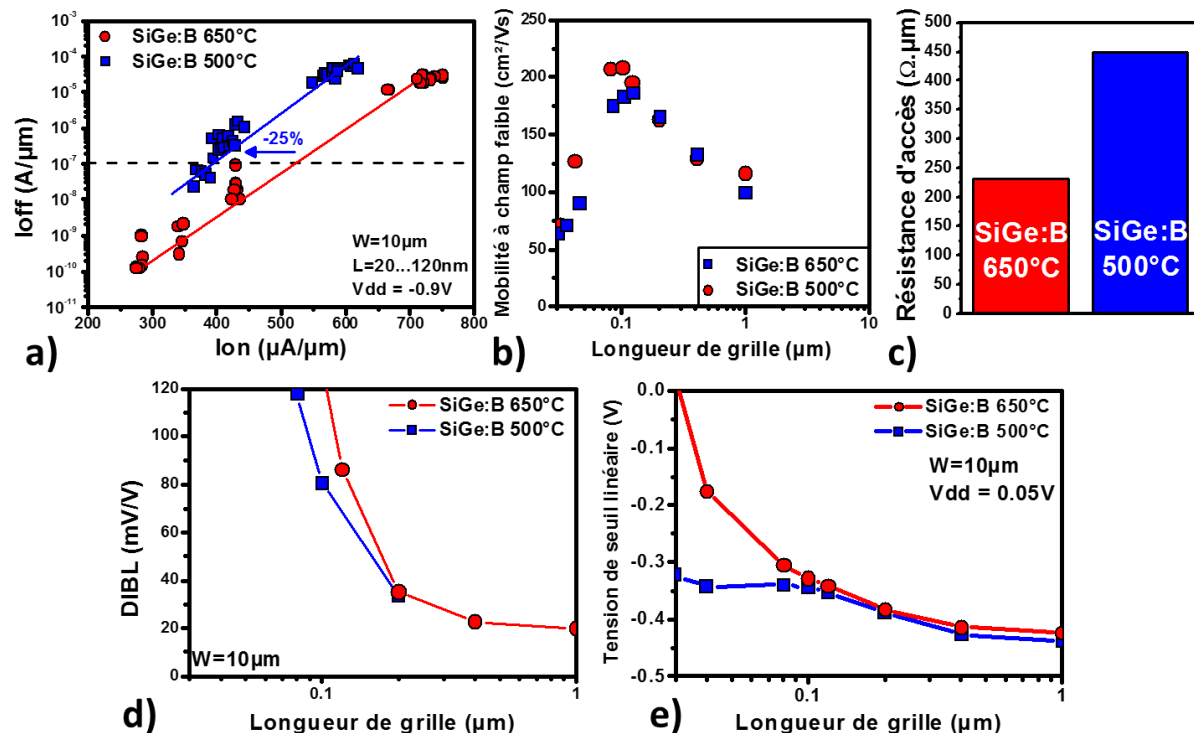


Figure 3.29 : Résultats électriques sur des transistors pMOS intégrant des S/D SiGe:B à 650°C (co-flow) ou à 500°C (Dep/Etch). a) Compromis $I_{\text{on}}/I_{\text{off}}$ mesuré sur des transistors larges ($W=10\mu\text{m}$); b) mobilité des trous à champ faible et c) résistance d'accès extraites par la méthode Y; d) variation du DIBL et e) de la tension de seuil à $V_{\text{dd}}=0.05\text{V}$ en fonction de la longueur de grille dessinée [Lu16b].

Conclusions du Chapitre 3

Le développement d'une épitaxie de surélévation des sources et drains à basse température est donc confronté à deux problématiques : la préparation de surface et la croissance épitaxiale.

Nous avons vu que **dans certains cas, la préparation de surface standard (nettoyage chimique « HF-last » suivi d'un recuit H₂ 650°C 2min+750°C 30s) peut s'avérer insuffisante**, ceci malgré un budget thermique relativement élevé. L'état de surface avant épitaxie dépend en effet fortement de toutes les étapes précédentes : gravure de la grille ou des espaceurs, efficacité des nettoyages chimiques et des retraits polymères, etc. L'intérêt d'une préparation de surface Siconi *ex-situ* (basée sur un plasma NH₃/NF₃) a donc été démontrée et une séquence optimale identifiée (i.e. un nettoyage Siconi « complet » *ex-situ* suivi d'un HF-last et d'un recuit H₂). L'utilisation de cette combinaison de préparation de surface sèche/humide a notamment permis de passer d'un recuit H₂ hybride 650°C 2min+750°C 30s à un recuit 650°C 2min seul sans impacter les propriétés électriques. Néanmoins, **une chambre Siconi *in-situ* est obligatoire afin de diminuer le budget thermique du recuit H₂ en-dessous de 500°C.**

Par ailleurs, la difficulté d'une croissance épitaxiale à basse température par stratégie « co-flow » est liée à la nécessité de conserver une vitesse de croissance correcte (de l'ordre du nm/min) tout en préservant la sélectivité et la qualité cristalline de l'épitaxie. **Une stratégie d'épitaxie cyclée (ou non) de dépôt suivi d'une gravure sélective a donc été présentée, permettant la croissance de sources et drains Si, Si:P et SiGe:B à 500°C** pour la première fois (Figure 3.30). L'épitaxie CDE permet notamment l'utilisation de précurseurs hydrogénés intrinsèquement non sélectifs mais produisant des vitesses de croissance plus élevées lors du dépôt, les phases non cristallines étant ensuite retirées sélectivement par du Cl₂ ou du HCl. Beaucoup de paramètres doivent cependant être pris en compte et être optimisés : la pression de la chambre et les pressions partielles des différents gaz, le taux d'ouverture, les temps des étapes de dépôt et de gravure, etc. L'absorption thermique dépendant des couches présentes est également un paramètre important à prendre en compte, en particulier dans le cas d'une intégration 3D, puisqu'elle modifie la température « réelle » de la surface de la plaque. Par ailleurs, si l'épitaxie dopée *in-situ* est très intéressante d'un point de vue intégration, **une optimisation de la quantité de dopants introduite est nécessaire afin d'éviter les problèmes de clusterisation**, en particulier lorsque suivie d'une activation de dopants à basse température.

		Haute température (référence)	Basse température	
Si intrinsèque		Epitaxie co-flow SiH ₂ Cl ₂ + HCl - 750°C (standard Leti)	Epitaxie cyclée dépôt/gravure Si ₂ H ₆ /Cl ₂ - 500°C	
Epitaxie pMOS	SiGe	Epitaxie co-flow SiH ₂ Cl ₂ + GeH ₄ + HCl - 650°C (standard Leti)	SiGe:B	Epitaxie dépôt/gravure Si ₂ H ₆ + GeH ₄ + B ₂ H ₆ + HCl/HCl - 500°C [Lu16b]
	SiGe:B	Epitaxie co-flow SiH ₂ Cl ₂ + GeH ₄ + HCl + B ₂ H ₆ - 650°C (standard Leti)		
Epitaxie nMOS	SiC:P	Epitaxie cyclée dépôt/gravure Si ₃ H ₈ + SiH ₃ CH ₃ + PH ₃ - 550°C/Cl₂ - 550°C [Bauer06] Si ₂ H ₆ + SiH ₃ CH ₃ + PH ₃ - 550°C/HCl + GeH₄ - 600°C [Hartmann13b]	Si:P	Epitaxie cyclée dépôt/gravure Si ₂ H ₆ + PH ₃ /Cl ₂ - 500°C (Développements en cours)

Figure 3.30 : Résumé des différentes épitaxies étudiées.

Chapitre 4 – Propriétés de l’empilement de grille

Introduction au Chapitre 4

L'empilement de grille est une des briques les plus importantes et complexes à optimiser dans un transistor MOSFET. En effet, c'est la polarisation sur l'électrode de grille qui contrôle l'établissement (voire la disparition dans le cas d'un transistor à appauvrissement) du canal de conduction entre la source et le drain permettant le passage d'un courant. Comme son nom l'indique, un transistor à effet de champ MOS est composé d'un métal sur lequel la polarisation de grille est appliquée, d'un semi-conducteur (dans notre cas, du silicium) à travers lequel le courant pourra passer ou non et d'un oxyde de grille isolant ces deux derniers. Tout du long de la miniaturisation des transistors, divers empilements de grilles ont été utilisés afin d'améliorer la performance des composants.

La grille SiO₂/poly-Si

Le couple oxyde de silicium/silicium polycristallin dopé (SiO₂/poly-Si) a pendant longtemps été utilisé comme diélectrique et électrode de grille pour les transistors MOS. Les avantages du SiO₂ sont nombreux : il peut être formé facilement et de manière contrôlée par un recuit à haute température du silicium sous atmosphère oxydante, il forme une interface quasi optimale avec le silicium, comporte peu de défauts de volume et possède un gap d'énergie large, établissant une excellente barrière de potentiel (ΔE_C et $q\Phi_B$) pour les deux types de porteurs (Figure 4.1a). De même, le poly-Si peut être déposé facilement, est stable thermiquement et ne réagit pas avec l'oxyde de silicium. Le dopage du poly-Si permet de plus d'augmenter (implantation type N, e.g. phosphore, arsenic), ou de diminuer (implantation type P, e.g. bore) son travail de sortie Φ_M , permettant de contrôler la **tension de seuil** (V_T) du transistor (Figure 4.1b). Il permet par ailleurs d'améliorer conséquemment la conductivité du poly-Si qui devient alors un semi-conducteur dégénéré.

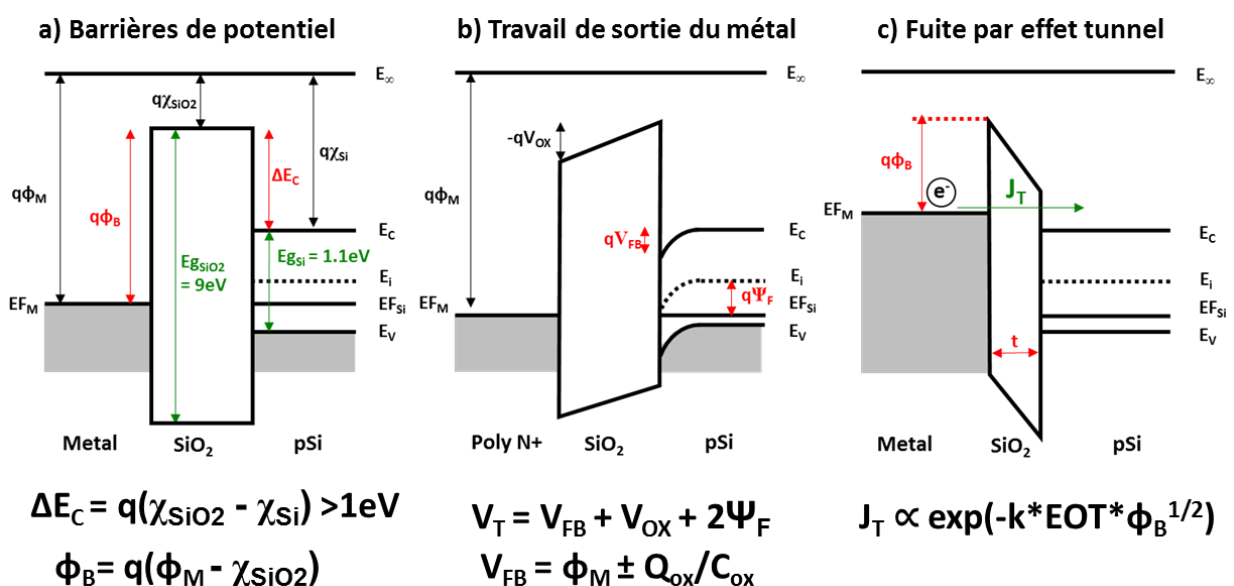


Figure 4.1 : Diagrammes de bandes d'énergie d'un MOS de type P représentant a) les barrières de potentiel du côté du métal et du semiconducteur, b) l'effet du travail de sortie du métal sur les courbures de bande du silicium lors de l'alignement des niveaux de Fermi, c) le phénomène d'effet tunnel à travers l'oxyde.

Diélectriques à haute permittivité

Le passage à des nœuds technologiques plus avancés et la réduction de la longueur de grille associée impose alors de diminuer l'épaisseur de l'oxyde afin de conserver une capacité surfacique constante et un bon contrôle du canal. Cependant, à partir du nœud 90nm environ, la réduction de l'épaisseur de SiO₂ atteint des limites technologiques d'une part, avec la nécessité de former un oxyde de bonne qualité de quelques monocouches seulement, et physiques d'autre part, avec l'apparition d'effets quantiques dégradant les performances des transistors. En particulier, l'effet tunnel (J_T), consistant au passage des porteurs à travers une couche très mince même si ceux-ci ont une énergie inférieure à la barrière de potentiel, augmente drastiquement les **courants de fuite** à travers l'oxyde (Figure 4.1c). Afin de conserver une épaisseur d'isolant plus conséquente et une capacité de grille élevée, la permittivité de l'oxyde doit être augmentée. On peut alors introduire le paramètre d'**épaisseur équivalente d'oxyde** (EOT – *Equivalent Oxide Thickness*) qui représente l'épaisseur de SiO₂ nécessaire pour obtenir la même capacité d'oxyde (C_{ox}) qu'avec un diélectrique de permittivité k et d'épaisseur t_{high-k} plus importantes et qui s'exprime par :

$$EOT = t_{high-k} \frac{k_{SiO_2} = 3.9}{k}$$

Une première solution pour augmenter la permittivité du SiO₂ est d'incorporer de l'azote par nitruration [Wilk01]. La permittivité des couches SiO_xN_y ainsi formées reste toutefois assez faible et nécessairement inférieure à celle d'un nitrure de silicium (k=7), limitant leur utilisation au nœud 65nm. Pour aller plus loin, de nouveaux matériaux dits high-k doivent être introduits. Le choix du high-k est alors soumis à plusieurs critères : avoir une permittivité assez élevée pour pouvoir être utilisé sur plusieurs nœuds successifs et être économiquement viable, avoir des propriétés électriques proches du SiO₂ et être stable thermiquement [Gusev01]. Un certain nombre de diélectriques high-k ont été considérés, en particulier ceux dérivés du groupe III du tableau périodique (e.g. Al₂O₃), du groupe IV (e.g. TiO₂, HfO₂, ZrO₂), du groupe V (e.g. Ta₂O₅) ou à base de terres rares (e.g. La₂O₃). La nécessité d'un diélectrique avec un bandgap large (généralement supérieur à 5eV) afin de limiter les fuites de courant limite déjà le choix du matériau. En effet, l'énergie de la bande interdite des oxydes est généralement inversement proportionnelle à leur constante diélectrique, limitant donc la valeur de permittivité utilisable (Figure 4.2a). Une deuxième problématique est la stabilité thermodynamique du diélectrique avec le silicium. Certains oxydes comme le Ta₂O₅ ont en effet tendance à réagir avec le silicium lors du dépôt ou sous l'effet d'un recuit, formant une couche de SiO₂ interfacial de faible permittivité qui dégrade l'EOT et limite donc l'intérêt du high-k [Alers98].

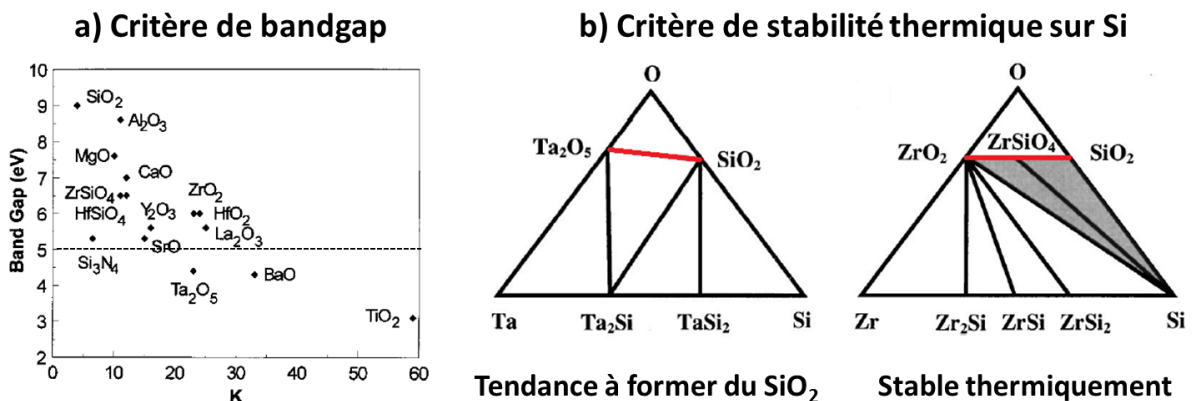


Figure 4.2 : a) Compromis entre l'énergie de bandgap et la permittivité de différents diélectriques [Robertson06] et b) Diagrammes de phase de systèmes ternaires Ta-Si-O et Zr-Si-O [Wilk01].

Au final, la plupart des industriels, et en particulier Intel [Mistry07], ont convergé vers des oxydes à base d'hafnium essentiellement du fait de leur permittivité élevée (k~25) et de leur relativement bonne stabilité avec le silicium.

Silicate d'hafnium et nitruration

D'autres problématiques demeurent néanmoins avec l'utilisation du high-k. Tout d'abord, le HfO_2 , déposé sous forme amorphe, cristallise à une température assez faible (400°C), formant alors des grains polycristallins lors d'un recuit. Les joints de grains formés peuvent alors agir comme des chemins privilégiés pour le passage des porteurs ce qui augmente les courants de fuite à travers le diélectrique, mais aussi comme des chemins de diffusion pour les dopants provenant du poly-Si, menant à des variations locales de la permittivité du matériau [Watanabe05]. Afin d'augmenter la température de cristallisation, des silicates d'hafnium (HfSiO_x) voire une variante nitrurée (HfSiON_x) ont été utilisés dans un premier temps malgré une permittivité plus faible ($k \sim 15$) que le HfO_2 . La nécessité d'avoir une permittivité toujours plus élevée ($n_{\text{œud}} < 28\text{nm}$) oblige néanmoins de revenir au HfO_2 afin d'obtenir des EOT plus agressifs. Plusieurs autres pistes existent afin d'augmenter la température de cristallisation du HfO_2 . Une première est de diminuer l'épaisseur du high-k sous une valeur critique au-dessous de laquelle le HfO_2 ne cristallise pas. Cette épaisseur critique a été mise en évidence dans [Besson08] par des études de consommation du HfO_2 dans un mélange HF/HCl. Au-dessus d'une épaisseur comprise entre 15Å et 40Å , la vitesse de gravure de la couche HfO_2 déposée par ALD et recuite à 600°C 15min devient quasi-nulle et est corrélée à la formation de phases cristallines qui se consomment moins rapidement que la phase amorphe. Par ailleurs, l'incorporation d'azote [Chang08] ou de zirconium [Kuo06] peut être une alternative afin de favoriser la forme amorphe du HfO_2 .

Oxyde piédestal et réduction de l'EOT

Une autre problématique est que l'interface formée par le high-k avec le silicium est moins parfaite que celle entre un SiO_2 et le silicium et comporte de nombreuses liaisons pendantes qui peuvent agir comme des défauts actifs électriquement. Cette **densité d'états d'interface** (D_{it}) élevée peut alors décaler la tension de seuil du transistor ou bien dégrader la **mobilité des porteurs** dans le canal via des phénomènes de *scattering* [Fang68]. Afin d'obtenir une meilleure interface, une fine couche de SiO_2 interfacial (aussi appelé oxyde piédestal) est généralement formée avant le dépôt du high-k soit de manière humide à basse température (e.g. UV- O_3 , SC1), soit thermiquement à plus haute température. De plus, un recuit aux alentours de 400°C sous atmosphère N_2/H_2 (aussi appelé FGA – *forming gas anneal*) est souvent effectué afin de passiver les liaisons pendantes par l'hydrogène [Yeow75]. Les liaisons Si-H formées étant métastables, ce recuit est généralement réalisé vers la fin de l'intégration au niveau de la formation des contacts afin d'éviter leur dépassivation (i.e. cassure de la liaison) par des budgets thermiques élevés. Cette couche de SiO_2 interfacial contribue néanmoins fortement à l'épaisseur équivalente de l'empilement $\text{SiO}_2/\text{HfO}_2$ due à sa faible permittivité. De plus, il a été calculé dans [Tang08] qu'au moins trois monocouches sont nécessaires afin d'obtenir toutes les propriétés isolantes du SiO_2 , correspondant à une EOT de $6-7\text{Å}$. C'est ce qui justifie également qu'il est moins nécessaire de travailler sur des diélectriques de très haute permittivité que d'optimiser l'interface du silicium avec l'oxyde piédestal tout en diminuant l'épaisseur de ce dernier. La technique de *scavenging* est notamment parfois utilisée afin de réduire davantage l'EOT sous $6-7\text{Å}$. Lorsqu'un oxyde est recuit en présence d'un élément dit « *oxygen scavenger* », le SiO_2 peut se décomposer, les atomes d'oxygène diffusant vers le métal à travers le high-k pour l'oxyder (si la réaction est énergétiquement favorable) tandis que le silicium est réincorporé dans le bulk [Kim04]. Plusieurs éléments *scavenger* existent mais le plus commun est le Ti qui nécessite une énergie thermique relativement faible ($T=300^\circ\text{C}$) pour dissocier l'oxyde et diffuser l'oxygène. Un *scavenging* complet du SiO_2 est alors possible mais aux dépens de la qualité d'interface entre l'oxyde de grille et le silicium [Ando12].

Grilles métalliques

Une dernière problématique avec l'introduction des high-k est liée à l'interface formée avec la grille poly-Si. Outre le problème de stabilité thermique entre les deux couches (de la même manière qu'avec le canal Si), le dépôt du poly-Si sur le HfO_2 fait apparaître un phénomène dit de *Fermi Level Pinning* au

niveau de l’interface. Cet effet, attribué à la formation de liaisons Si-Hf et/ou à la présence de défauts particuliers (lacunes d’oxygène) près de l’interface, mène à un écrantage du travail de sortie du poly-Si [Hobbs04]. Dans la pratique, cela signifie que le dopage du poly-Si ne permet plus de contrôler correctement les courbures de bandes du silicium et mène à des tensions de seuil plus élevées quel que soit le type de transistor (n ou p). De plus, la miniaturisation des transistors implique de diminuer la résistance de contact du poly-Si, ce qui est obtenu en le dopant plus fortement. En contrepartie, des effets de diffusion de dopants (en particulier du bore) à travers l’oxyde de grille peuvent apparaître, modifiant le niveau de dopage du canal et donc la tension de seuil. Ce problème est néanmoins en parti résolu par l’incorporation d’azote dans les diélectriques de grille qui permet de former une barrière de diffusion. Une dernière problématique est la formation d’une zone de déplétion image du canal dans le poly-Si (aussi appelé *poly-depletion*) qui introduit alors une capacité en série et qui dégrade l’épaisseur équivalente (Figure 4.3a). A noter qu’un effet similaire dit de « *darkspace* » apparaît à l’interface canal Si/SiO₂. Lors du régime d’inversion, les bandes d’énergie du silicium se courbent pour former le canal d’inversion. Lorsque la courbure de bande est élevée, une discrétisation des niveaux d’énergie apparaît menant à une probabilité de présence des porteurs légèrement éloignée de l’interface entre l’oxyde de grille et le silicium (Figure 4.3b). De fait, cet effet quantique augmente également l’épaisseur équivalente et n’est plus négligeable pour des EOT sous 1nm [Skotnicki03].

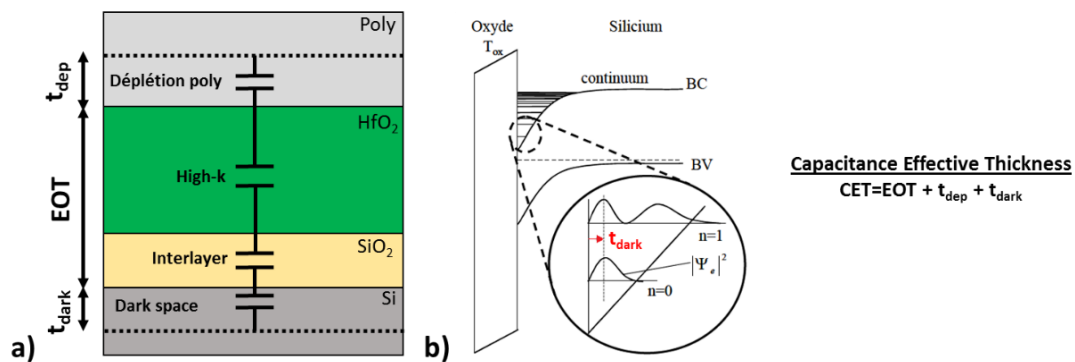


Figure 4.3 : a) Schéma équivalent représentant les différentes capacités en série de la grille et b) Phénomène quantique à l’interface entre le SiO₂ et le Si menant à l’apparition d’un darkspace [Skotnicki03].

Ces problématiques justifient alors l’introduction de grilles métalliques. La forte concentration de porteurs libres limite en effet la formation de la zone de déplétion dans la grille et donc la dégradation de l’épaisseur équivalente. Le choix du métal de grille, outre le critère de dépôt qui doit se faire sans endommager le diélectrique, dépend essentiellement de son travail de sortie qui détermine la tension de seuil du transistor. On peut alors soit utiliser un métal commun aux n et pMOS dont le travail de sortie se situe au milieu du gap du silicium (e.g. TiN) ou bien utiliser deux métaux avec des travaux de sortie favorables à chaque type de transistor (e.g. Ti pour nMOS ou Mo pour pMOS [Yeo01]). Un métal avec un travail de sortie faible (respectivement élevé) sera dit de type N (P) car il décale l’énergie de Fermi du silicium vers la bande de conduction (valence), diminuant la tension de seuil des nMOS (pMOS). Une autre possibilité pour faire varier la tension de seuil est la formation de dipôles (IDE – *Interface Dipole Engineering*). Sur le principe, lorsque des atomes d’électronégativités différentes sont liés, l’atome plus électronégatif tend à attirer un électron de l’atome moins électronégatif, créant localement un dipôle. A l’échelle du volume d’une couche isotrope, ces dipôles tendent à s’annuler, mais à l’interface entre deux matériaux différents, les dipôles se comportent comme une capacité introduisant un potentiel supplémentaire, modifiant la tension de bande plate (V_{FB}) de la capacité MOS [Huang12]. Dans la pratique, plusieurs méthodes existent pour former des dipôles d’interface, par exemple en insérant une couche de La₂O₃ entre le HfO₂ et le métal de grille pour les nMOS [Guha07] ou en dopant le high-k avec de l’aluminium [Sharia08] pour les pMOS. Dans tous les cas, si l’intégration

dite *dual metal-gate* semble plus intéressante, elle est par contre plus complexe et donc plus coûteuse avec l’ajout d’étapes technologiques supplémentaires pour séparer les deux types de transistors.

Intégration « gate-last »

Ceci étant dit, le travail de sortie d’un métal dépend non seulement de la technique de dépôt (e.g. CVD, PVD), de sa composition et de son épaisseur, mais également des procédés subséquents, et en particulier les budgets thermiques élevés (e.g. activation de dopants). Par exemple, un décalage négatif du travail de sortie du TiN est observé lorsque celui-ci est recuit, décalage dont l’origine est discutée : effet de *Fermi Level Pinning* [Yu04] ou modification des dipôles à l’interface métal/high-k [Charbonnier10] par exemple. Dans tous les cas, ce décalage augmente donc la tension de seuil des transistors pMOS. D’autre part, la réduction des épaisseurs d’oxyde de grille fait également apparaître un effet de « roll-off », correspondant à un décalage de la tension de seuil (ou de bandes plates sur les capacités) vers des valeurs négatives lorsque l’EOT diminue, ce qui augmente également la tension de seuil des pMOS. A noter que cet effet est corrélé à la réduction de l’épaisseur du SiO₂ et non du high-k [Charbonnier10]. Quoiqu’il en soit, la nécessité d’un meilleur contrôle de la tension de seuil, en particulier pour les transistors pMOS, justifie l’introduction de l’intégration « *gate-last* » (aussi appelée *replacement gate*), par opposition à l’intégration « *gate-first* » qui a été présentée dans le Chapitre 1. L’idée principale est alors d’éviter d’appliquer un budget thermique sur l’empilement de grille et la variation du travail de sortie de la grille [Mistry07]. Pour cela, une fausse grille (*dummy-gate*) en poly-Si est d’abord déposée, suivi du dépôt des espaceurs de grille, de l’épithaxie des sources et drains et de l’implantation et de l’activation des dopants à haute température. La fausse grille est ensuite retirée, puis remplacée par l’empilement de grille final qui ne voit donc aucun budget thermique élevé (Figure 4.4). Le flow *gate-last*, s’il permet de résoudre le problème de la tension de seuil des pMOS, reste néanmoins très complexe, notamment dû à l’utilisation de multiples étapes de CMP (*Chemical-Mechanical Planarization*), par exemple avant la gravure de la fausse grille ou encore après remplissage de l’empilement de grille final pour séparer les grilles, qui sont difficiles à contrôler [Shao13].

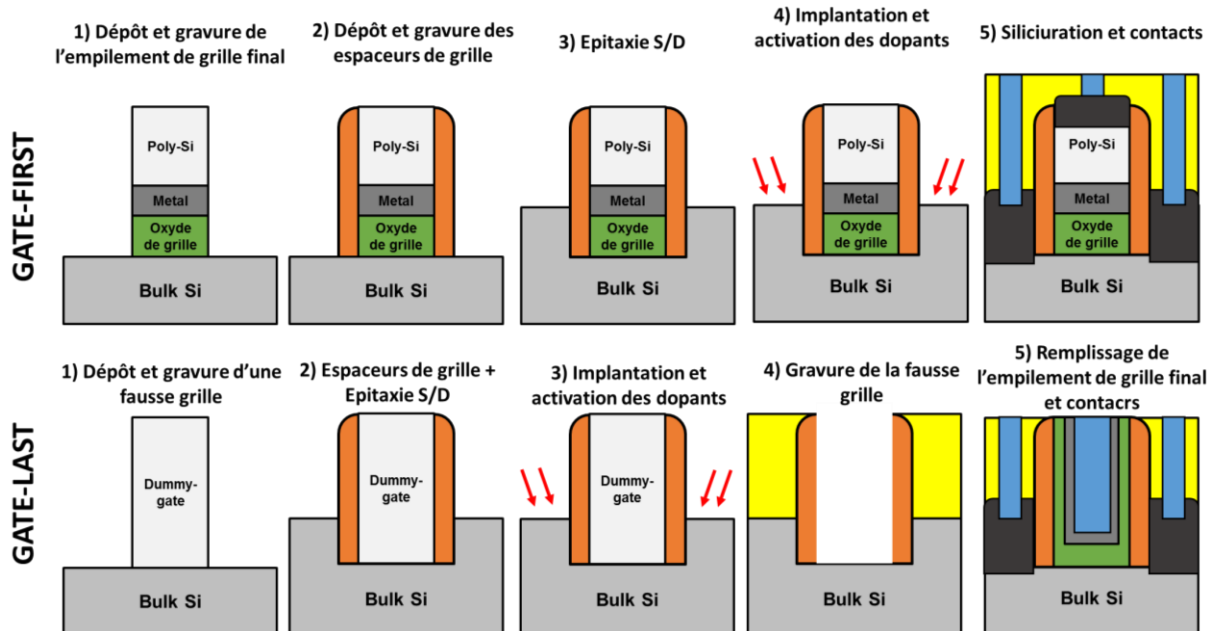


Figure 4.4 : Process flows simplifiés pour des intégrations *gate-first* et *gate-last*.

Fiabilité des transistors

Une dernière propriété essentielle à discuter est la fiabilité des transistors. De manière générale, la fiabilité correspond à la variation des caractéristiques du transistor (e.g. tension de seuil, courant de drain) au cours du temps et de son utilisation. Dans ce manuscrit, nous traiterons essentiellement du

phénomène de BTI (*Bias-Temperature Instability*). Lorsqu’une contrainte électrique (i.e. une tension élevée pendant une durée relativement longue) est appliquée sur la grille, une apparition de charges au niveau des diélectriques de grille peut faire varier la tension de seuil (et par extension, le courant de drain linéaire). Il faut alors dissocier deux paramètres correspondant au signe de la contrainte électrique : Positive (PBTI) et Negative (NBTI) Bias-Temperature Instability. Les porteurs mis en jeu selon la contrainte électrique (électrons pour PBTI, trous pour NBTI) étant différents, les phénomènes physiques varient également selon le paramètre étudié.

- NBTI :

Lorsqu’une contrainte électrique négative est appliquée sur la grille sur des transistors pMOS, une augmentation de la tension de seuil est observée qui est fonction entre autres de la durée de la contrainte, de la tension appliquée et de la température [Deal67]. Cette augmentation a d’abord été expérimentalement corrélée à la formation d’états d’interface (ΔD_{it}) entre le SiO₂ et le Si. Bien que ces défauts d’interface aient un caractère amphotère (i.e. peuvent agir soit comme un donneur, soit comme un accepteur d’électron), les mesures expérimentales montrent que les états d’interfaces formés se trouvent principalement dans la bande interdite du SiO₂ entre le midgap et la bande de valence du silicium, correspondant donc à un état donneur des défauts [Goetzberger73]. Ainsi, lorsqu’une tension négative est appliquée, une permutation entre l’électron du défaut donneur et un trou de la couche d’inversion peut se produire, créant une variation de charge nette positive dans l’oxyde, un décalage de la tension de bande plate vers des valeurs négatives et donc une augmentation de la tension de seuil des pMOS. Le développement de méthodes de mesures plus rapides (dites dynamiques) de la dégradation NBTI fait néanmoins apparaître une composante supplémentaire qui n’est pas observée lors de mesures plus longues (ou statiques) et qui ne provient pas des états d’interface. Cette composante est attribuée à un phénomène rapide de piégeage de trous soit sur des défauts préexistants dans l’oxyde de grille, soit sur des défauts créés par la contrainte électrique et qui fait donc également apparaître une charge positive dans l’oxyde [Ang06]. Lorsque la contrainte négative est retirée ou lorsqu’une tension positive est appliquée sur la grille (par exemple lors d’un balayage capacité-tension), le phénomène inverse se produit (dépiégeage des trous) menant à une relaxation du ΔV_T qui diminue (Figure 4.5a).

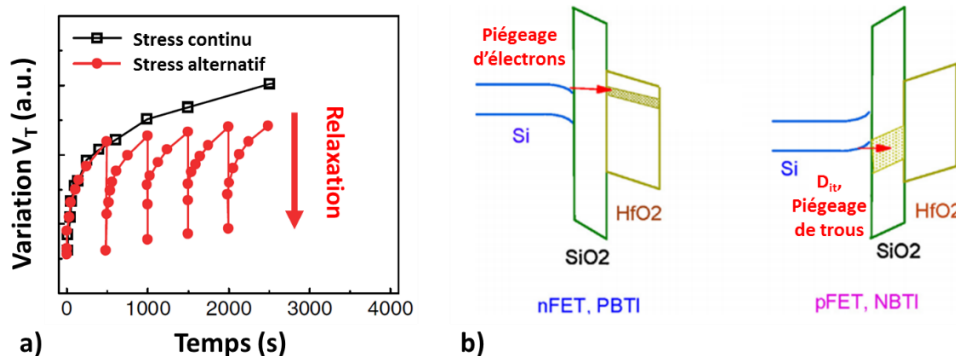


Figure 4.5 : a) Mesure de la dégradation de la tension de seuil suivant un stress continu ou un stress alternatif [Huard03] et b) Schéma de bandes représentant les défauts créés et leur niveau d’énergie selon la polarité de la contrainte électrique [Robertson15].

Cette composante « recouvrable » n’est donc pas observable lors de mesures plus longues. A noter aussi qu’un faible effet de relaxation des D_{it} a également été observé dans [TsujiKawa03]. Il est donc aujourd’hui généralement admis que la dégradation NBTI est liée à la formation d’états d’interface et au piégeage rapide de trous sur des défauts préexistants ou non. Cependant, les contributions exactes de ces composantes et les phénomènes physiques mis en jeu sont largement discutés, menant à divers modèles afin d’expliquer les mesures expérimentales voire de prévoir la dégradation NBTI [Mahapatra13].

- PBTI :

L’introduction des high-k fait également apparaître une instabilité de la grille lorsqu’une tension positive est appliquée, menant à une augmentation de la tension de seuil sur les transistors nMOS. Cette dégradation est généralement attribuée au piégeage d’électrons sur des défauts contenus dans le HfO₂ (similairement à la composante recouvrable du NBTI). L’origine et la contribution de ces défauts sont néanmoins sujets à discussion : oxygène interstitiel, lacunes d’oxygène, etc. [Ribes05]. Toutefois, nous verrons que le PBTI est généralement peu problématique comparé à la fiabilité NBTI.

De manière générale, du fait du type de porteur mis en jeu, il est admis et vérifié expérimentalement que la fiabilité NBTI est principalement liée à la qualité de la couche SiO₂ interfacial (défauts proches de la bande de valence du silicium et dans le volume du SiO₂) tandis que la fiabilité PBTI est liée à la qualité du high-k (défauts proches de la bande de conduction) (Figure 4.5b).

Pour résumer cette introduction, les principaux enjeux actuels de l’empilement de grille sont donc : la réduction de l’EOT, la diminution des courants de fuite, l’obtention d’une excellente interface entre l’oxyde de grille et le canal, le contrôle de la tension de seuil (en particulier pour les transistors pMOS) et l’amélioration de la fiabilité. Dans ce chapitre, nous présenterons avant tout les procédés expérimentaux mis en place pour étudier les propriétés de la grille (Section 4.1). Nous évaluerons ensuite dans un premier temps l’influence de la diminution des budgets thermiques subséquents (espaceurs de grille, surélévation des S/D, activation de dopants) sur les propriétés de la grille dans la Section 4.2. Ensuite, nous détaillerons dans la Section 4.3 différents procédés leviers développés à bas budget thermique et leurs effets sur l’empilement de grille : nitruration du high-k, densification de l’oxyde chimique et recuit de passivation deutérium. En outre, l’utilisation de recuits alternatifs introduits dans le Chapitre 1 (micro-ondes, laser) sera étudiée dans la Section 4.4. Enfin, nous discuterons dans la Section 4.5 des pistes pour améliorer la fiabilité NBTI dans le cadre d’une intégration 3D Séquentielle.

4.1 Procédés expérimentaux

La majorité de cette étude a été réalisée sur des capacités MOS (MOSCAPs), plus simples et donc plus rapide à fabriquer que des transistors MOSFETs. Le flow de fabrication simplifié d’un MOSCAP est donné dans la Figure 4.6.

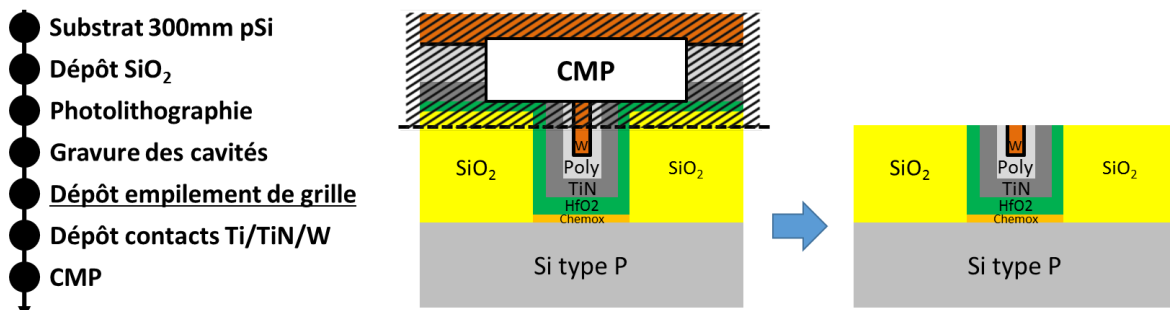


Figure 4.6 : Process flow simplifié et schéma d’une structure MOSCAP.

Des substrats silicium de type P ont été utilisés, représentant le fonctionnement d’un canal pMOS et permettant d’étudier la dégradation NBTI qui est plus problématique. Des capacités sur substrat de type N ont également été fabriqués mais n’ont pas pu être analysés rigoureusement dû à des

problématiques de fuites de grille trop élevées. Une solution serait d’utiliser des capacités plus petites mais nécessite alors des modifications dans l’intégration pour les mesurer ce qui n’a pas pu être fait dans cette thèse. Des cavités d’oxyde sont ensuite formées dans lesquelles l’empilement de grille et les contacts W précédés d’une barrière Ti/TiN sont déposés, suivi d’une CMP avec arrêt sur oxyde afin d’isoler les capacités MOS. L’empilement de grille, dans sa forme la plus simple, sera composé d’un oxyde piédestal formé chimiquement (aussi appelé chemox – *chemical oxide*), d’un dépôt HfO_2 2nm par ALD à 300°C, d’une grille TiN 6.5 nm déposée par PVD et d’un poly-Si 50nm déposé par LPCVD (*Low Pressure CVD*) à 525°C. La couche de poly-Si sert notamment de contact sur les transistors puisque celle-ci est siliciurée (cf. Chapitre 1). Une implantation d’arsenic est donc également effectuée après le dépôt poly-Si afin de diminuer sa résistivité.

Les mesures capacité-tension (C-V), intensité-tension (I-V) et conductance-tension (G-V) permettent alors de remonter aux paramètres d’EOT, de tension de bande plate (V_{FB}), de courant de fuite de grille (J_g) et de densité d’états d’interface (Figure 4.7).

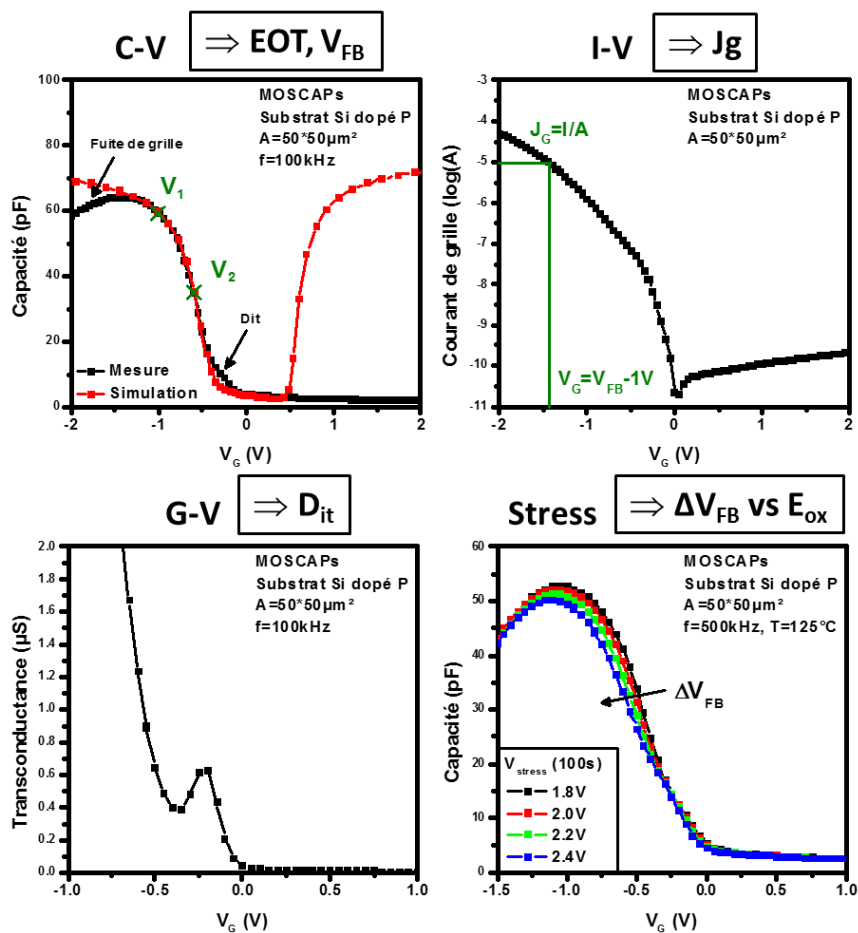


Figure 4.7 : Exemples de mesures C-V, I-V, G-V et de stress permettant d’extraire l’EOT, la tension de bande plate (V_{FB}), le courant de fuite de grille (J_g), la densité d’états d’interface (D_{it}) et d’évaluer la fiabilité NBTI des capacités MOS.

L’extraction de ces paramètres est effectuée par un traitement statistique et semi-automatique décrit dans [Charbonnier11] sur une dizaine de capacités MOS. En particulier, les mesures expérimentales C-V sont fittées par un modèle analytique entre deux tensions, une dans la zone d’accumulation forte (V_1), où des fuites élevées peuvent perturber la mesure de la capacité, et une après le début de la zone d’accumulation (V_2) afin d’exclure la contribution des états d’interface. Le modèle utilisé ainsi que les méthodes d’analyse de la courbe simulée permettent alors d’une part de corriger les effets quantiques (e.g. la contribution du *darkspace*) afin d’extraire l’EOT de l’empilement $\text{SiO}_2/\text{HfO}_2$, mais aussi une

extraction plus fiable de la tension de bande plate [Leroux07]. Le courant de fuite de grille est ensuite extrait de la courbe I-V à une tension d’*overdrive* constante de $V_G=V_{FB}-1V$ afin de comparer différentes variantes de grille. A partir de ces extractions, les figures de mérites **Jg vs EOT** et **ΔV_{midgap} vs EOT** seront données pour chaque variante de grille. La tension ΔV_{midgap} correspond à l’écart entre la valeur de V_{FB} et celle du midgap du silicium (4.61eV). Une valeur positive de ΔV_{midgap} correspondra alors à une grille de type P (V_T faible sur pMOS) et inversement. En outre, la densité d’états d’interface peut être extraite à partir d’une modélisation du pic de la courbe G-V [Nicollian65]. Enfin, la fiabilité des capacités MOS sera évaluée en mesurant la variation de la tension de bande plate sur des courbes C-V après contrainte électrique par différentes tensions négatives appliquées pendant 100s. Afin de comparer les variantes de grille de manière plus rigoureuse, la variation ΔV_{FB} sera tracée en fonction du champ électrique à travers les diélectriques et donnée par la formule :

$$E_{ox} = \frac{V_{stress} - V_{FB}}{EOT}$$

Si ces études sur capacités MOS sont pratiques pour étudier une large gamme de variantes de grille rapidement, les mesures ne représentent néanmoins pas forcément de manière absolue les propriétés obtenues sur un transistor. Par exemple, les flancs de la grille sur un transistor n’étant pas immédiatement protégés après gravure, une recroissance latérale de l’oxyde piédestal peut se produire, menant à une EOT plus importante sur transistors en particulier pour des grilles courtes. Aussi, si la tension de seuil dépend bien de la tension de bande plate, d’autres effets peuvent modifier sa valeur sur des transistors : effets de canaux courts, diffusion de dopants des sources/drains dans le canal, etc. Enfin, une autre implication de l’utilisation des MOSCAPs se porte sur la mesure de la fiabilité. En effet, la détermination de la variation de V_{FB} par des mesures C-V sur MOSCAPs est relativement longue (~40ms), temps qui est suffisant pour que la composante recouvrable de la dégradation NBTI soit relaxée. Les dégradations de V_{FB} sur MOSCAPs sont donc essentiellement corrélées à la formation de D_{it} . Au contraire, l’évaluation de la fiabilité NBTI ou PBTI sur transistors est effectuée par une mesure ponctuelle ($t < 1ms$) du courant de drain linéaire à V_G =tension de stress et avec un temps de stress plus long ($t=1000s$). La variation de $I_{D,lin}$ peut alors être convertie en une variation de V_T qui est également tracée en fonction du champ électrique. Des critères de fiabilité sont alors choisis pour évaluer la viabilité d’une technologie. Ces critères peuvent être formulés de plusieurs manières équivalentes, par exemple dans le cas du nœud 14nm de STMicroelectronics (avec une tension d’utilisation de 0.9V) :

- Une dégradation de V_T inférieure à 50mV à $V_G=0.9V$ après 5 ans d’utilisation
- Une tension $V_{DD}=V_G$ utilisable supérieure à 0.9V lorsque la dégradation de V_T atteint 50mV après 5 ans d’utilisation
- Un temps de vie supérieur à 5 ans lorsque la dégradation de V_T dépasse 50mV à $V_G=0.9V$

A noter que les mesures de fiabilité sur transistors comme sur MOSCAPs doivent alors être effectuées à une température élevée ($T=125^\circ C$ dans notre cas) et avec des tensions de grilles bien plus élevées que celle utilisée ($V_G > V_{DD}=0.9V$) afin d’accélérer la dégradation de V_T . Des modèles peuvent ensuite être utilisés pour extrapoler ces mesures à des valeurs représentant les conditions de fonctionnement du transistor, par exemple pour extraire un temps de vie (TTF – *Time to Failure*) à $V_G=V_{DD}$ avec l’expression :

$$TTF = (\Delta V_{T_MAX})^{\frac{1}{n}} * V_G^{-\frac{\gamma}{n}} * \exp(-\frac{A_0}{n})$$

Où ΔV_{T_MAX} correspond au critère de dégradation de la tension de seuil autorisée (50mV), V_G la tension de contrainte électrique utilisée ($V_G=\pm 1.4V, 1.6V, 1.8V, 2.0V$) et A_0, γ et n sont des paramètres d’extrapolation du modèle utilisé.

Ainsi, bien que les mesures sur capacités MOS diffèrent quelque peu de celles sur transistors, elles permettent de rendre compte des variations des propriétés de l’empilement de grille, variations que l’on retrouvera également sur transistors.

4.2 Influence du budget thermique

Dans un premier temps, nous avons étudié les effets sur les différentes propriétés de l’empilement de grille de trois budgets thermiques correspondant à des étapes de l’intégration :

- **1050°C, 2s** : comme référence haute température (HT) correspondant au budget thermique d’une activation de dopants standard.
- **600°C, 1min** : correspondant au budget thermique d’une activation par SPER. Ce budget ne prend pas en compte les autres étapes (e.g. espaceurs de grille, épitaxie) et n’est pas réaliste vis-à-vis de la fabrication d’un transistor. Il servira toutefois de référence basse température (LT).
- **630°C, 5h** : correspondant à la température de dépôt des espaceurs de grille standard $\text{SiN}_{630^\circ\text{C}}$. La durée réelle du dépôt est plus faible (~2h) mais a été surévaluée dans cette partie.

Afin de simuler ces budgets thermiques, un recuit sous ambiance inerte (N_2) a été réalisé après l’implantation arsenic du poly-Si (Figure 4.8a).

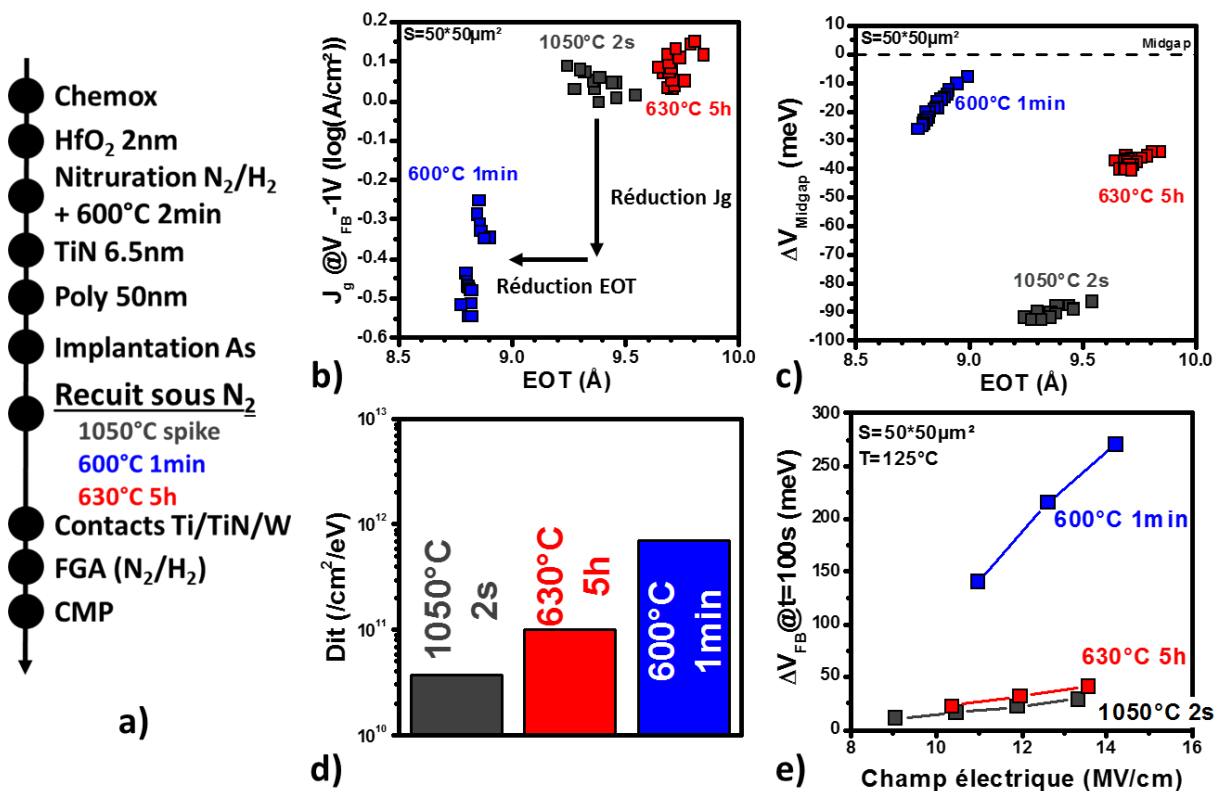


Figure 4.8: a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS avec des variantes de recuits après poly-Si.

Nous voyons tout d’abord que la réduction du budget thermique de 1050°C 2s à 600°C 1min mène à une diminution de l’EOT d’environ 0.5Å (Figure 4.8b). Cette diminution de l’EOT est liée à une épaisseur physique du SiO_2 plus fine comme confirmées par les mesures sur des observations TEM (Figure 4.9). Ceci indique que lors d’un recuit à haute température de la grille, une réoxydation du SiO_2 interfacial se produit. L’origine de l’oxygène peut être multiple et sera abordée dans d’autres sections. Les

observations TEM montrent également que le HfO_2 s'est partiellement cristallisé avec le recuit 1050°C 2s, ce qui peut expliquer l’augmentation du courant de fuite, les joints de grains du poly- HfO_2 pouvant agir comme des chemins préférentiels de conduction pour les porteurs [Watanabe05].

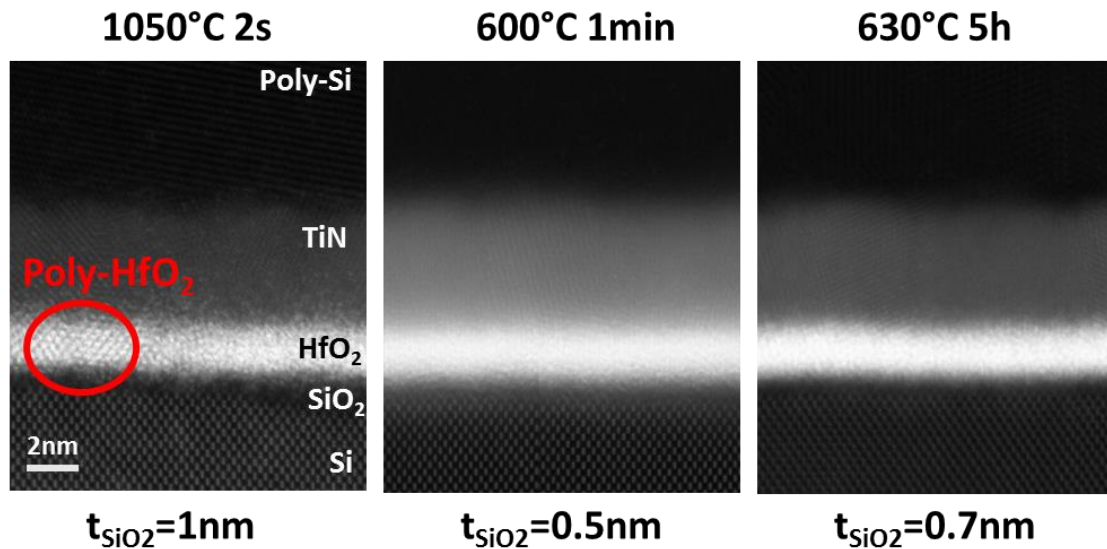


Figure 4.9 : Observations HAADF-STEM (High-Angle Annular Dark-Field Scanning TEM) de grilles après différents recuits four après poly-Si.

En revanche, le recuit intermédiaire 630°C 5h a mené à une dégradation de l'EOT et à un courant de fuite légèrement supérieur par rapport à la référence HT malgré un oxyde piédestal physiquement plus fin et un HfO_2 visuellement amorphe, signifiant alors que la constante diélectrique équivalente des couches $\text{SiO}_2/\text{HfO}_2$ a diminué. Ceci peut indiquer la formation d’un silicate de permittivité plus faible entre les deux oxydes sous l’effet d’un recuit de longue durée. Le recuit sous N_2 entraîne également un décalage de la tension de bande plate des MOSCAPs vers des valeurs négatives lorsque la température augmente (Figure 4.8c). Comme abordé dans l’introduction du chapitre, cette variation est essentiellement due au travail de sortie du métal TiN [Charbonnier10]. Par ailleurs, on remarque une densité d’états d’interface importante ($D_{it} \sim 10^{12} \text{cm}^2/\text{eV}$) dans le cas du recuit 600°C 1min (Figure 4.8d) et qui peut être due à un SiO_2 interfacial dont l’épaisseur est inférieure aux $6\text{-}7\text{\AA}$ nécessaires pour obtenir les propriétés isolantes optimales du SiO_2 [Tang08]. Le budget thermique intermédiaire 630°C 5h mène à des D_{it} plus faibles mais toujours plus élevés que la référence HT. La diminution des D_{it} avec l’augmentation du budget thermique a également été observée entre 1025°C et 1100°C et attribuée à un effet de guérison de défauts à haute température dans [Song06]. Enfin, les mesures de stress montrent une très faible variation de la tension de bande plate avec un recuit 1050°C 2s (Figure 4.8e), conforme à la température seuil à 800°C permettant d’obtenir une bonne fiabilité trouvée par [Veloso14], contrairement au budget thermique 600°C 1min dont la fiabilité est fortement dégradée. Il est en revanche intéressant de noter que le recuit 630°C 5h mène à une fiabilité équivalente au recuit HT.

Dans un second temps, nous avons diminué le budget thermique des recuits intermédiaires, tout d’abord en faisant varier uniquement la température de 600°C à 525°C , cette dernière correspondant à la température de dépôt du poly-Si (**Error! Reference source not found.a**). A noter que des études sont en cours afin de diminuer la température du dépôt poly-Si, notamment en remplaçant le précurseur SiH_4 standard par du Si_2H_6 permettant d’obtenir des vitesses de croissance plus rapides à basse température.

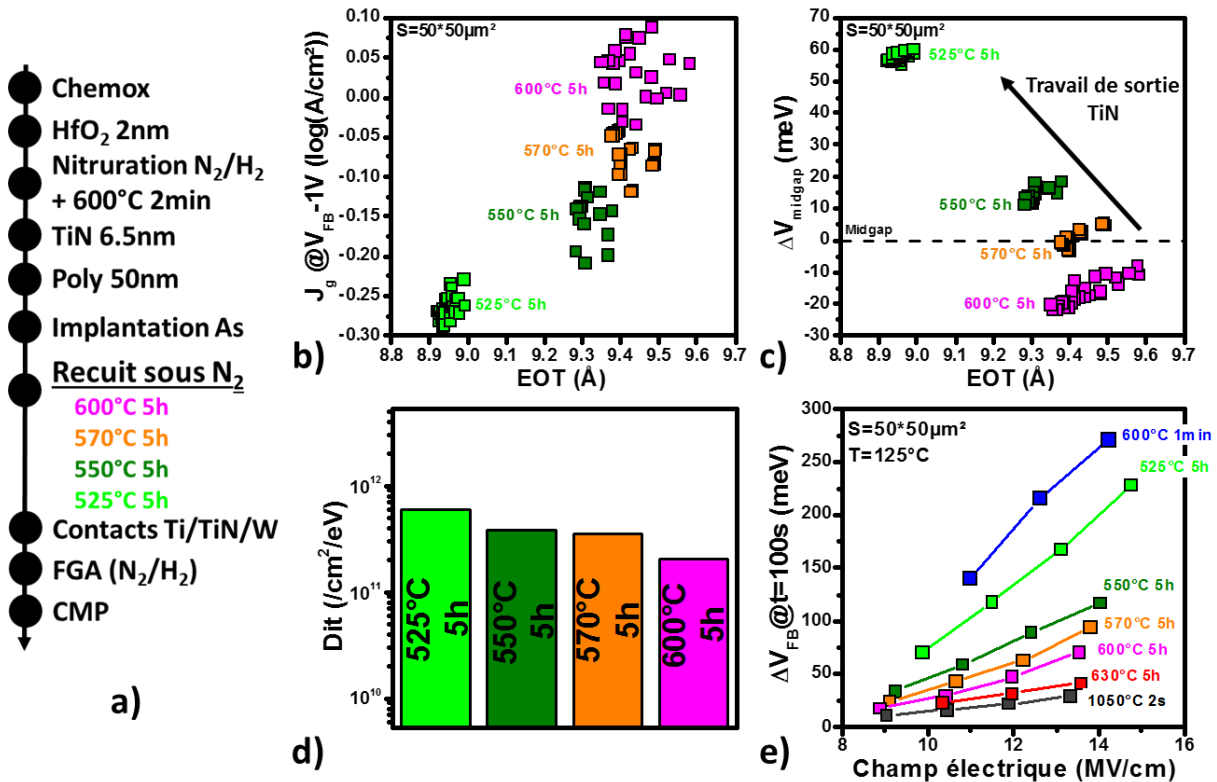


Figure 4.10 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS avec des variantes de recuits (budgets thermiques intermédiaires) après poly-Si.

La Figure 4.10b montre que la réduction du budget thermique mène à des gains sur l'EOT et le courant de fuite, ce qui peut étayer l'hypothèse précédente d'une formation d'un silicate défavorable lors d'un recuit de longue durée. De plus, la diminution de la température de 600°C à 525°C décale la tension de bande plate d'environ +100meV toujours lié au travail de sortie du TiN (Figure 4.10c). En outre, on remarque également une légère et linéaire augmentation de la densité d'états d'interface avec la diminution de la température (Figure 4.10d). Enfin, les mesures de stress montrent que si la réduction de la température dégrade progressivement la fiabilité, une température seuil semble apparaître entre 525°C et 550°C où la fiabilité est dégradée plus rapidement (Figure 4.10e). Ce résultat est important car il indique que la fiabilité peut être liée à une transition physique plutôt qu'à des effets thermique et/ou temporel (auquel cas une bonne fiabilité à bas budget thermique serait impossible). Des études supplémentaires sont néanmoins nécessaires et seront abordées plus loin dans ce chapitre.

Par ailleurs, la diminution du temps de recuit de 5h à 1h à une température fixe de 570°C (Figure 4.11a) confirme la plupart des effets du budget thermique observés lorsqu'on est au-dessus d'une température seuil (550°C) : réduction de l'EOT et du courant de fuite (Figure 4.11b), décalage positif du travail de sortie du TiN (Figure 4.11c) et légère dégradation de la fiabilité (Figure 4.11e) avec la réduction du temps. On remarque toutefois que si l'augmentation de la température diminuait la densité d'états d'interface, l'augmentation du temps à une température fixée augmente elle légèrement la quantité de liaisons pendantes (Figure 4.11d). Une explication possible est qu'il existe un compromis entre la passivation des états d'interface (limitée par l'énergie thermique) et leur dépassivation (limitée par la diffusion d'une espèce, par exemple l'hydrogène). Nous reviendrons dans la Section 4.5 sur les mécanismes qui peuvent exister à l'interface entre l'oxyde piédestal et le silicium.

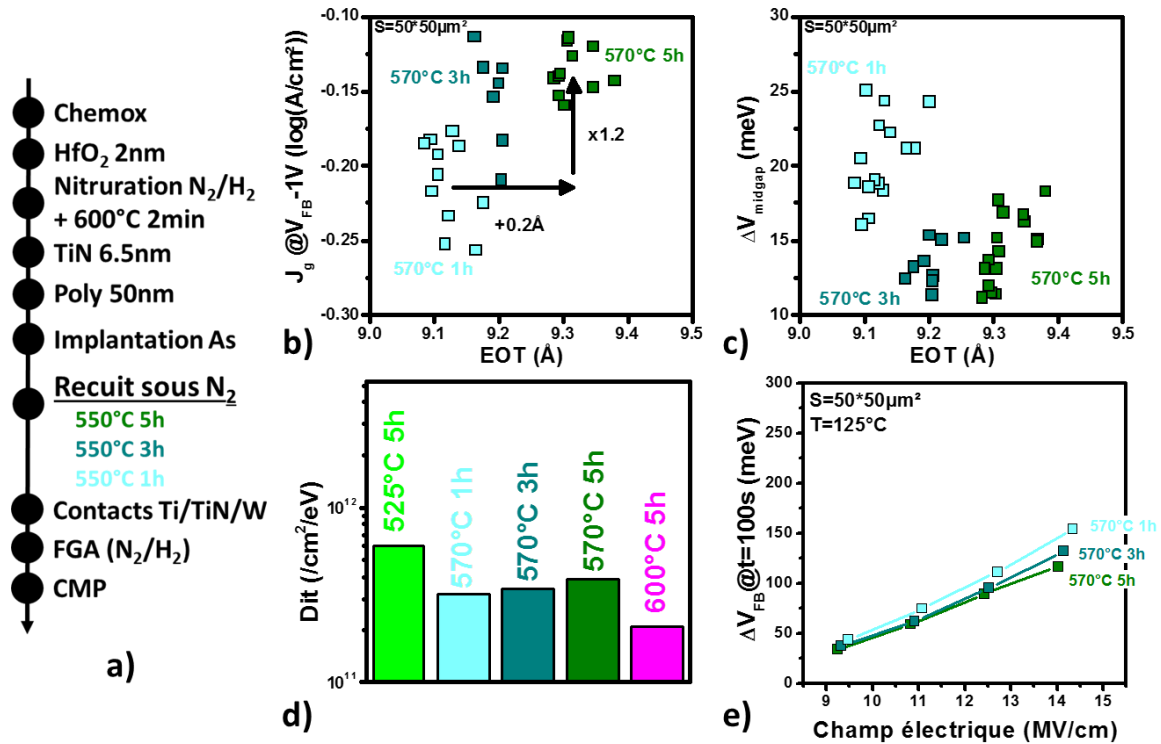


Figure 4.11 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS avec des variantes de recuits (variation de temps) après poly-Si.

Pour terminer, l'influence du budget thermique sur un même empilement de grille a également été étudiée sur des transistors (Figure 4.12a). Afin de réduire la température totale en dessous de 630°C (LT – Low Temperature), une épitaxie sources et drains SiGe (au lieu de Si à 750°C) et une activation de dopants par SPER (au lieu du recuit 1050°C 2s) ont été intégrées. Les espaceurs low-k (T=400°C avec le SiCO) ainsi que l'épithaxie dépôt/gravure cyclée (T=500°C) n'étant pas encore développés lors du début de cette étude, ils n'ont pas encore pu être intégrés pour réduire davantage le budget thermique.

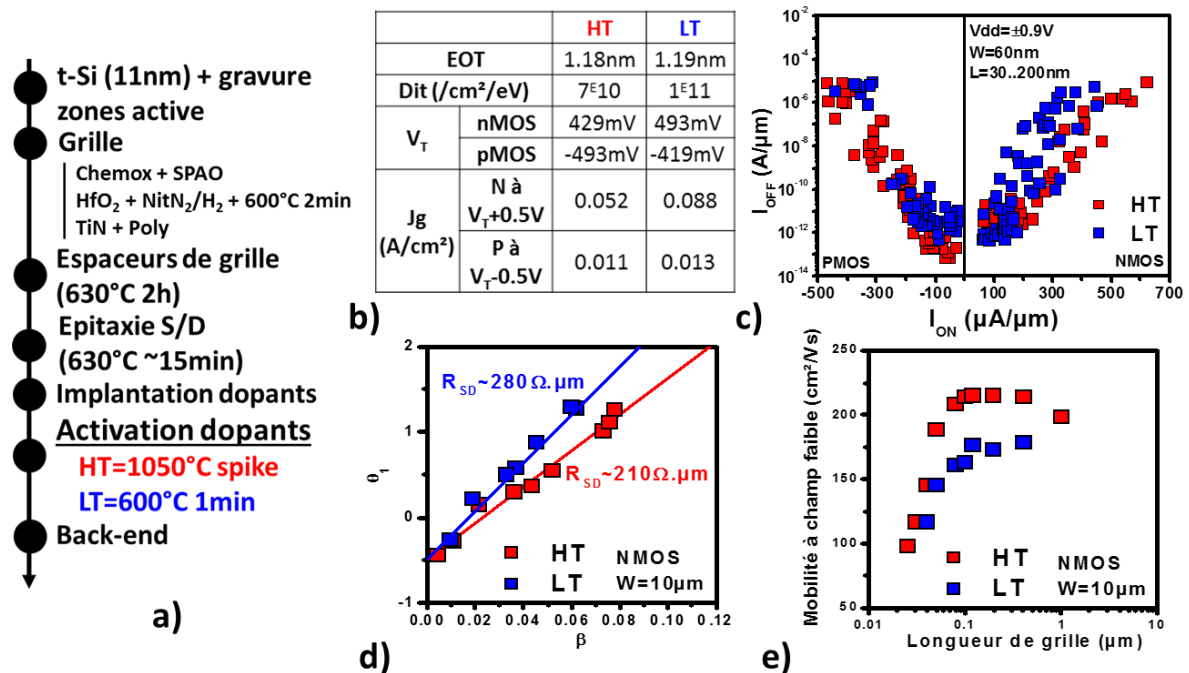


Figure 4.12 : a) Process flow, b) Tableau récapitulatif des propriétés de la grille (W/L=10µm/10µm), c) Compromis I_{ON}/I_{OFF}, d) Extraction de R_{access} et e) de mobilité à champ faible en utilisant la méthode-Y de transistors fabriqués avec un budget thermique élevé (HT) ou faible (LT).

La Figure 4.12b résume les valeurs d'EOT, de courant de fuite de grille (pris à $V_G=V_T\pm 0.5V$), de tension de seuil et de D_{it} extraites pour les transistors n et pMOS fabriqués avec un budget thermique élevé (HT) ou plus faible (LT). On retrouve alors sensiblement les mêmes résultats obtenus sur MOSCAPs (entre 1050°C 2s et 630°C 5h) lorsque le budget thermique est réduit : EOT, courant de fuite et D_{it} plus élevés ainsi qu'une tension de seuil pour les nMOS (respectivement pMOS) plus élevée (plus faible) dû au décalage positif du travail de sortie du TiN. Par ailleurs, la Figure 4.12c montre une très légère dégradation du compromis Ion/Ioff pour les transistors fabriqués à 630°C. Cette dégradation est corrélée à une résistance d'accès plus importante (Figure 4.12d) et attribuée à un manque de dopants sous les espaceurs lors d'une activation à basse température par SPER comme soulevé dans le Chapitre 1. En outre, la Figure 4.12e montre une mobilité à champ faible également réduite par rapport à la référence HT et qui peut être due à une densité d'états d'interface plus importante causant des effets de *scattering* [Fang68].

La fiabilité des transistors a ensuite été évaluée grâce à des mesures de PBTI sur les nMOS et de NBTI sur les pMOS.

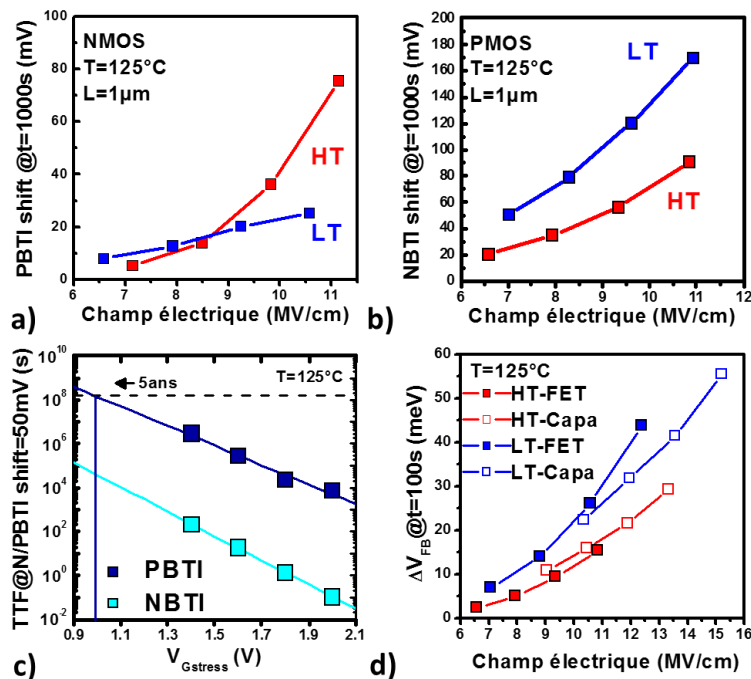


Figure 4.13 : Mesures de ΔV_T vs E_{ox} sous une contrainte électrique a) positive (nMOS) et b) négative (pMOS) pour des transistors fabriqués avec un budget thermique élevé ou faible, c) Extrapolation des temps de vie pour la variante LT et d) Comparaison des mesures de NBTI statiques pour des temps de stress de 100s sur transistors et MOSCAPs.

La Figure 4.13a montre une excellente stabilité vis-à-vis d'une contrainte électrique positive des nMOS fabriqués à basse température, avec un décalage de la tension de seuil inférieur à 50mV quel que soit la tension de stress appliquée. Ce résultat indique que la fiabilité PBTI, qui est liée à la présence de défauts dans le HfO_2 , ne devrait pas être problématique pour l'intégration 3D Séquentielle. En revanche, la Figure 4.13b montre une dégradation plus importante de la fiabilité NBTI par rapport à la référence HT. Les extractions de temps de vie montrent alors que si les transistors nMOS fabriqués avec un budget thermique faible passent largement le critère des 5ans (Figure 4.13c), les pMOS exhibent un temps de vie inférieur à un an, bien en deçà du temps de vie des transistors fabriqués à haute température (TTF=5ans). Ces résultats sont donc contradictoires avec ceux sur MOSCAPs où une bonne fiabilité était obtenue à 630°C et peuvent être expliqués par deux raisons. Tout d'abord, comme soulevé dans la Section 4.1, la mesure rapide du NBTI sur transistor met en jeu une composante

recouvrable non observée sur capacités MOS et qui peut être plus importante lorsque le budget thermique est diminué. La Figure 4.13d montre notamment que lorsque des mesures statiques (i.e. lente) du NBTI pour un même temps de stress de 100s sont effectuées, des résultats similaires entre transistors et capacités MOS sont obtenus, avec un décalage sur le champ électrique dû à des différences de tension de seuil (ou bande plate) et d’EOT. On note tout de même une dégradation légèrement plus prononcée de la variante LT sur transistor. Les différences peuvent donc également être attribuées à un budget thermique plus faible (630°C ~2h) comparé au recuit four effectué (630°C 5h) sur capacités MOS.

En conclusion, nous avons vu dans cette section que le budget thermique subi par la grille modifie fortement ses propriétés. La réduction du budget thermique est bénéfique pour la réduction de l’EOT et du courant de fuite et qui pourrait être liée d’une part à une réoxydation moins importante du SiO₂ interfacial, et d’autre part à une plus faible diffusion du silicium dans le high-k pour former un silicate. De plus, le décalage positif du travail de sortie du TiN avec la diminution de la température peut être intéressant pour l’obtention d’une faible tension de seuil pour les transistors pMOS. En revanche, un budget thermique faible dégrade fortement la qualité de l’interface entre le canal et l’oxyde piédestal SiO₂ ainsi que la mobilité des porteurs dans le canal. Enfin, si la fiabilité des transistors nMOS (PBTI) ne semble pas poser de problème, la fiabilité NBTI des transistors pMOS est la principale difficulté pour une grille froide. Une température seuil entre 525°C et 550°C a toutefois été identifiée au-dessus de laquelle la fiabilité semble s’améliorer, indiquant la présence d’une transition physique.

Dans les sections suivantes, nous allons présenter et étudier l’influence de différents procédés leviers à basse température sur les propriétés de la grille, avec pour objectif principal l’amélioration de la fiabilité NBTI.

4.3 Procédés leviers à bas budget thermique

Dans cette section, nous présenterons l’effet sur les propriétés de la grille de trois procédés différents : la nitruration du high-k, la densification de l’oxyde chimique et la passivation par deutérium. Ces procédés sont déjà relativement connus et utilisés dans la littérature, mais nous verrons ici d’une part des procédés dont le budget thermique est limité (en particulier pour la nitruration et la densification), et d’autre part l’influence de ces procédés dans un process flow froid.

4.1.1 Nitruration du HfO₂

Comme nous l’avons vu dans l’introduction, l’incorporation d’azote dans l’oxyde de grille a plusieurs bénéfices : augmentation de la permittivité du SiO₂, formation d’une barrière de diffusion pour le bore du poly-Si, augmentation de la température de cristallisation du HfSiO ou HfO₂. Par ailleurs, des études de nitruration plasma sous ambiance N₂ ou de nitruration thermique (T>650°C) sous NH₃ d’un empilement SiON/HfO₂ ont été réalisées dans [Dai13] sur lesquelles nous reviendrons dans cette section.

Dans un premier temps, nous avons évalué un procédé de nitruration plasma sur l’empilement SiO₂/HfO₂ développé au Leti et réalisé à 250°C. Deux ambiances différentes ont été utilisées : N₂ seul ou N₂/H₂ (Figure 4.14a). De plus, dans le cas de la nitruration N₂ seul, deux variantes supplémentaires

ont été étudiées : une avec un temps de procédé rallongé et une autre avec une pression réduite, permettant dans les deux cas d’augmenter la quantité d’azote incorporée. La nitruration plasma incorpore l’azote en volume et en position interstitielle du HfO_2 . Un recuit PNA (*Post-Nitridation Anneal*) est alors nécessaire afin de greffer l’azote sous forme de liaisons Hf-O-N, sans quoi l’azote exodiffuse avec le temps.

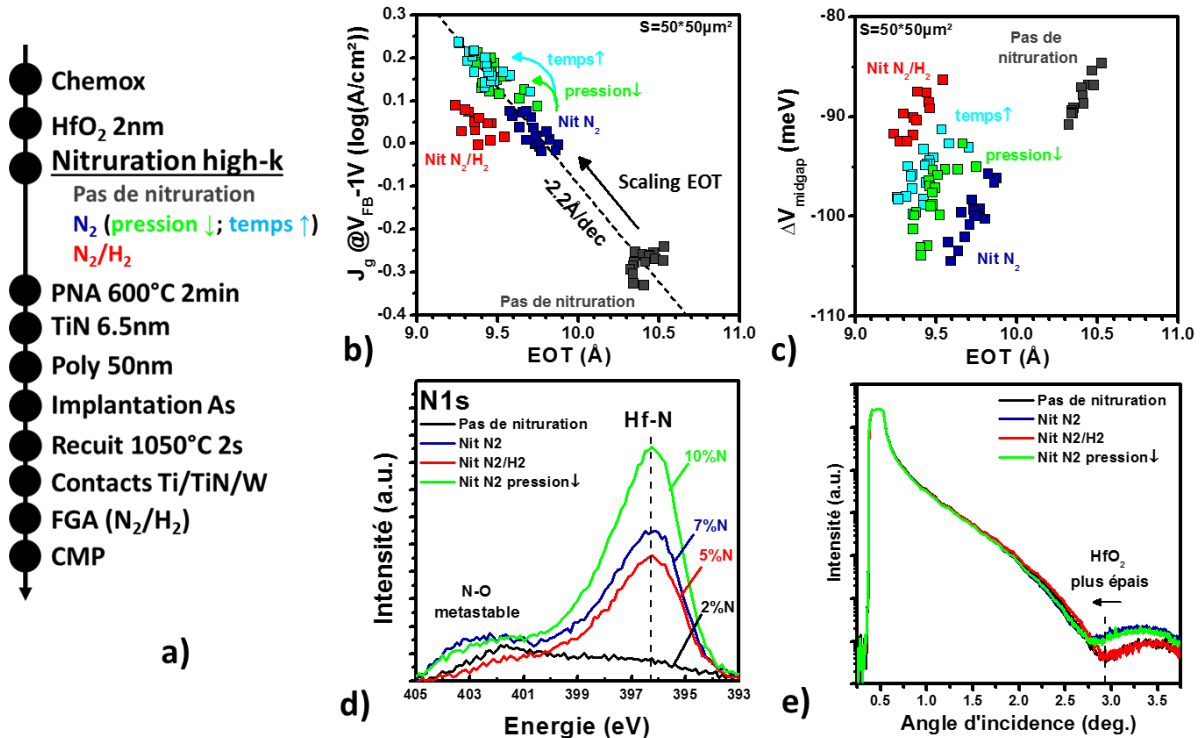


Figure 4.14 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT de capacités MOS et d) Spectres XPS et e) Spectres XRR d’un empilement $\text{SiO}_2/\text{HfO}_2$ avec des variantes de nitruration du high-k.

Tout d’abord, la Figure 4.14b montre une réduction de l’EOT ainsi qu’une augmentation du courant de fuite avec la nitruration N_2 seul, indiquant une réduction de l’épaisseur physique des diélectriques (SiO_2 et/ou HfO_2) qui s’accompagne alors d’une augmentation de l’effet tunnel. Cette variation suit une pente d’une décade de dégradation en courant pour une diminution de 2.2Å en EOT et correspond en fait essentiellement à une réduction de l’épaisseur du SiO_2 [Leroux07]. On remarque de plus une corrélation entre l’EOT et la dose d’azote incorporée (extraite par des mesures XPS – *X-Ray Photoelectron spectrometry* – sur l’empilement $\text{SiO}_2/\text{HfO}_2$ seulement) lorsque la pression diminue (Figure 4.14d). Ceci peut indiquer que l’azote formant des liaisons Hf-N dans le high-k agit comme une barrière de diffusion à l’oxygène qui permet de limiter la recroissance du SiO_2 lors du recuit après poly-Si. Une réduction de l’EOT avec la nitruration plasma sous N_2 a également été observée dans [Dai13]. Celle-ci a été attribuée à une augmentation de la permittivité des couches $\text{SiO}_2/\text{HfO}_2$ uniquement, ce qui n’expliquerait cependant pas l’augmentation du courant de fuite.

La nitruration dans une ambiance N_2/H_2 mène, elle, à une EOT plus faible et un courant de fuite équivalent à la nitruration N_2 malgré une dose d’azote incorporée plus faible. Une explication possible est que dans le cas de la nitruration sous N_2 , des liaisons N-O sont formées qui sont observables sur les spectres XPS. Ces liaisons se forment dans le HfO_2 ce qui peut expliquer une épaisseur plus importante du high-k observée sur les spectres XRR (Figure 4.14e). Les liaisons N-O étant métastables, le recuit post poly-Si peut casser ces liaisons et libérer de l’oxygène qui peut induire une légère recroissance du SiO_2 . Au contraire, l’ajout d’hydrogène peut réagir avec l’oxygène pour former des liaisons O-H ou de l’eau qui sont ensuite désorbées lors du recuit PNA, aucun pic lié aux liaisons N-O n’apparaissant en

effet après une nitruration N_2/H_2 . Toutefois, la diminution de l'épaisseur physique du SiO_2 avec une nitruration N_2/H_2 n'est probablement pas le seul effet en jeu dans la mesure où le courant de fuite ne suit pas la pente de $2.2\text{\AA}/\text{décade}$. Cette fuite plus faible qu'attendue dans le cas d'un SiO_2 plus fin peut alors signifier une constante diélectrique plus élevée, par exemple grâce à une inhibition de la formation d'un silicate de permittivité plus faible. Enfin, le très léger décalage de la tension de bande plate avec la nitruration peut éventuellement être attribué au phénomène de roll-off avec la diminution de l'EOT (Figure 4.14c).

Le recuit 1050°C 2s après poly-Si utilisé pour ces premières études masquant les effets de la nitruration du HfO_2 sur la densité d'états d'interface et sur la fiabilité, le budget thermique a été par la suite diminué à 525°C 5h (Figure 4.15a). De plus, nous avons également évalué l'influence d'une nouvelle technique de nitruration sous ambiance NH_3 assistée par un recuit laser DSA (cf. Chapitre 1) à 800°C ou à 1100°C pendant 0.5ms. A noter tout de même que le budget thermique du recuit DSA à 1100°C est trop élevé pour l'intégration 3D Séquentielle, mais sera utile pour évaluer les effets du laser.

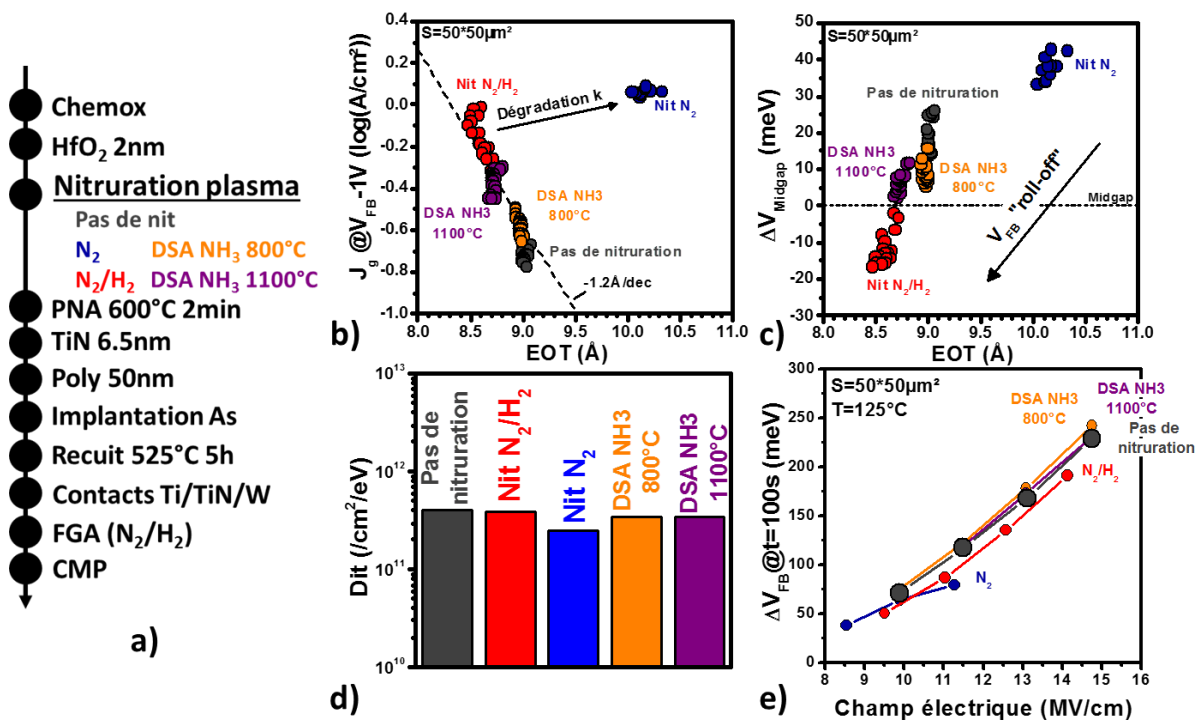


Figure 4.15 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{OX} de capacités MOS avec des variantes de nitruration du high-k.

Tout d'abord, la Figure 4.15b montre que si la nitruration N_2/H_2 présente une diminution de l'EOT relativement similaire à ce qui a été vu précédemment (avec cependant une pente différente), la nitruration sous N_2 a dégradé l'EOT avec un courant de fuite équivalent, indiquant une permittivité plus faible des couches SiO_2/HfO_2 . L'origine de cette constante diélectrique plus faible peut être liée à un délai important (~ 2 mois) entre le dépôt du HfO_2 et la nitruration sur ces échantillons qui occasionne une absorption plus importante d'humidité et/ou d'oxygène. La nitruration en présence d'hydrogène permettrait alors de réduire l'oxygène et/ou l'eau qui sont ensuite désorbés par le recuit post nitruration et donc d'être plus robuste. A noter que la différence de pente ($1.2\text{\AA}/\text{dec}$ au lieu de $2.2\text{\AA}/\text{dec}$ auparavant) peut indiquer tout de même une légère dégradation de la permittivité même avec la nitruration N_2/H_2 .

En ce qui concerne la nitruration sous NH_3 assistée par DSA, on remarque une faible variation à 800°C avec la référence. En revanche, la nitruration à 1100°C montre une légère réduction de l'EOT associée

à un courant de fuite plus important bien que la quantité d'azote incorporée semble quasiment identique pour les deux températures (Figure 4.16a), indiquant une réduction de l'épaisseur physique à plus haute température. A noter que comme dans le cas d'une nitruration sous N_2/H_2 , la nitruration sous NH_3 ne mène pas à la formation de liaisons métastables N-O. Les spectres XRR associés montrent que c'est en réalité l'épaisseur du HfO_2 qui a diminué avec le recuit $1100^\circ C$ (Figure 4.16b), expliquant également que la réduction de l'EOT se fait suivant une pente plus importante de $1.2\text{\AA}/\text{dec}$.

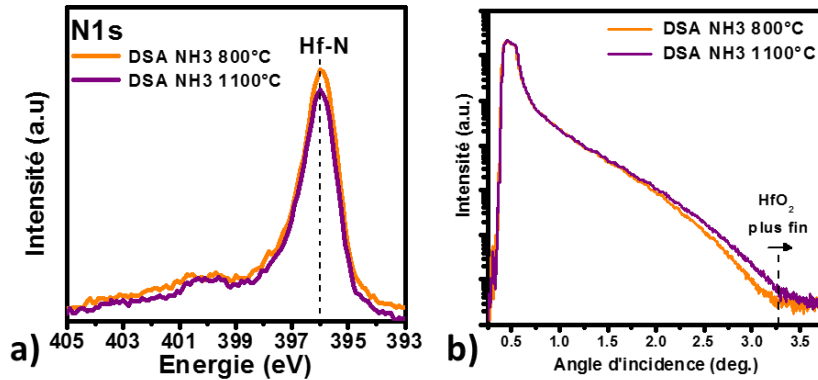


Figure 4.16 : Spectres a) XPS et b) XRR pour un empilement SiO_2/HfO_2 avec nitruration NH_3 assisté par DSA à 800 ou $1100^\circ C$.

Par ailleurs, la Figure 4.15c montre cette fois un plus fort effet de roll-off de la tension de bande plate qui peut être dû à la diminution de l'EOT sous une valeur seuil d'environ $\sim 10\text{\AA}$. En outre, la faible variation de la densité d'états d'interface quelle que soit la méthode de nitruration indique que ni le recuit post nitruration, ni le budget thermique du recuit laser ne diffuse l'azote jusqu'à l'interface SiO_2/Si (Figure 4.15d). Par conséquent, la nitruration du HfO_2 n'a pas non plus impacté de manière significative la fiabilité des MOSCAPs (Figure 4.15e).

Pour terminer, l'influence de la nitruration sous ambiance N_2/H_2 a été évaluée sur des transistors (Figure 4.17a).

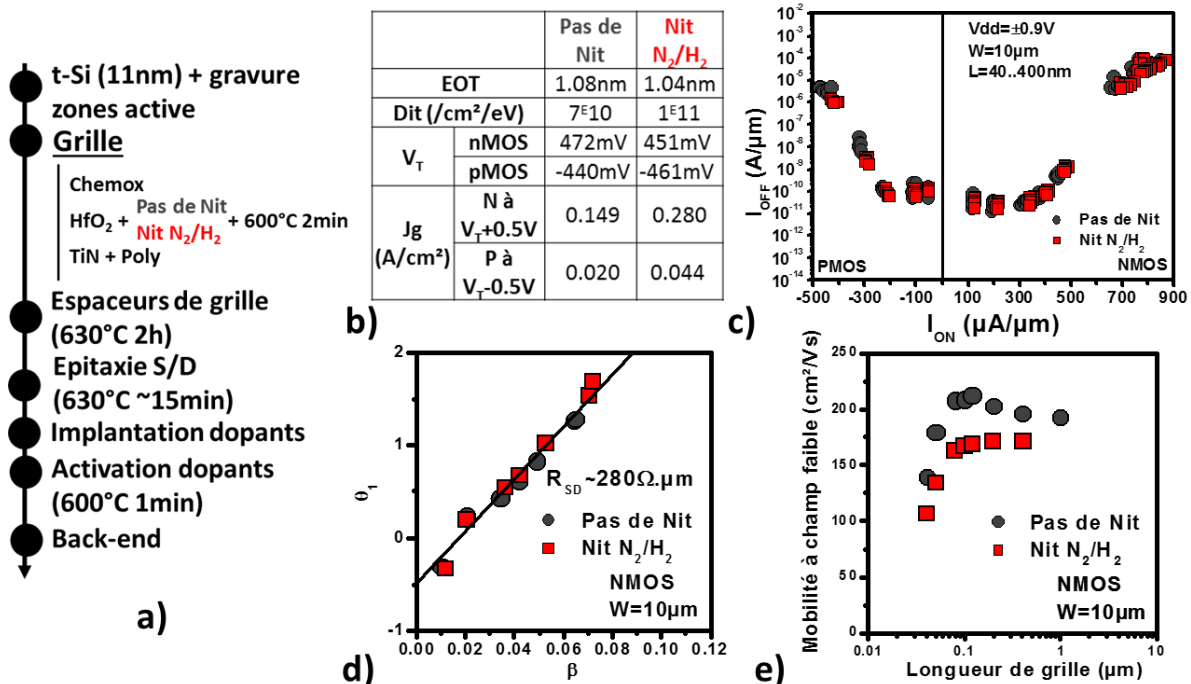


Figure 4.17 : a) Process flow, b) Tableau récapitulatif des propriétés de la grille ($W/L=10\mu m/10\mu m$), c) Compromis I_{off}/I_{on} , d) Extraction de R_{access} et e) de mobilité à champ faible en utilisant la méthode-Y de transistors fabriqués avec ou sans nitruration N_2/H_2 du high-k.

La Figure 4.17b confirme les comportements observés sur MOSCAPs : diminution de l'EOT et augmentation du courant de fuite ainsi que de faibles variations de la densité d'états d'interface et de la tension de seuil lorsque l'EOT est supérieure à 10Å avec la nitruration N₂/H₂. Néanmoins, malgré la réduction de l'EOT et des résistances d'accès similaires (Figure 4.17b et d), les transistors avec grille nitrurée exhibent un compromis Ion/Ioff équivalent à une grille non nitrurée (Figure 4.17c). Ceci est d'une part dû à un courant de fuite de grille plus élevé augmentant le Ioff, et d'autre part lié à une mobilité des porteurs plus faible qui diminue le Ion (Figure 4.17e). La formation de liaisons Si-N près de l'interface SiO₂/Si lors de la nitruration et observée sur les mesures XPS (Figure 4.18) tend en effet à dégrader la mobilité des porteurs dû à des effets de Coulomb *scattering* [Casse11].

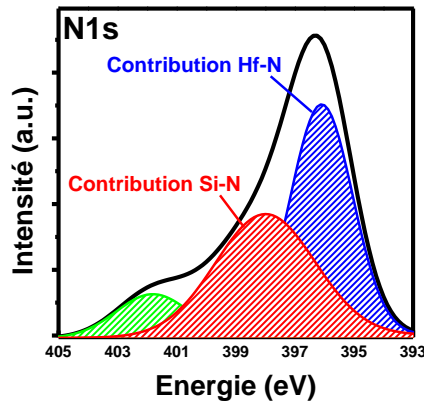


Figure 4.18 : Spectre XPS déconvolué d'un empilement SiO₂/HfO₂ après nitruration N₂/H₂ et recuit PNA 600°C 2min montrant les différentes contributions de liaisons avec l'azote.

Par ailleurs, l'incorporation d'azote dégrade la fiabilité NBTI des transistors, contrairement à ce qui était observé sur capacités MOS (Figure 4.19b). Cette dégradation est attribuée à la formation de pièges dans le volume du SiO₂ par l'azote [Garros09]. Le phénomène de piégeage/dépiégeage de trous sur ces défauts étant rapides, ils ne sont pas observables lors des mesures C-V pour les capacités MOS.

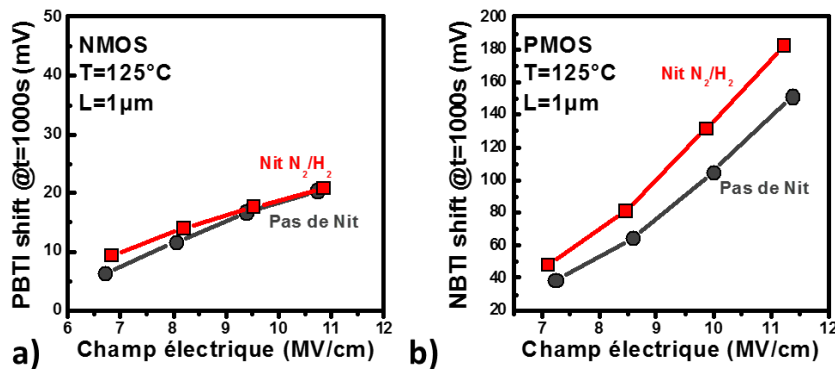


Figure 4.19 : Mesures de ΔV_T vs E_{ox} sous une contrainte électrique a) positive (nMOS) et b) négative (pMOS) pour des transistors fabriqués avec ou sans nitruration N₂/H₂ du high-k.

Ainsi, l'incorporation d'azote par nitruration permet de former une barrière de diffusion à l'oxygène, limitant la recroissance du SiO₂ interfacial lors d'un recuit. De plus, l'ajout d'hydrogène permet de réduire l'humidité et/ou l'oxygène résiduel absorbé dans le volume du HfO₂ lorsqu'il est à l'air et permet un gain supplémentaire sur l'EOT. Cependant, la formation de liaisons Si-N près de l'interface avec le canal mène à des effets de *scattering* des porteurs diminuant leur mobilité. Des optimisations peuvent néanmoins être apportées afin de contrôler le profil d'azote et limiter leur diffusion (e.g. diminution de la température du PNA). Enfin, la nitruration a également un effet néfaste sur la fiabilité NBTI, l'azote ayant tendance à favoriser la formation de défauts électriques dans le SiO₂.

4.1.2 Densification de l’oxyde chimique

L’intégration de high-k nécessite la formation d’un SiO₂ interfacial d’une part pour favoriser la nucléation du HfO₂ lors du dépôt, et d’autre part pour obtenir une meilleure interface avec le silicium. Afin d’obtenir un SiO₂ à la fois fin pour l’EOT et de bonne qualité, l’oxyde est généralement formé à haute température. En particulier, l’oxydation par dépôt eISSG (*enhanced in-situ steam generation*) permet d’obtenir un SiO₂ dense et bien contrôlé, mais requiert une température assez élevée (T>~800°C). Afin d’obtenir un oxyde piédestal de bonne qualité à basse température, des études de densification de l’oxyde chimique ont été réalisées. Quatre techniques de densification seront évaluées (Figure 4.20a) :

- SPAO (*Slot Plane Antenna Oxidation*) : oxydation plasma assistée par micro-ondes à 200°C
- DPO (*Decoupled Plasma Oxidation*) : oxydation plasma à 500°C
- RPO (*Remote Plasma Oxidation*) : oxydation plasma déporté (i.e. formation des radicaux dans une chambre qui sont ensuite déportés vers la chambre contenant la plaque et permettant de limiter les endommagements liés au plasma) à 500°C
- Laser millisecondes DSA sous une ambiance 50%N₂-50% O₂. Encore une fois, le budget thermique utilisé ici (1200°C, 0.5ms) est trop élevé pour l’intégration 3D Séquentielle mais permettra d’évaluer l’effet d’un budget thermique élevé sur l’oxyde piédestal.

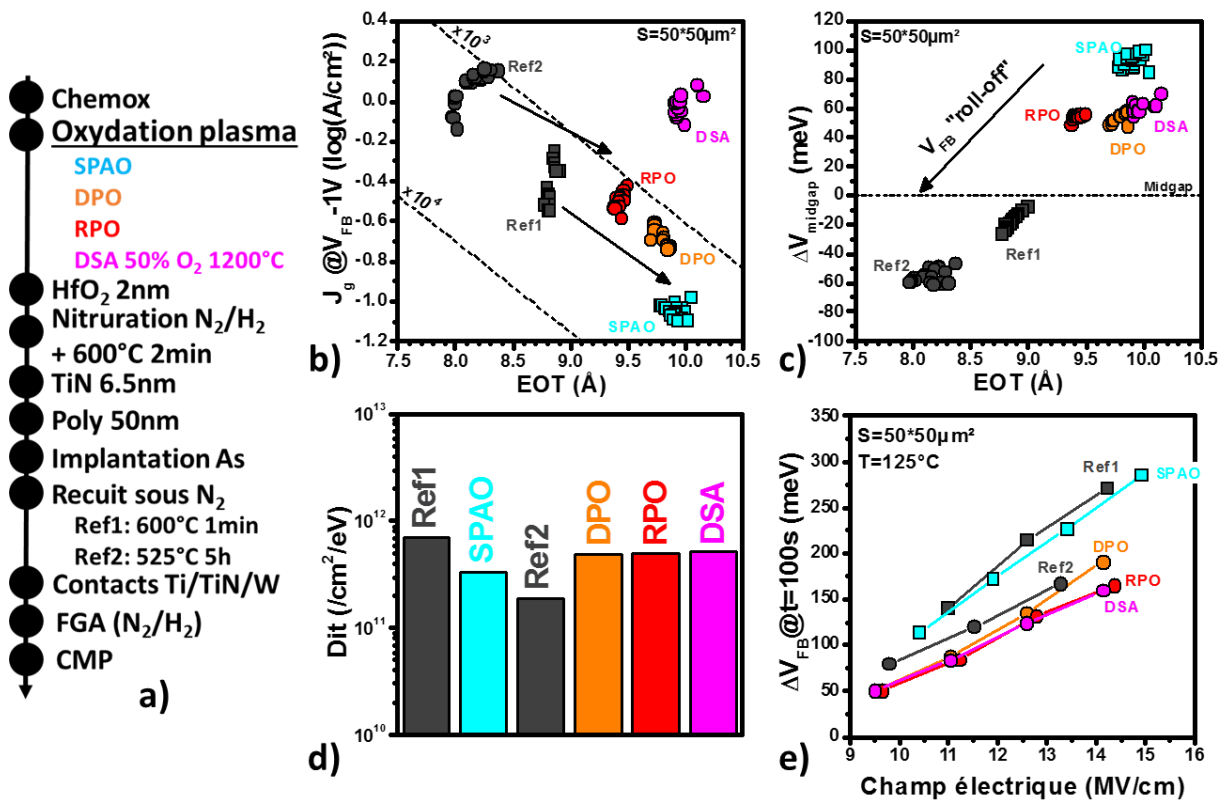


Figure 4.20 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) Dit et e) ΔV_{FB} vs EOX de capacités MOS avec des variantes de densification du SiO₂.

En particulier, concernant les techniques DPO, RPO et DSA, des études préliminaires chez Applied Materials ont été réalisées afin de trouver des conditions permettant de densifier l’oxyde tout en conservant une épaisseur mesurée par XPS supérieure à 6Å (sauf pour la variante DSA) afin de conserver de bonnes propriétés d’oxyde (Figure 4.21).

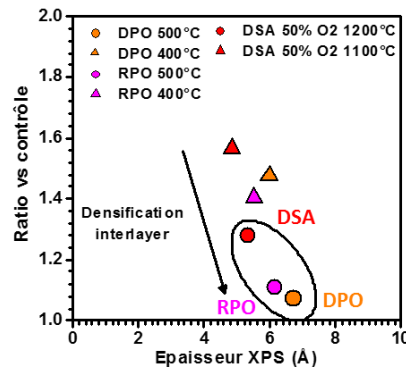


Figure 4.21 : Etude de la densification de l’oxyde chimique (par rapport à un oxyde déposé à haute température). Les procédés à plus haute température mène à une meilleure densification de l’oxyde tout en conservant une épaisseur mesurée par XPS supérieure à 6Å (excepté pour le DSA).

La Figure 4.20b montre une augmentation de l’EOT suivant une pente de $-2.2\text{Å}/\text{dec}$ due à une forte recroissance de l’oxyde piédestal quel que soit le procédé d’oxydation, sauf pour la variante DSA qui exhibe une dégradation du courant de fuite plus importante et qui est probablement dû à un SiO_2 inférieur à 6Å. Cette augmentation de l’EOT est accompagnée d’un effet de roll-off sur la tension de bande plate, avec un V_{FB} plus élevé pour une EOT plus grande (Figure 4.20c). Les mesures de D_{it} montrent une réduction de la densité de liaisons pendantes avec le SPAO grâce à une meilleure qualité d’oxyde (Figure 4.20d). En revanche, les procédés DPO, RPO et DSA ont tous les trois menés à une augmentation des D_{it} . Ceci peut être dû à un temps d’enchaînement non négligeable entre l’oxydation chimique et la densification de l’oxyde par rapport à la référence non densifiée. Enfin, la Figure 4.20e montre que l’obtention d’un oxyde piédestal plus dense améliore légèrement leur stabilité lors d’une contrainte électrique. Cette variation reste toutefois insuffisante comparée à la référence à haute température.

Le procédé SPAO à 200°C du Leti a ensuite été intégré sur des transistors (Figure 4.22a).

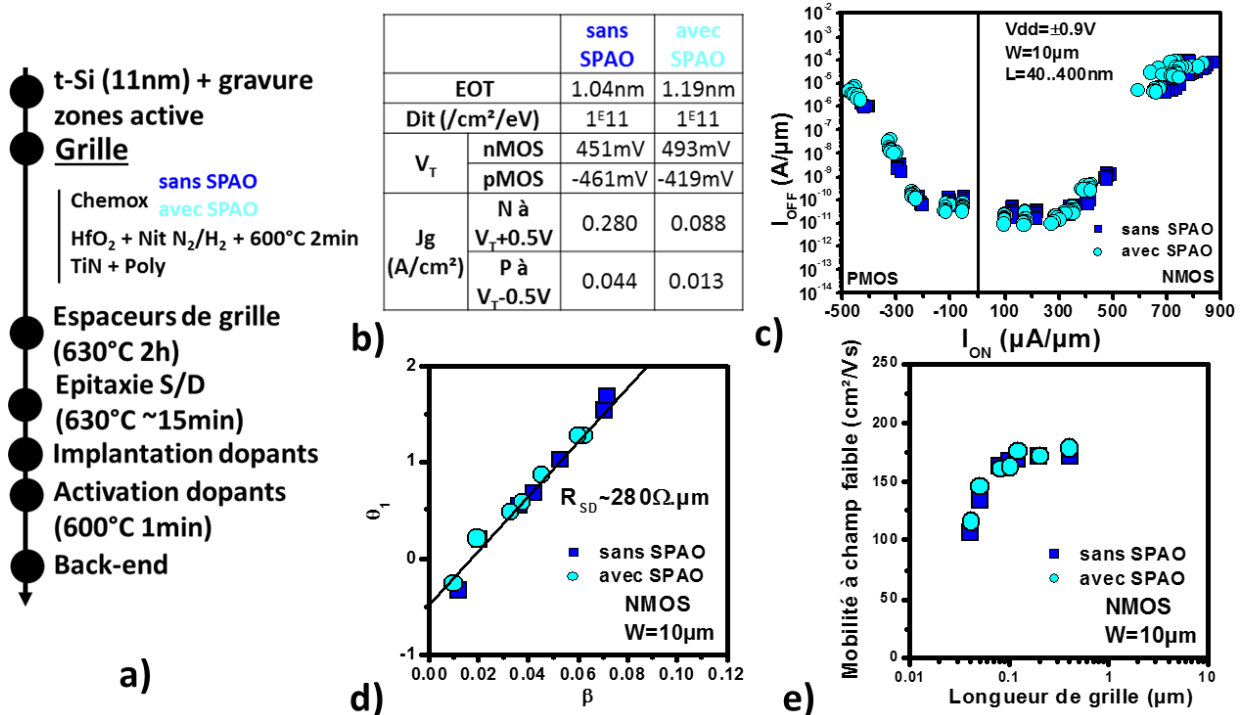


Figure 4.22 : a) Process flow, b) Tableau récapitulatif des propriétés de la grille ($W/L=10\mu\text{m}/10\mu\text{m}$), c) Compromis I_{on}/I_{off} , d) Extraction de R_{access} et e) de mobilité à champ faible en utilisant la méthode-Y de transistors fabriqués avec ou sans oxydation plasma SPAO du SiO_2 .

La Figure 4.22b confirme une recroissance de 1.5Å en EOT observée sur MOSCAPs et accompagnée d’une réduction du courant de fuite. Une augmentation (respectivement réduction) de la tension de seuil des nMOS (pMOS) est également observée dû à l’effet de roll-off. En revanche, la densité d’états d’interface mesurée sur transistors n’a pas été améliorée par l’oxydation plasma. Les mobilités et les résistances d’accès similaires résultent alors en des compromis Ion/Ioff équivalents avec ou sans oxydation plasma malgré une EOT plus élevée dans le premier cas (Figure 4.22c à e).

Enfin, les mesures de PBTI et NBTI ne montrent aucune amélioration de la fiabilité avec un SiO₂ interfacial plus épais (Figure 4.23). Ceci indique probablement que le SPAO, s’il épaissit le SiO₂, ne l’a peut-être pas réellement densifié dû à une température de procédé trop faible.

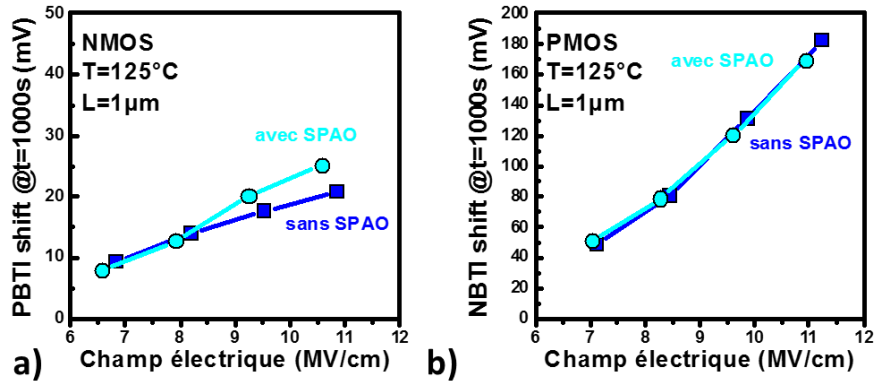


Figure 4.23 : Mesures de ΔV_T vs E_{ox} sous une contrainte électrique a) positive (nMOS) et b) négative (pMOS) pour des transistors fabriqués avec ou sans oxydation plasma SPAO du SiO₂.

En conclusion, l’oxydation plasma à très faible température (200°C) dégrade l’EOT sans pour autant améliorer la stabilité de l’oxyde piédestal, indiquant qu’il ne suffit pas d’avoir un SiO₂ épais pour obtenir une bonne qualité d’oxyde et une bonne fiabilité. Une température plus élevée (500°C voire 1200°C avec le DSA) permet effectivement de densifier l’oxyde chimique mais ne mène qu’à une légère amélioration de la fiabilité.

4.1.3 Passivation par deutérium

Un recuit aux alentours de 400°C sous une ambiance N₂/H₂ est généralement effectué afin de passiver les liaisons pendantes à l’interface SiO₂/Si et réduire la densité d’états d’interface [Yeow75]. L’utilisation de deutérium, isotope plus lourd que l’hydrogène, permet de former des liaisons plus stables et mène à une meilleure passivation des D_{it} [Cheng01]. Le deutérium étant plus lourd, un recuit à forte pression est nécessaire afin de pouvoir faire diffuser l’atome jusqu’à l’interface SiO₂/Si. Si l’utilisation du recuit deutérium a été largement démontrée, étudiée et utilisée sur des technologies avancées, nous allons ici évaluer son utilisation dans une intégration à bas budget thermique.

La passivation par N₂/H₂ a donc été remplacé par un recuit deutérium sous forte pression (HPD2 – High Pressure Deuterium) sur deux variantes d’empilement de grille, dans un premier temps avec un budget thermique très faible de 600°C 1min (Figure 4.24a).

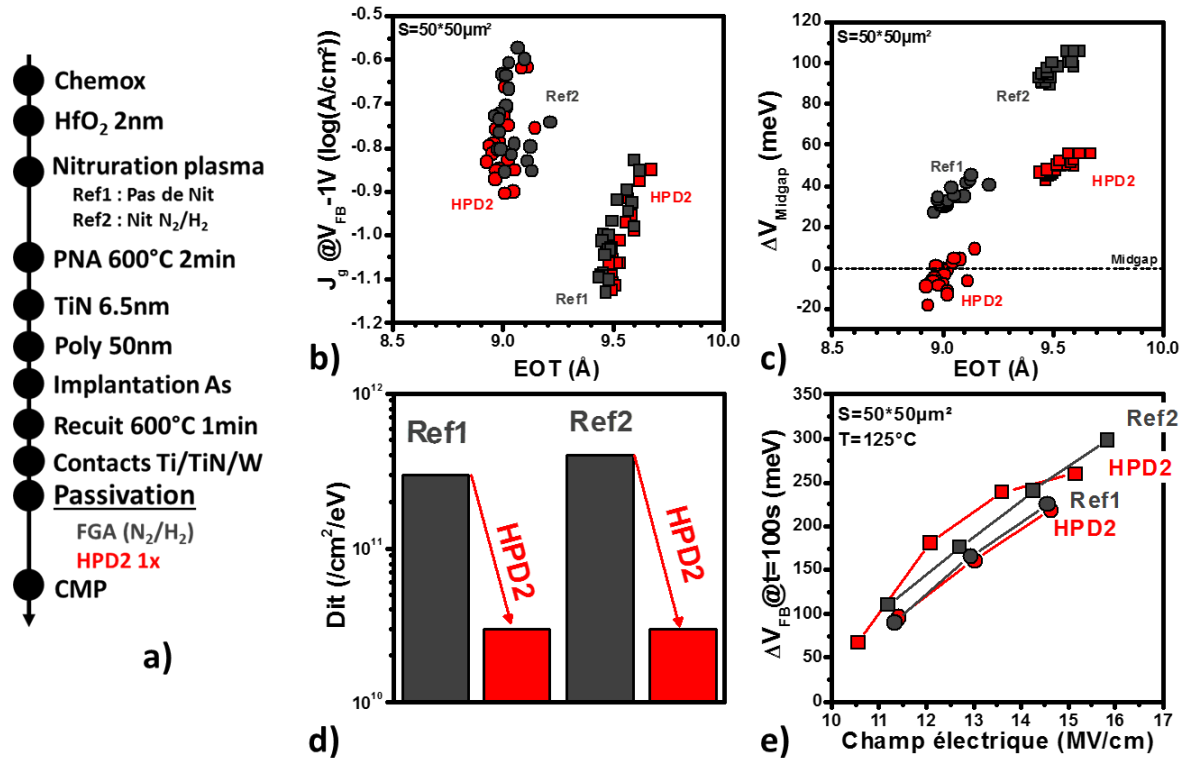


Figure 4.24 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS avec passivation par FGA ou par recuit HPD2.

La Figure 4.24d montre comme attendu une nette réduction de la densité d'états d'interface de l'ordre d'une décade quelle que soit la variante (avec ou sans nitruration du HfO₂). Cette réduction de D_{it} est accompagnée d'un décalage de la tension de bande plate d'environ 50meV vers des valeurs négatives dans les deux cas (Figure 4.24c) et qui peut soit indiquer un effet du deutérium sur le travail de sortie du TiN, soit l'apparition d'une charge nette positive dans la grille. Dans le dernier cas, il peut s'agir par exemple de la création de pièges de trous proches de la bande de valence, ou bien de la passivation de défauts accepteurs proches de la bande de conduction. En revanche, la Figure 4.24b ne montre aucune variation de l'EOT ni du courant de fuite avec le recuit HPD2. Enfin, malgré la formation de liaisons Si-D à l'interface SiO₂/Si censées être plus stables, la fiabilité des MOSCAPs n'a pas été améliorée (Figure 4.24e).

Dans un second temps, la quantité de deutérium incorporée a été augmentée en répétant l'étape de recuit HPD2 plusieurs fois, cette fois avec un budget thermique intermédiaire de 525°C 5h (Figure 4.25b).

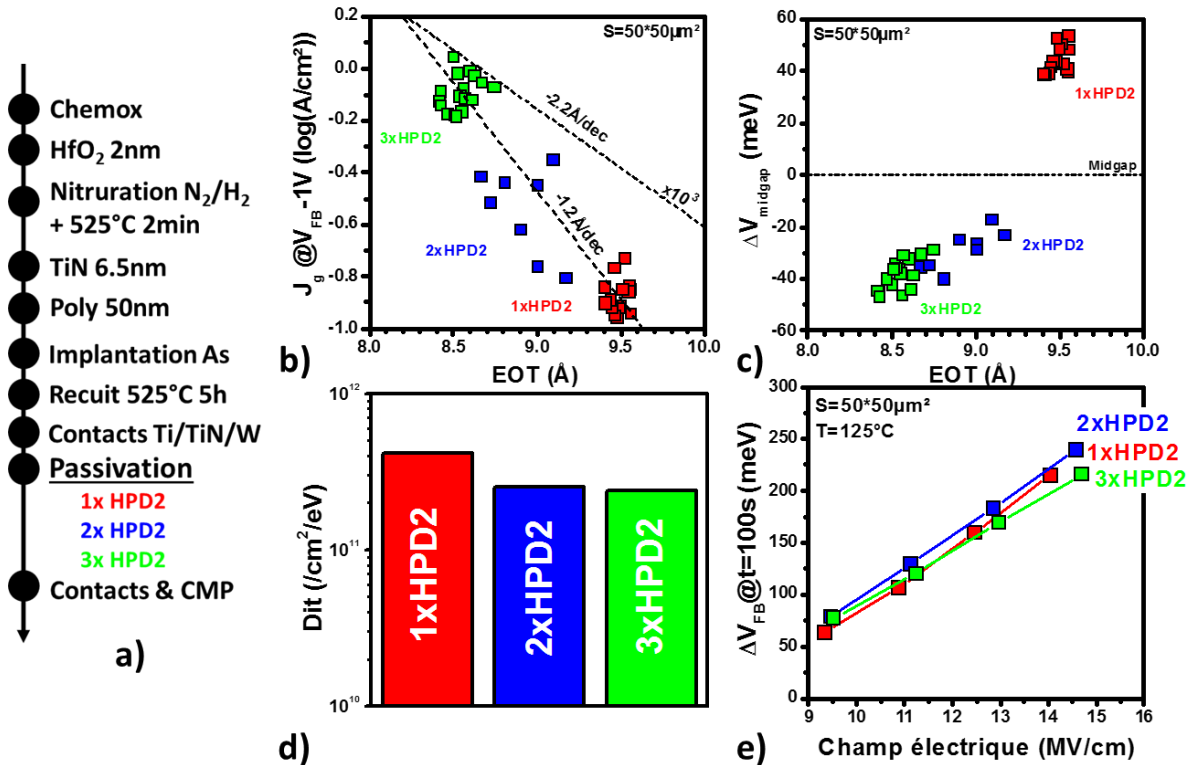


Figure 4.25 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS avec des variantes de recuit HPD2.

On remarque cette fois une diminution de l'EOT associée à une augmentation du courant de fuite selon une pente de -1.2 \AA/dec avec la répétition du nombre de recuit HPD2. Cette pente est identique à celle observée sur la Figure 4.15b où le recuit DSA à 1100°C a entraîné une diminution de l'épaisseur du HfO₂. Ceci peut alors indiquer soit un effet réducteur du deutérium sur l'épaisseur physique du high-k, soit une densification de la couche liée à l'utilisation d'une pression importante lors du recuit. On note également un décalage négatif de la tension de bande plate avec la répétition du recuit HPD2, bien que ce décalage semble saturer au bout de la troisième occurrence (Figure 4.25c). L'effet de roll-off ne devrait pas contribuer à cette variation, celui-ci étant essentiellement lié à un scaling du SiO₂ et non au HfO₂ [Charbonnier10]. Par ailleurs, la Figure 4.25d montre que la passivation de l'interface SiO₂/Si semble également stagner à partir du troisième recuit HPD2, avec de plus une valeur absolue après un recuit 525°C 5h ($D_{it} \sim 2 \times 10^{11} \text{ cm}^{-2} \cdot \text{eV}^{-1}$) supérieure au cas avec un budget thermique de 600°C 1min ($D_{it} \sim 3 \times 10^{10} \text{ cm}^{-2} \cdot \text{eV}^{-1}$). Ceci peut indiquer une plus grande difficulté à diffuser le deutérium à travers l'empilement jusqu'à l'interface SiO₂/Si lorsque certaines couches (e.g. HfO₂) sont plus denses. Une meilleure interface pourrait donc éventuellement être obtenue si le recuit forming gas ou deutérium était effectué plus tôt (ce qui n'a pas pu être fait ici pour des raisons de niveau de contamination de l'équipement) mais pose le risque d'une dépassivation lors d'un recuit subséquent. Dans tous les cas, la Figure 4.25e confirme que l'incorporation de deutérium n'améliore pas la fiabilité des capacités MOS.

Pour terminer, la passivation par deutérium a également été intégrée sur des transistors (Figure 4.26a). Les propriétés de la grille sur les transistors n'ont pas pu être extraites dans ce cas-ci. Néanmoins, la Figure 4.26d montre que si la densité d'états d'interface a probablement été réduite avec le recuit HPD2, la mobilité des porteurs à champ faible reste équivalente à un recuit N₂/H₂. Par ailleurs, le compromis Ion/Ioff (Figure 4.26b) et les résistances d'accès extraites (Figure 4.26c) sont similaires quelle que soit la passivation. A noter que les résistances d'accès sont ici très élevées

($R_{SD} \sim 600 \Omega \cdot \mu m$ contre $\sim 280 \Omega \mu m$ auparavant) probablement dû à un problème de siliciuration lors de l’intégration.

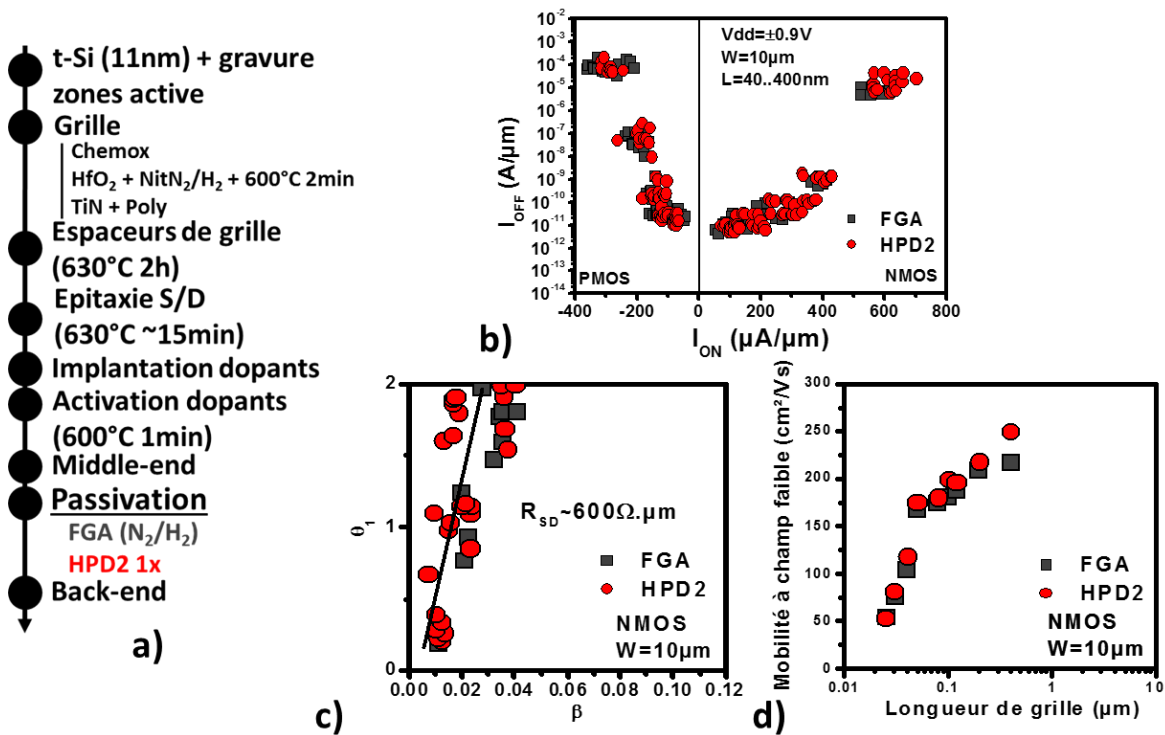


Figure 4.26 : a) Process flow, b) Compromis Ion/Ioff, c) Extraction de R_{access} et d) de mobilité à champ faible en utilisant la méthode-Y de transistors fabriqués avec passivation par FGA ou recuit HPD2.

La Figure 4.27 montre toutefois une dégradation de la fiabilité PBTI et NBTI sur les transistors. Des résultats différents sont obtenus dans la littérature, par exemple dans [Hook05], où aucune influence de la passivation par deutérium sur la fiabilité NBTI n’est trouvée. Au contraire, une amélioration de la fiabilité NBTI a été observée dans [Kimizuka00]. La dégradation de la fiabilité dans notre cas peut indiquer une création de défauts dans le volume du SiO₂ (piège de trous) dans le cas du NBTI et qui peut expliquer les variations négatives de tension de bande plates observées précédemment, ainsi que dans le volume du HfO₂ (piège d’électrons) dans le cas du PBTI. Néanmoins, ces différences de résultats obtenus par rapport à la littérature montrent que les phénomènes liés à l’hydrogène et/ou au deutérium sont encore relativement peu compris. De plus, des différences de procédés (e.g. température, pression du recuit) peuvent induire les différences observées.

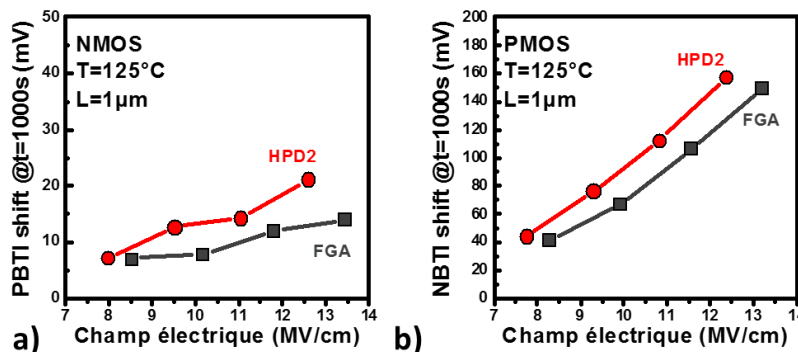


Figure 4.27 : Mesures de ΔV_T vs E_{ox} sous une contrainte électrique a) positive (nMOS) et b) négative (pMOS) pour des transistors fabriqués avec passivation par FGA ou recuit HPD2.

En conclusion, l’utilisation d’une passivation par le deutérium, plus lourd que l’hydrogène, permet de réduire la densité d’états d’interface SiO₂/Si, même si le procédé semble moins efficace dans un process flow avec un budget thermique intermédiaire (e.g. 525°C quelques heures sur capacités MOS).

En revanche, malgré la formation de liaisons Si-D censées être plus stables, une légère dégradation de la fiabilité PBTI et NBTI a été observée qui pourrait être liée à la formation de pièges de trous et d’électrons dans le SiO₂ et le HfO₂ respectivement.

4.4 Recuits alternatifs

Nous avons traité différents procédés leviers à bas budget thermique et évalué leur influence sur les propriétés de la grille. Cependant, aucun de ces procédés ne permet d’améliorer la fiabilité NBTI de manière significative. C’est pourquoi nous avons également évalué l’influence de différents recuits alternatifs.

4.1.4 Recuit micro-ondes

Comme introduit dans le Chapitre 1, le recuit micro-ondes est un recuit volumique, contrairement au recuit four qui est essentiellement surfacique. Le recuit micro-ondes permet donc de chauffer sélectivement certains matériaux selon leur conductivité et leur absorptivité à ces longueurs d’ondes. A noter tout de même que les temps de recuits de l’ordre de la minute sont largement suffisants pour diffuser l’énergie thermique et uniformiser la température. De plus, le recuit micro-ondes sur l’empilement de grille a été évalué afin de tirer profit de l’effet de polarisation et de l’échauffement local des défauts annoncés par l’équipementier. Le but ici est donc principalement de voir si ce recuit permet de diminuer la densité d’états d’interface. Pour cela, des échantillons MOSCAPs terminés comprenant les contacts W ont été envoyés chez l’équipementier DSG et recuits à différentes puissance et temps de micro-ondes (Figure 4.28a).

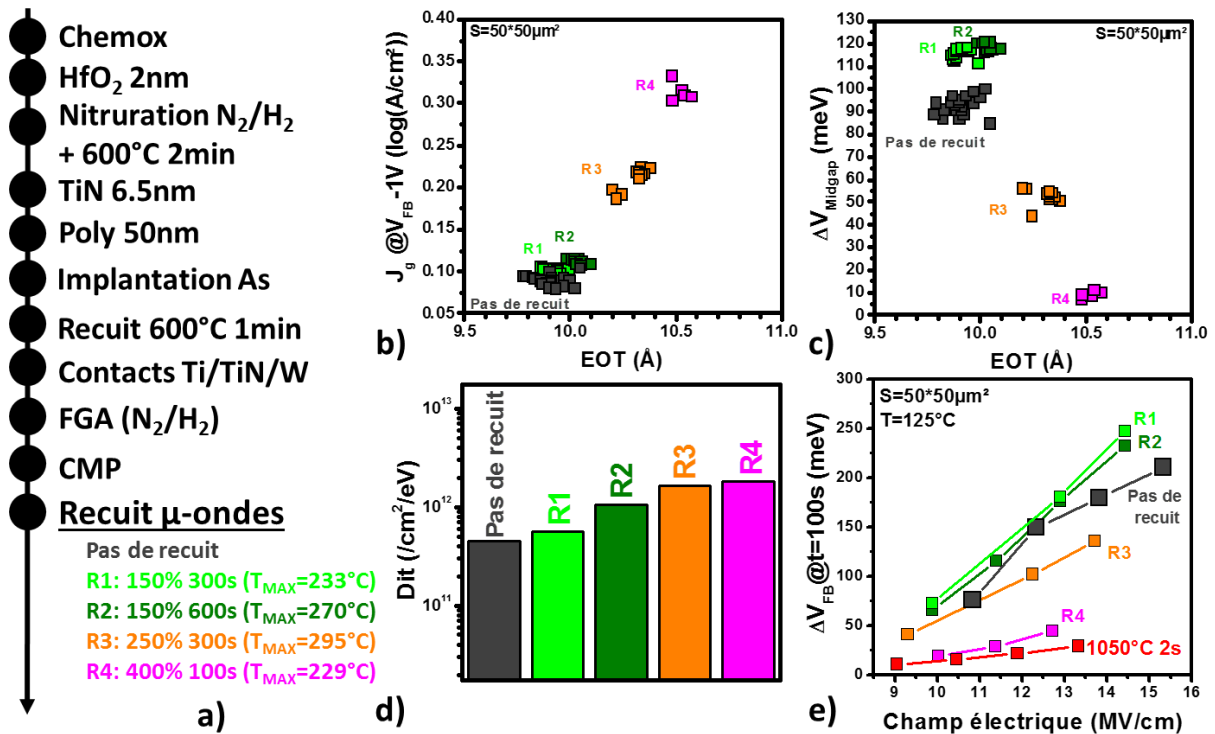


Figure 4.28 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) Dit et e) ΔV_{FB} vs E_{OX} de capacités MOS avec des variantes de recuit micro-ondes.

Les températures de la surface des plaques selon les variantes mesurées par l'équipementier sont également indiquées et sont toutes largement inférieures à 500°C. Contrairement à ce qui était attendu, la Figure 4.28d montre une dégradation de la densité d'états d'interface lorsque la puissance des micro-ondes est augmentée. Cette dégradation est accompagnée d'une augmentation de l'EOT et du courant de fuite pour les échantillons recuits (Figure 4.28b). Par ailleurs, la Figure 4.28c montre un décalage de la tension de bande plate d'environ -80meV entre la référence et l'échantillon recuit avec la puissance la plus élevée (variante R4), décalage équivalent à un recuit 1050°C 2s tel qu'observé auparavant (Figure 4.28c). Ceci peut indiquer que le TiN, métallique donc conducteur, a été chauffé à une température bien plus élevée que celle mesurée à la surface de la plaque. Cette hypothèse est étayée par la mesure de stress où le recuit avec la puissance la plus élevée mène à une excellente fiabilité de la capacité MOS et quasiment équivalente à un recuit 1050°C 2s.

L'utilisation du recuit micro-ondes n'ayant jamais été validée dans le cadre d'une intégration 3D Séquentielle, des conditions de recuits identiques ont été appliquées sur des transistors et un niveau de métal (cuivre/ultra low-k) standards représentant les couches inférieures. Les observations SEM des lignes en cuivre après un recuit correspondant à la variante R2 (où peu d'effets sont visibles sur les propriétés de la grille) montrent déjà une dégradation de la morphologie des lignes avec un effet d'agglomération du cuivre (Figure 4.29a). Cette dégradation est confirmée par les mesures électriques qui montrent une forte dégradation de la résistance des lignes et une modification de la capacité de l'ULK (Figure 4.29b). En revanche, les performances des transistors n'ont pas ou peu été impactées par le recuit micro-ondes, indiquant que les lignes métalliques au-dessus ont presque totalement absorbé et réfléchi les ondes électromagnétiques et que les siliciures des composants n'ont pas été dégradés (Figure 4.29c).

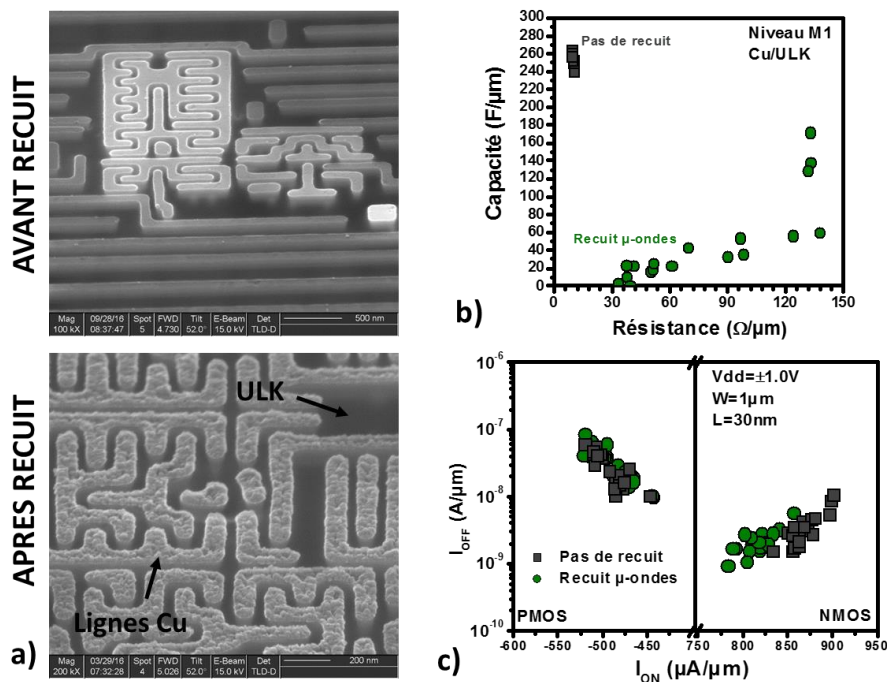


Figure 4.29 : a) Observations SEM, b) Mesures de capacité et résistance des interconnexions Cu/ULK et c) Compromis Ion/Ioff des transistors inférieurs avant et après recuit micro-ondes.

Ainsi, si des conditions de recuit micro-ondes permettant d'améliorer la fiabilité NBTI ont pu être trouvées, ces conditions ne coïncident pas avec la fenêtre de procédé des couches inférieures. En particulier, comme on pouvait s'y attendre, les couches métalliques (TiN de la grille, niveaux de métaux en cuivre) absorbent totalement les micro-ondes et chauffent à une température trop élevée. Des études sont néanmoins en cours avec des optimisations du recuit micro-onde, notamment en pulsant

les ondes électromagnétiques afin d’éviter l’effet de peau des métaux. A noter également que les contacts W sur les échantillons n’ont pas ou peu été impactés par le recuit micro-ondes et pourraient être plus stables vis-à-vis de ces longueurs d’ondes.

4.1.5 Recuit laser

Un autre recuit alternatif déjà introduit auparavant est le DSA qui permet de monter à une température de 800°C sur des temps de l’ordre de la milliseconde sans dégrader les performances des transistors inférieurs. Nous avons déjà évalué les effets de ce recuit sur l’oxyde piédestal (DSA 50% O₂ à 1200°C) ainsi que sur le high-k (DSA NH₃ à 1100°C) et vu que le budget thermique appliqué au niveau de ces couches n’améliorait pas ou peu la fiabilité. Par la suite, nous étudierons ces recuits après dépôt du métal de grille TiN ainsi que sur la couche poly-Si. Encore une fois, un recuit à une température de 1100°C sera également utilisé bien que n’étant pas acceptable dans l’intégration 3D Séquentielle. Nous évaluerons de plus les effets d’un recuit laser nanosecondes vert ($\lambda=532\text{nm}$) développé par Applied Materials. Ce laser permet alors de monter à des températures très élevées ($T>1000^\circ\text{C}$) sur un temps très court (1 pulse=60ns), limitant la diffusion thermique vers les couches inférieures sur une intégration 3D Séquentielle.

Recuit laser après dépôt TiN

Dans un premier temps, les recuits DSA (800°C et 1100°C, 0.5ms) et nanosecondes (densité d’énergie et nombre de pulses variables – A à D) ont été effectués après dépôt TiN (Figure 4.30a). Un screening des conditions de recuit laser nanosecondes a été préalablement réalisé afin de ne pas endommager l’empilement de grille.

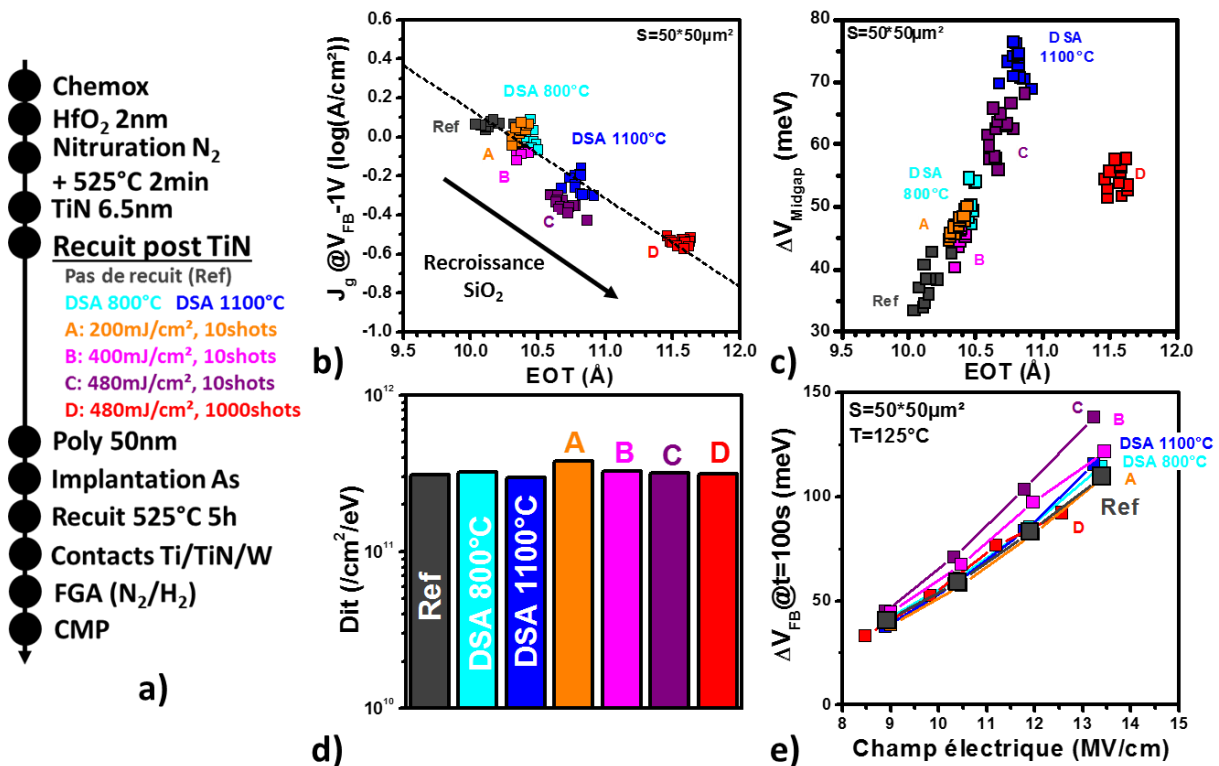


Figure 4.30 : a) Process flow, b) J_g vs EOT, c) ΔV_{Midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs Eox de capacités MOS avec des variantes de recuit laser micro ou nanosecondes après dépôt TiN.

La Figure 4.30b montre alors une recroissance du SiO₂ quel que soit le recuit laser. En particulier, l’augmentation de la température du recuit DSA ainsi que de la densité d’énergie ou du nombre de pulses pour le recuit laser nanosecondes amplifie la dégradation de l’EOT. Ceci indique notamment que la source principale d’oxygène résiduel menant à une réoxydation de l’oxyde piédestal lors du recuit après poly-Si provient du TiN. Le TiN a en effet une tendance à s’oxyder facilement à l’air et donc à absorber de l’oxygène dans son volume. C’est la raison pour laquelle le dépôt poly-Si est généralement précédé d’un nettoyage HF afin de limiter la quantité d’oxygène dans l’empilement de grille. Par ailleurs, la Figure 4.30c montre un décalage progressif du V_{FB} vers des valeurs positives lorsque le budget thermique appliqué sur le TiN non encapsulé est augmenté, contrairement à ce qui était observé dans la Section 4.2, et qui pourrait être lié à une oxydation du TiN modifiant son travail de sortie. En revanche, si le recuit sur le TiN fait croître l’oxyde piédestal, ni la densité d’états d’interface (Figure 4.30d), ni la fiabilité des MOSCAPs (Figure 4.30e) n’ont été modifiées.

Recuit laser après dépôt poly-Si

Dans un second temps, les recuits DSA et laser nanosecondes ont été effectuées après dépôt poly-Si et implantation d’arsenic (Figure 4.31a). A noter que pour toutes les variantes, un recuit 525°C 5h a tout de même été effectué après les recuits laser [Lu16c].

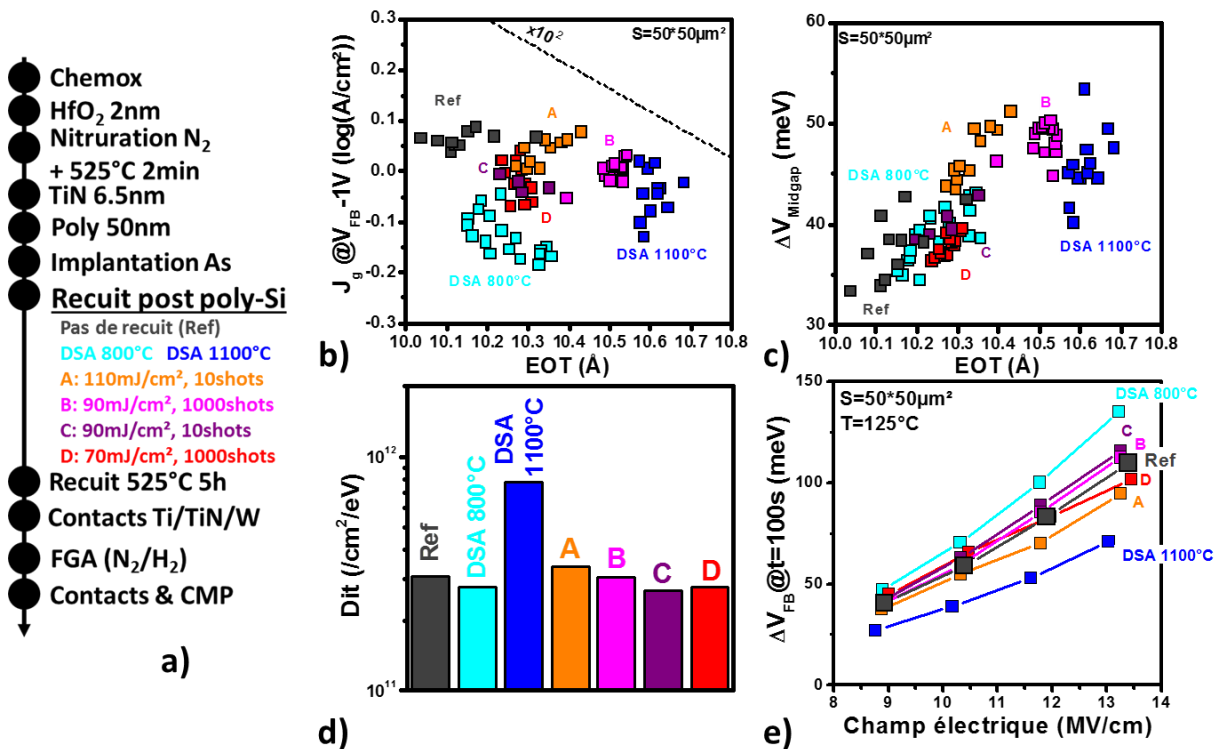


Figure 4.31 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS avec des variantes de recuit laser micro ou nanosecondes après dépôt poly-Si.

On ne remarque cette fois qu’une très légère recroissance de l’EOT (Figure 4.31b) ainsi que de faibles variations de la tension de bande plate (Figure 4.31c) avec les recuits laser sur le poly-Si. En revanche, la Figure 4.31d montre une forte dégradation de la densité d’états d’interface avec le recuit DSA 1100°C. Cette dégradation des D_{it} est néanmoins associée à une amélioration de la fiabilité des capacités MOS par rapport à la référence non recuite (Figure 4.31e).

Ainsi, ces études où des recuits laser ont été effectués après chaque couche de l’empilement de grille montrent que seul le recuit après poly-Si a amélioré la fiabilité, indiquant que ce dernier pourrait avoir une influence sur la stabilité du SiO₂ lors d’une contrainte électrique. Dans la section suivante, nous discuterons plus en détails de l’influence possible du poly-Si sur la fiabilité.

4.5 Pistes pour l’amélioration de la fiabilité NBTI

La fiabilité NBTI est essentiellement liée à la qualité du SiO₂ interfacial. Dans nos études, nous avons vu que les procédés leviers à bas budget thermique n’ont pas ou peu amélioré la fiabilité des MOSCAPs, que l’épaisseur du SiO₂ ait été réduite (nitration plasma du high-k) ou augmenté (densification de l’oxyde), ou que les D_{it} initiaux soient élevées (budget thermique 600°C 1min) ou faibles (passivation deutérium). En revanche, un recuit de 550°C 5h ou de 1100°C 0.5ms par DSA après poly-Si a mené à une nette amélioration de la stabilité du SiO₂, indiquant une possible contribution du poly-Si.

Afin d’étayer cette hypothèse, un budget thermique élevé (1000°C 15s) a été appliqué sur l’oxyde chimique et suivi d’un recuit à bas budget thermique (600°C 1min) après dépôt poly-Si (Figure 4.32a). A noter que ce recuit a été initialement mis en place comme recuit post nitration N₂/H₂ du SiO₂ (qui n’est pas étudié dans ce manuscrit).

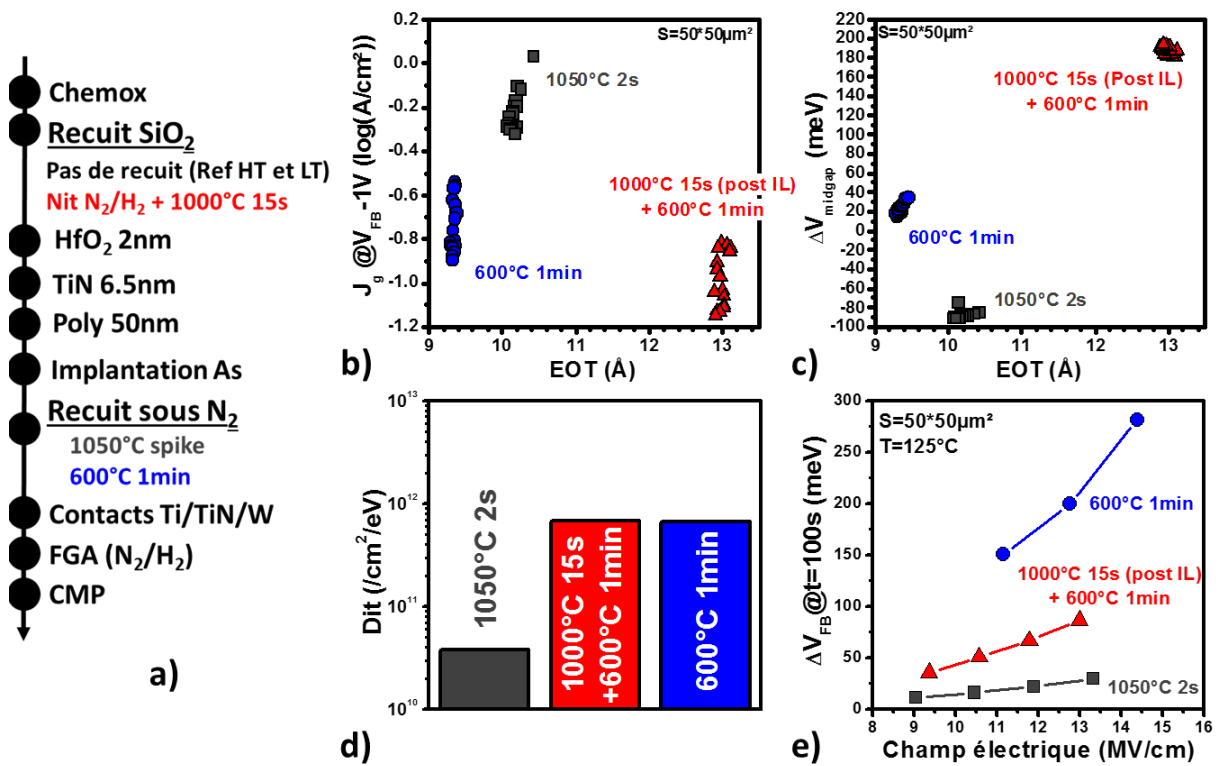


Figure 4.32 : a) Process flow, b) J_g vs EOT, c) ΔV_{midgap} vs EOT, d) D_{it} et e) ΔV_{FB} vs E_{ox} de capacités MOS comparant l’influence d’un budget thermique élevé appliqué sur l’oxyde piédestal (IL – interlayer) ou sur le poly-Si.

La Figure 4.32b montre une très forte recroissance de l’EOT dû au budget thermique élevé sur le SiO₂. En revanche, ce budget thermique n’a pas amélioré la densité d’états d’interface (Figure 4.32d). Par ailleurs, la Figure 4.32c confirme qu’aucun budget thermique élevé n’a été appliqué sur le TiN, avec une tension de bande plate proche de la bande de conduction du silicium. Enfin, les mesures de stress montrent que malgré un budget thermique plus élevé, le recuit 1000°C 15s, s’il a bien amélioré la stabilité du SiO₂, mène tout de même à une moins bonne fiabilité qu’un recuit 1050°C 2s après poly-Si. Ceci signifie alors que la contribution du poly-Si est non négligeable pour la stabilisation de l’oxyde.

Une hypothèse possible est que la fiabilité NBTI serait liée à la cristallinité du poly-Si. La Figure 4.33 montre les observations TEM de l’empilement de grille après différents recuits. On note tout d’abord que dans le cas du recuit 600°C 1min, le poly-Si est en réalité amorphe (couche qu’on continuera tout de même d’appeler poly-Si par simplification). En effet, la température de dépôt par LPCVD (525°C) ne

permet pas d’obtenir une couche polycristalline directement. Ensuite, on remarque un semblant de corrélation entre la formation de grains de poly-Si et la fiabilité NBTI. En particulier, le recuit DSA 1100°C mène à une couche mi-amorphe/mi-poly et à une fiabilité située entre les recuits les plus faibles (600°C 1min, 525°C 5h) où la couche est restée amorphe et plus élevés (1050°C 2s, 630°C 5h) où des grains sont visibles.

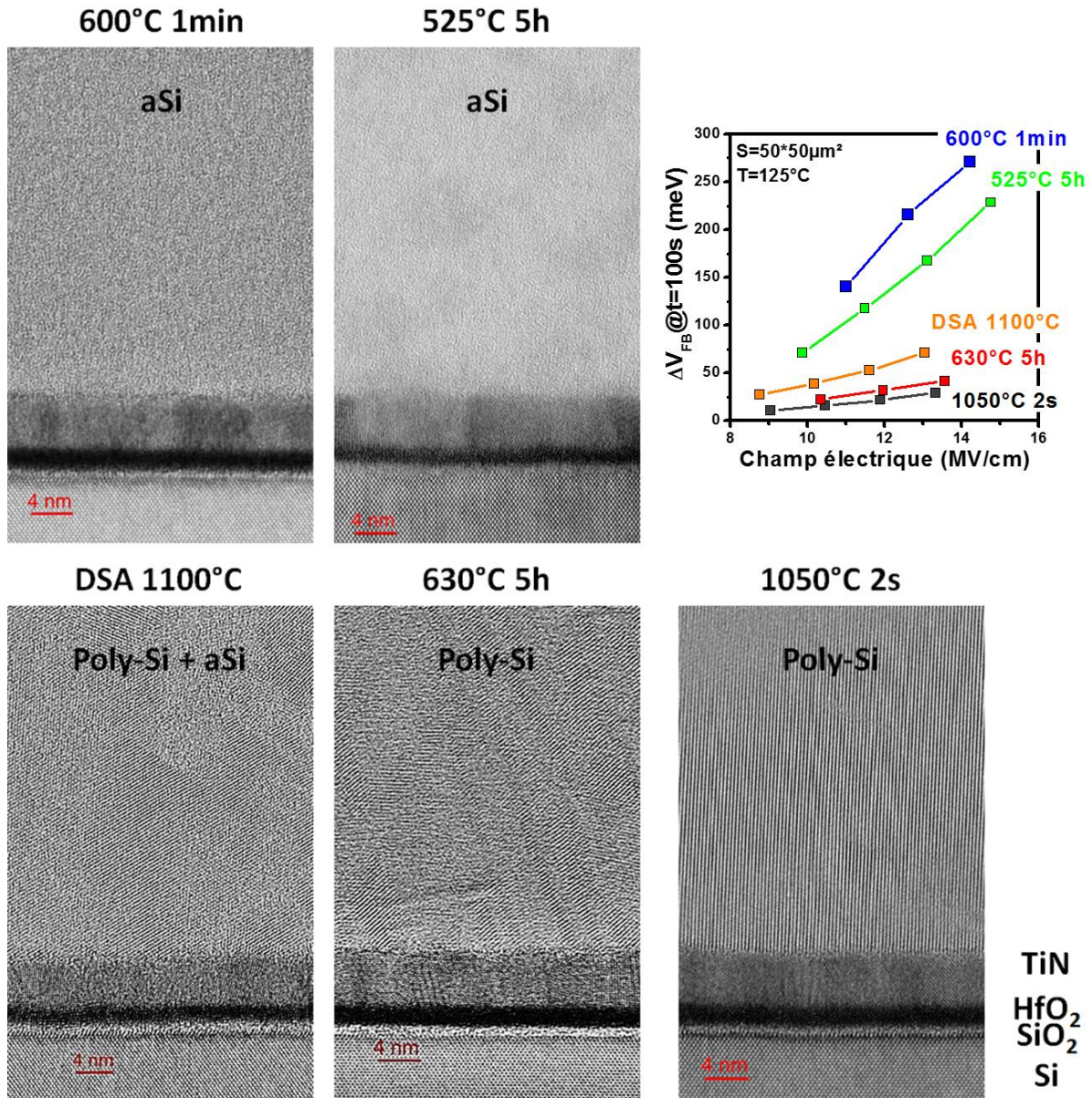


Figure 4.33 : Observations TEM d’empilements de grilles après différents budgets thermiques. Une corrélation entre la cristallinité du poly-Si et la fiabilité des capacités MOS est observée.

Deux hypothèses peuvent alors être émises pour expliquer l’amélioration de la fiabilité avec la cristallisation du poly-Si.

Une première est liée à la distribution d’oxygène dans l’empilement de grille. La présence d’une couche polycristallisée au-dessus du TiN peut en effet former une meilleure barrière à la diffusion de l’oxygène résiduel hors de la grille contrairement à une couche amorphe. L’oxygène qui redescend dans l’empilement peut alors soit recroûtre l’oxyde chimique, soit éventuellement favoriser la formation d’un silicate avec l’oxyde d’hafnium et des atomes de silicium provenant du canal. La formation du

silicate pourrait alors avoir un effet sur l’état de liaisons à l’interface entre le SiO_2 et le Si et améliorer leur stabilité.

Une autre hypothèse est liée à la distribution de l’hydrogène dans l’empilement de grille. L’influence de l’hydrogène sur la fiabilité NBTI peut en effet être anticipée puisqu’à *minima*, c’est la cassure de liaisons Si-H qui mène à la formation d’états d’interface qui dégradent la tension de seuil. Le modèle RD (Réaction-Diffusion) est généralement utilisé afin d’expliquer la formation des D_{it} lors d’une contrainte électrique négative. Initialement proposé par [Jeppsen77], puis constamment optimisé [Mahapatra13], le modèle RD repose principalement sur deux phénomènes (Figure 4.34a) :

- La cassure des liaisons Si-H à l’interface entre le SiO_2 et le Si assistée par les trous, le champ électrique et la température qui affaiblissent l’énergie de liaison.
- La diffusion des espèces hydrogènes (H) vers le poly-Si et leur recombinaison pour former du H_2 .

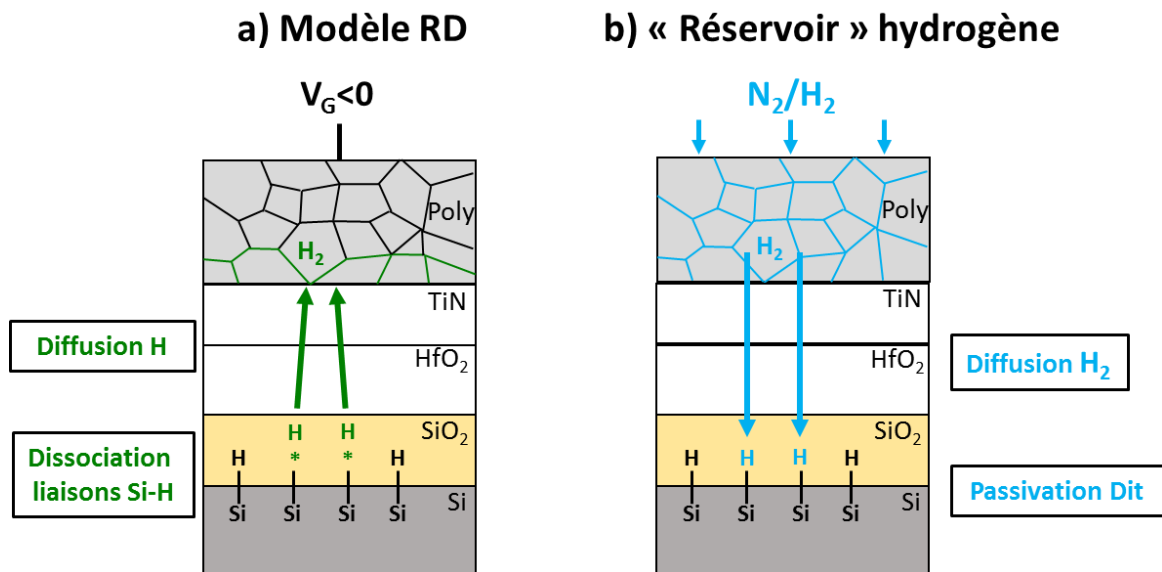


Figure 4.34 : Schémas simplifiés a) du modèle de réaction-diffusion proposé dans [Alam05] et b) du possible phénomène de « réservoir » d’hydrogène permettant une repassivation des liaisons pendantes.

Ce modèle a été initialement proposé pour décrire les phénomènes d’un empilement $\text{SiO}_2/\text{poly-Si}$. L’hydrogène étant un atome petit, sa diffusion à travers les couches HfO_2 et TiN ne devrait cependant pas être entravée. A noter également que ce modèle reste incomplet car il n’explique pas la présence d’une composante recouvrable observée lors de mesures rapides. Néanmoins, les études sur MOSCAPs ne permettent pas d’observer cette composante et le modèle peut donc bien être appliqué ici. Dans la gamme de conditions de mesures utilisée, il a été montré que le phénomène limitant de la formation de D_{it} est la diffusion de l’hydrogène loin de l’interface SiO_2/Si plutôt que la cassure des liaisons Si-H [Alam05].

Une possible raison de la corrélation entre cristallinité du poly-Si et la fiabilité est alors que l’hydrogène peut s’accumuler au niveau des grains de poly-Si, tandis qu’une couche totalement amorphe permettrait d’exodiffuser l’hydrogène hors de l’empilement de grille (Figure 4.34b). Lorsque la concentration d’hydrogène sature aux joints de grains, la diffusion de l’hydrogène ralentit. La limitation du phénomène de diffusion peut alors limiter la formation des D_{it} et donc la dégradation de la fiabilité. Ceci peut notamment expliquer que si la densification du SiO_2 permet de former des liaisons Si-H plus stables et d’améliorer la fiabilité, le recuit sur le poly-Si prédomine dans la dégradation NBTI. En outre, l’incorporation d’hydrogène durant le recuit FGA pourrait également commencer à saturer les grains de poly-Si. Cependant, cela n’expliquerait pas la fiabilité équivalente sur capacités MOS obtenue avec un recuit HPD2. Afin de compléter le modèle RD, [Alam05] propose également que lorsque la contrainte négative est retirée, les molécules H_2 formées dans le poly-Si peuvent redescendre vers

l’interface SiO_2/Si , se décomposer et repasser des liaisons pendantes précédemment créées. La présence d’un « réservoir » d’hydrogène au niveau du poly-Si peut alors amplifier cet effet et mener à une dégradation de la fiabilité « apparente » plus faible.

Par ailleurs, l’utilisation d’un traitement réducteur sur le TiN avant le dépôt du poly-Si permet également d’améliorer la fiabilité. Ce traitement consiste à effectuer une rampe en température jusqu’à 525°C sous une ambiance N_2/H_2 afin de retirer l’oxyde natif du TiN suivi du dépôt poly-Si sans remise à l’air. L’idée première de ce traitement est de limiter la concentration d’oxygène résiduel dans l’empilement de grille et donc d’empêcher une recroissance d’EOT. Toutefois, la Figure 4.35 montre que ce traitement permet également d’améliorer sensiblement la fiabilité des MOSCAPs quel que soit le recuit effectué après implantation As (600°C 1min, 525°C ou 630°C 5h). Une hypothèse est que le TiN pourrait avoir un effet inducteur sur la cristallisation du poly-Si, effet appelé *Metal-Induced Crystallization*. Ainsi, la réduction du TiO_2 interfacial par le traitement N_2/H_2 peut permettre de mettre en contact direct le métal et le poly-Si, diminuant la température nécessaire pour le cristalliser. Néanmoins, les observations TEM ne montrent pas de présence de grains de poly-Si après recuit.

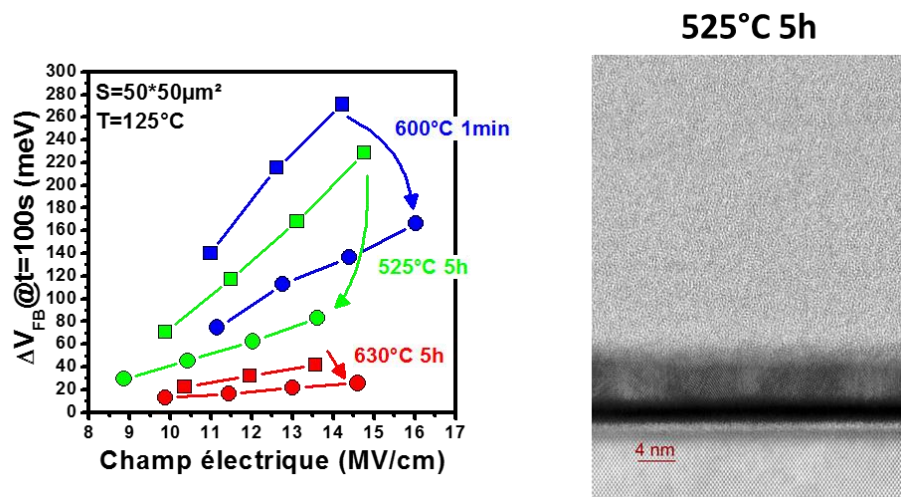


Figure 4.35 : Mesure de fiabilité sur capacités MOS avec ou sans traitement N_2/H_2 avant dépôt poly-Si pour différents budgets thermiques et observation TEM après traitement N_2/H_2 et recuit 525°C 5h.

Ces résultats peuvent soit signifier la présence de grains de poly-Si non détectables par TEM, soit que l’amélioration de la fiabilité est ici liée au procédé N_2/H_2 permettant d’incorporer de l’hydrogène plus tôt dans l’empilement de grille et d’accélérer sa saturation lors des mesures NBTI. Des études sont donc en cours pour effectuer le traitement N_2/H_2 plus tôt, par exemple au niveau du HfO_2 .

Dans le cas où une cristallisation de la couche amorphe serait nécessaire pour améliorer la fiabilité, le recuit laser Excimer ($\lambda=308\text{nm}$) nanosecondes apparaît alors comme une solution prometteuse et a déjà été démontrée dans [Ouerghi16] pour la recristallisation d’une couche amorphe plus épaisse. La forte absorptivité du silicium amorphe comparé au silicium (poly)cristallin à cette longueur d’onde (comparé au laser vert) permet en effet d’atteindre plus facilement la température de fusion du silicium amorphe. Une fois fondue, la couche recristallise alors sous forme poly-Si lorsqu’elle refroidit. Néanmoins, le recuit laser étant fortement dépendant de la structure utilisée, les capacités MOS ne permettent pas d’évaluer l’effet de ce recuit. L’énergie du laser est en effet essentiellement absorbée par les couches présentes sur les « plots » d’oxyde, comme confirmé par des simulations thermiques et les observations TEM (Figure 4.36). Des études sont donc également en cours pour évaluer le recuit laser Excimer sur des transistors.

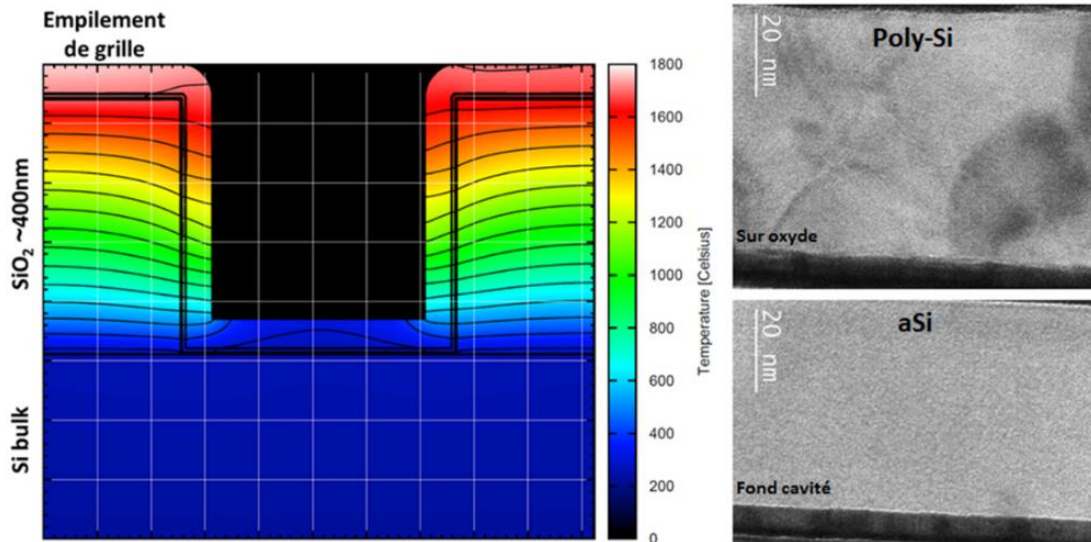


Figure 4.36 : Simulations thermiques de la température atteinte lors d'un recuit laser Excimer nanosecondes avec une densité d'énergie de $0.3\text{J}/\text{cm}^2$ d'une structure MOSCAP et observations TEM des différentes zones (sur oxyde ou au fond de la cavité). Ces simulations ont été obtenues grâce à un outil décrit dans [Mathieu15].

Conclusions du Chapitre 4

L’étude de l’empilement de grille est très complexe car elle met en jeu plusieurs phénomènes : formation d’un silicate entre le high-k et l’oxyde piédestal, roll-off de la tension de seuil ou de bande plate, recroissance d’EOT dû à l’oxygène résiduel, modification du travail de sortie du métal etc.

Nous avons tout d’abord vu que **la diminution du budget thermique des étapes subséquentes à la grille dégrade la qualité et la stabilité de l’oxyde chimique**, menant à une densité d’états d’interface, une mobilité et surtout une fiabilité NBTI dégradée. Néanmoins, **un budget thermique plus faible permet également d’obtenir une EOT plus faible ainsi qu’une réduction du courant de fuite de grille**. De plus, la réduction du budget thermique augmente le travail de sortie du TiN, facilitant l’obtention d’une tension de seuil faible sur les transistors pMOS. Par ailleurs, nous avons étudié différents procédés leviers à bas budget thermique qui peuvent être utilisés selon l’application visée (Figure 4.37). La nitruration de l’HfO₂ permet de réduire la recroissance de l’EOT lors d’un recuit. Des optimisations sont néanmoins nécessaires afin de contrôler la distribution de l’azote et d’éviter une dégradation de la mobilité par la formation de liaisons Si-N près de l’interface SiO₂/Si. Les procédés d’oxydation plasma à 500°C (DPO et RPO) permettent en outre d’obtenir un SiO₂ densifié permettant de légèrement améliorer la fiabilité au prix d’une recroissance de l’EOT. Enfin, la passivation par deutérium en lieu et place du recuit forming gas permet de tirer profit de la meilleure stabilité des liaisons Si-D et de réduire la densité d’état d’interface.

En revanche, si la fiabilité des transistors nMOS (PBTI) semble poser peu de problèmes, **la fiabilité NBTI des transistors pMOS reste clairement la principale difficulté dans une intégration 3D Séquentielle**. Un budget thermique seuil entre 525°C et 550°C a toutefois été identifié permettant d’améliorer la fiabilité. De plus, les études de recuit laser sur les couches de l’empilement de grille indique que **la contribution du poly-Si semble prédominer sur la dégradation NBTI dans une intégration froide**. Enfin, la désoxydation du TiN par un traitement N₂/H₂ avant le dépôt poly-Si permet également de sensiblement améliorer la fiabilité. Ces résultats indiquent une possible corrélation entre la formation de grains de poly-Si et la fiabilité NBTI, auquel cas l’utilisation d’un recuit laser Excimer nanosecondes serait très prometteur. Une étude fine de la distribution de l’hydrogène dans l’empilement de grille selon les variantes reste néanmoins nécessaire afin d’évaluer la contribution de l’hydrogène sur la fiabilité.

	Réduction de l’EOT	Réduction du courant de fuite	V _T faible sur pMOS	Interface SiO ₂ /Si (D _{it} , mobilité)	Fiabilité NBTI
Réduction BT	++	++	++	--	--
Nitruration high-k	++	=	=	--	-
Densification chemox	--	++	+	=	+
Passivation deutérium	=	=	-	++	-

Figure 4.37 : Récapitulatifs des différents leviers pour l’amélioration des propriétés de l’empilement de grille.

Chapitre 5 – Conclusions du manuscrit

Dans le Chapitre 1, nous avons vu que la miniaturisation des transistors suivant la Loi de Moore devient de plus en plus difficile, avec des solutions technologiques existantes (e.g. introduction de nouveaux matériaux, nouvelles architectures) mais de plus en plus complexes (e.g. photolithographie) et aussi et surtout plus coûteuses. En cela, la 3D Séquentielle – et en particulier l'intégration CoolCube™ dans notre cas – apparaît comme une alternative prometteuse afin d'augmenter la densité de composants et le gain en puissance sans se résoudre à développer un nouveau nœud technologique. En revanche, deux verrous technologiques doivent alors être levés : l'obtention d'une couche de silicium cristallin de bonne qualité pour la fabrication des composants du haut et la préservation des transistors et des interconnexions du niveau bas. **Dans ce manuscrit, nous nous sommes essentiellement intéressés à la diminution du budget thermique des transistors du haut afin de permettre une intégration sans dégrader les couches inférieures.** A partir du process flow de départ, quatre briques technologiques critiques ont été identifiées : l'activation des dopants, les espaceurs de grille, l'épitaxie de surélévation des sources et drains et l'empilement de grille. Le but était alors de diminuer le budget thermique de ces briques en-dessous de 500°C pour des procédés de l'ordre de l'heure, 800°C pour quelques millisecondes voire des températures plus élevées pour des temps de l'ordre de la nanoseconde, le tout sans dégrader les performances du transistor fabriqué à bas budget thermique. L'activation des dopants ayant été traitée dans un autre travail et des solutions trouvées (e.g. SPER à 600°C 1-2min, recuit laser nanosecondes, intégration « extension-first »), nous nous sommes focalisés dans notre étude sur les trois autres briques technologiques.

Dans le Chapitre 2, nous avons étudié et comparé différents diélectriques déposés à basse température comme alternative au nitrure de silicium ($\text{SiN}_{630^\circ\text{C}}$) pour les espaceurs de grille. Cependant, la réduction de la température de dépôt mène à une dégradation des propriétés isolantes des matériaux, à une moins bonne stabilité thermique ainsi qu'à une plus grande sensibilité aux chimies humides et sèches menant à l'apparition de trous dans les espaceurs et à la formation d'un dépôt parasite sur la grille lors de l'épitaxie des sources et drains. Néanmoins, la présence d'un plasma lors du dépôt permet d'améliorer les propriétés des diélectriques grâce à l'apport d'une énergie supplémentaire à l'énergie thermique. Par ailleurs, l'introduction de matériaux low-k permet de diminuer les capacités parasites liées aux espaceurs de grille et ainsi de réduire les délais de propagation des signaux dans les circuits. **Ainsi, le SiCO ($k=4.5$) déposé par PECVD à 400°C et le SiOCN ($k=4.7$) déposé à 480°C par PEALD ont été intégrés sur des transistors, menant à des performances statiques équivalentes et un gain sur les délais de propagation des signaux démontré dans le cas du $\text{SiCO}_{400^\circ\text{C}}$.**

Dans le Chapitre 3, nous avons vu que la problématique d'une surélévation des sources et drains à basse température peut être scindée en deux : la préparation de surface et la croissance épitaxiale. La réduction de la température du recuit sous H_2 mène en effet à l'apparition de défauts dans la couche épitaxiale subséquente due à la présence de

contaminants à la surface du silicium initiale. Néanmoins, combinée aux nettoyages humides standards, l'utilisation d'un nettoyage Siconi permet de retirer efficacement l'oxyde natif du silicium ainsi que les autres contaminants provenant des étapes précédentes. De plus, **une Siconi *in-situ* permet d'éviter une réadsorption d'oxygène avant la croissance épitaxiale, auquel cas le budget thermique du recuit hydrogéné peut être diminué voire supprimé.** Par ailleurs, la faisabilité d'une épitaxie à la fois sélective et à basse température est principalement compliquée par la faible vitesse de croissance et/ou de gravure des chimies chlorées. Ainsi, la stratégie de dépôt/gravure en deux temps permet d'utiliser des précurseurs plus favorables pour la vitesse de dépôt (e.g. Si_2H_6) suivie d'une gravure sélective en utilisant des chimies chlorées (e.g. HCl ou Cl_2). **En utilisant la stratégie dépôt/gravure cyclée, des épitaxies à 500°C ont été développées et validées morphologiquement (Si non dopé, Si dopé phosphore) voire électriquement (SiGe dopé bore).** Des optimisations, notamment en termes de résistance d'accès des sources et drains, doivent toutefois encore être apportées.

Enfin, dans le Chapitre 4, les challenges liés à la diminution des budgets thermiques « vus » par l'empilement de grille ont été soulevés. Nous avons vu que la réduction du budget thermique permet d'améliorer l'EOT et le courant de fuite et de diminuer la tension de seuil des transistors pMOS mais dégrade fortement la qualité de l'interface entre le SiO_2 et le Si et sa stabilité vis-à-vis d'une contrainte électrique négative, menant à une mobilité et une fiabilité NBTI moins bonnes qu'à haute température. En revanche, la fiabilité PBTI des transistors nMOS ne semble pas problématique dans une intégration 3D Séquentielle. En parallèle, différents procédés leviers à bas budget thermique ont été étudiés afin de modifier les propriétés de la grille selon l'application visée : nitruration du high-k pour l'EOT, densification de l'oxyde piédestal pour améliorer la qualité du SiO_2 ou passivation par deutérium pour réduire la densité d'états d'interface. La fiabilité NBTI des pMOS reste néanmoins la difficulté principale d'une grille froide. Nous avons vu qu'un budget thermique supérieur à 550°C 5h permettait d'améliorer sensiblement la fiabilité des capacités MOS. De plus, l'utilisation d'un recuit DSA sur le poly-Si ou d'un traitement N_2/H_2 avant son dépôt permet également d'améliorer la stabilité de l'oxyde piédestal, indiquant une contribution non négligeable du poly-Si sur la fiabilité NBTI. Dans le cas où une (re) cristallisation du poly-Si serait nécessaire, **l'utilisation d'un recuit laser nanosecondes peut être prometteuse pour améliorer la fiabilité tout en limitant l'impact thermique sur les niveaux inférieurs de l'intégration 3D Séquentielle.**

Toutes ces études nous mènent donc à un process flow idéal d'un « MOS froid » donné dans la Figure 5.1. En particulier, le recuit laser nanosecondes afin de recristalliser la couche de poly-Si peut être faite avant la gravure de l'empilement de grille afin de minimiser l'impact de la topographie sur le faisceau.

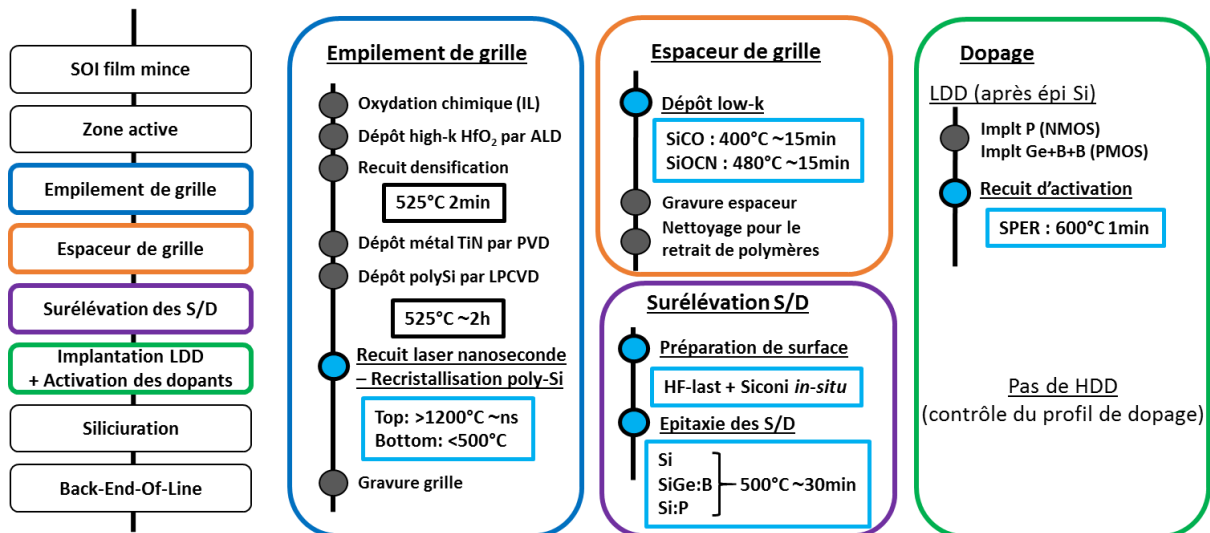


Figure 5.1 : Process flow idéal d'un transistor « MOS froid ». Les étapes dont le budget thermique a été réduit sont en bleu.

Ce process flow permettrait alors la fabrication de transistors performants et fiables tout en étant compatible avec la stabilité des transistors et des interconnexions intermédiaires des niveaux bas [Lu17a]. Ceci ouvre la voie à une large gamme de configurations possibles, par exemple cMOS sur cMOS, unipolaire (i.e. nMOS ou pMOS) sur cMOS, unipolaire sur unipolaire, etc. L'intégration cMOS sur cMOS est idéale car elle permet d'empiler n'importe quel circuit (e.g. mémoire, porte logique) l'un sur l'autre permettant donc *en théorie* de doubler la densité de composants par unité de surface. Néanmoins, une difficulté est d'interconnecter la grande densité de composants sur deux niveaux, nécessitant alors des outils de design dédiés (PDK – *Process Design Kit*) permettant de placer les lignes métalliques tout en respectant les règles de dessin (e.g. espacement entre les grilles, entre les contacts, etc.) du nœud technologique considéré. Ces outils n'étant pas encore développés pour une intégration 3D cMOS sur cMOS, cette option est pour le moment difficile à évaluer. En revanche, il est possible d'adapter les outils de design planaires existants pour rajouter un niveau de transistors unipolaire, permettant ainsi d'évaluer les éventuels gains à utiliser l'intégration 3D Séquentielle.

Pour donner un exemple d'application, des évaluations sur des cellules SRAMs ont été effectuées afin de réduire leur surface. Les SRAMs peuvent en effet représenter parfois plus de 50% de la surface de certaines puces telles que les microprocesseurs, justifiant le besoin de réduire leur surface et d'augmenter leur densité. Des travaux ont été menés sur des cellules SRAMs 4T (4 transistors – 2 nMOS et 2 pMOS), menant à des gains de 30% sur la surface de la cellule unitaire en empilant les deux transistors nMOS sur les deux transistors pMOS [Brocard16]. Néanmoins, les cellules 4T sont relativement peu utilisées dues à leur consommation importante, contrairement aux cellules 6T (6 transistors) plus usitées. On peut alors distinguer les cellules 6T 4P2N (4 pMOS et 2nMOS) ou 4N2P (4 nMOS et 2pMOS), ces dernières étant plus communes car les transistors nMOS sont généralement plus performants. Le circuit électrique d'une cellule SRAM 6T 4N2P est donné dans la Figure 5.2.

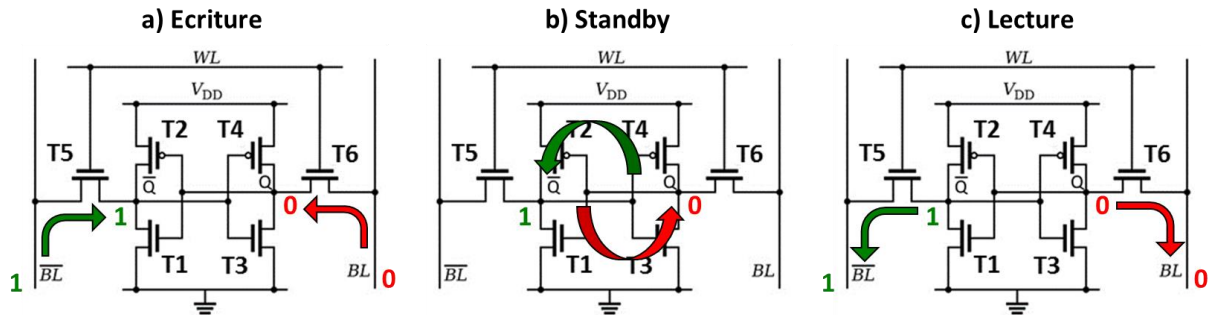


Figure 5.2 : Circuit électrique et principe de fonctionnement d'une SRAM : a) Ecriture d'un état 0 dans la SRAM (WL=1), b) SRAM en standby (WL=0) et c) Lecture de la SRAM (WL=1).

Cette cellule est essentiellement constituée de trois paires de transistors :

- T5 et T6 appelées *pass-gate*
- T1 et T3 appelées *pull-down*
- T2 et T4 appelées *pull-up*

La cellule est donc composée de deux inverseurs (T1/T2 et T3/T4) qui vont permettre de stocker l'information ainsi que deux transistors de contrôle (T5 et T6) qui servent à écrire ou à lire l'information. Pour faire simple, lorsque le signal WL (Word Line) est à un niveau haut (e.g. à V_{DD}), les transistors T5 et T6 deviennent passants et transfèrent le niveau de signal BL (Bit Line) et /BL sur les nœuds Q et /Q respectivement (Figure 5.2a). Par convention, le nœud Q correspond à l'information stockée dans la mémoire SRAM. Ainsi, si on désire stocker un état 0, le signal BL doit être bas tandis que le signal /BL sera haut. Lorsque le signal WL est à un niveau bas (e.g. à 0V), les transistors T5 et T6 sont bloquants et les deux inverseurs T1/T2 et T3/T4 permettent de conserver l'information stockée sur le nœud Q tant qu'ils sont alimentés (Figure 5.2b). Afin de lire l'information stockée, le signal WL est de nouveau mis à un niveau haut, permettant de transférer l'information stockée vers les signaux BL et /BL (Figure 5.2c). Généralement, plusieurs SRAMs sont alors généralement regroupées dans une matrice afin de traiter des informations plus importantes, chaque cellule représentant un bit.

En termes de design, les six transistors peuvent être agencés de différentes manières et ont été regroupés dans 4 différentes catégories dans [Ishida98], voire dans 5 dans [Balobas15]. Ces catégories ont notamment été construites à partir de la position des deux inverseurs. Ces deux inverseurs comportent chacun une grille partagée représentée par les « barreaux » de poly-Si qui doivent être séparées d'une distance minimale appelée CPP (*Contacted Poly Pitch*) dépendant du nœud technologique utilisé. Une fois la configuration des deux inverseurs choisis, il y a alors un nombre limité de façons de rajouter les autres composants de la SRAM (pass-gate, word line, bit lines). La Figure 5.3 résume les différentes configurations possibles ainsi que la surface occupée par les cellules SRAMs en considérant une matrice de 16 bits et des règles de dessin au nœud 32nm. La configuration 4, où les 2 pMOS sont au milieu des 4 nMOS, est donc la plus optimisée en terme de surface et celle sur laquelle les évaluations 3D ont été effectuées.

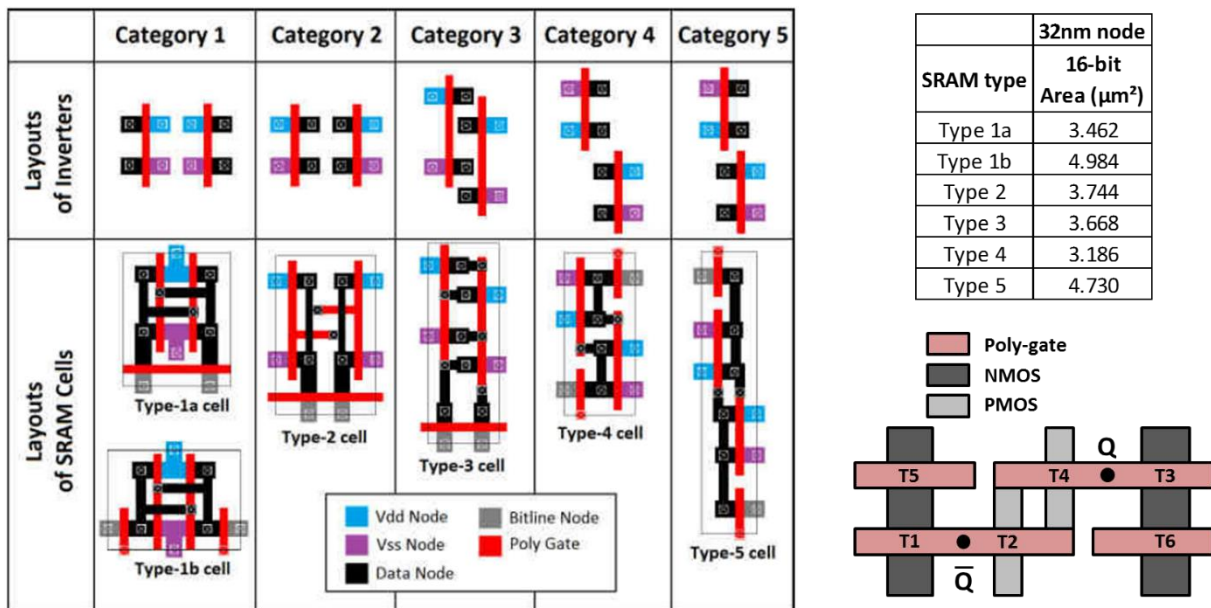


Figure 5.3 : Différentes catégories d'arrangement d'une cellule SRAM et surface associée d'une matrice de 4x4 SRAMs (16bits) en considérant les règles de dessin 32nm [Balobas15]. Le lay-out simplifié de la SRAM considérée dans notre étude (Catégorie 4) est également représenté.

A partir de ce circuit, plusieurs configurations 3D sont envisageables. Tout d'abord, une intégration unipolaire sur unipolaire (e.g. 4 nMOS en haut et 2 pMOS en bas) a été proposée dans [Liu12], permettant un gain de 33% sur la surface de la cellule. De plus, une intégration unipolaire est plus simple et moins coûteuse qu'une intégration CMOS, l'implantation ionique des dopants ne nécessitant pas d'étape de photolithographie. Néanmoins, cette configuration nécessite l'utilisation de vias (*Vertical Interconnect Accesses*) dits internes connectant directement les zones actives des deux niveaux (Figure 5.4a). Ces vias internes sont technologiquement difficilement réalisables et il est en réalité nécessaire de déporter les interconnexions non seulement entre les zones actives (Figure 5.4b), mais également au niveau des grilles, occasionnant une perte de surface et donc de l'utilité de l'empilement 3D.

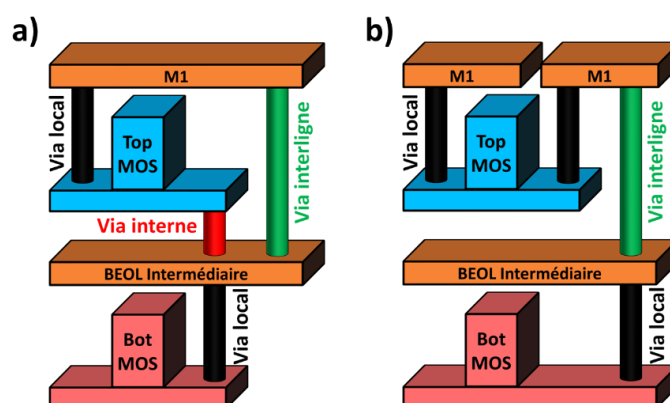


Figure 5.4 : a) Schéma de l'empilement 3D proposé par [Liu12] intégrant des vias interne entre les zones actives des deux niveaux (non réalisable technologiquement actuellement) et b) Schéma 3D avec déport des vias.

Une autre configuration possible est alors une intégration unipolaire sur CMOS [Lu17b]. Plus particulièrement, une intégration de transistors nMOS sur le niveau supérieur permet de profiter de la meilleure fiabilité PBTI à bas budget thermique comparée aux transistors pMOS et leur fiabilité NBTI dégradée comme vu dans le Chapitre 4. Ces derniers peuvent alors être

réalisés à haute température au niveau de la couche inférieure. Encore une fois, il y a différentes manières de séparer la cellule SRAM. Une façon intuitive de réarranger la SRAM est de couper de manière symétrique, c'est-à-dire de déplacer sur le niveau haut soit les deux pull-down (T1 et T3), soit les deux pass-gate (T5 et T6). Néanmoins, lorsque les deux pull-down sont reportés sur le niveau du haut, aucun gain n'est obtenu sur le niveau du bas dont la surface dépend des quatre transistors restants (Figure 5.5a). De plus, dans cette configuration, 4 nœuds internes au circuit sont coupés et doivent être remplacés par des vias entre les deux niveaux. Ces quatre vias devant respecter les contraintes d'espacement lié aux règles de dessin, aucun gain de surface n'est obtenu sur le niveau supérieur non plus. De manière similaire, si le report des deux pass-gate ne nécessite que deux vias pour remplacer les nœuds internes coupés, la surface du bas reste dépendante des deux inverseurs restants (Figure 5.5b). Au final, seule une configuration asymétrique où un pass-gate et un pull-down (e.g. T5 et T1) sont reportés sur le dessus permet effectivement de réduire la surface de la cellule SRAM (Figure 5.5c).

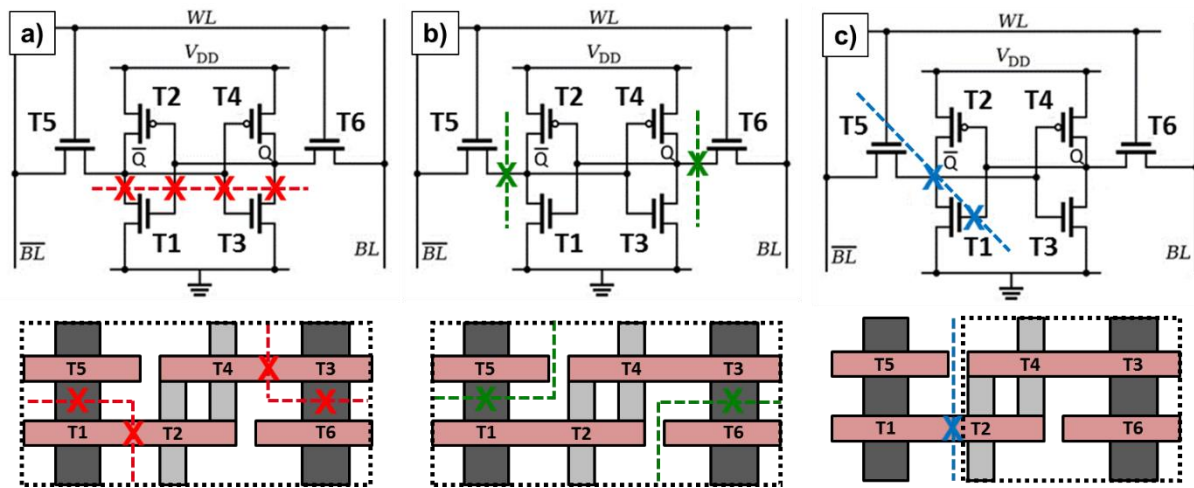


Figure 5.5 : Schémas représentant les différents découpages possibles d'une SRAM 6T. a) Déport des deux pull-down, b) des deux pass-gate et c) d'un pass-gate et d'un pull-down sur le niveau supérieur. La surface de la cellule 3D est représentée en pointillé.

La Figure 5.6 montre alors un schéma 3D de la cellule SRAM proposée incluant les lignes métalliques permettant d'acheminer les différents signaux de contrôle de la SRAM (word line et bit lines) ainsi que les alimentations (V_{DD} et Ground). En prenant en compte les règles de dessin du 14FDSOI de STMicroelectronics, cette configuration permet alors de réduire la surface de la cellule SRAM 6T de $0.078\mu\text{m}^2$ à $0.057\mu\text{m}^2$ en 3D, soit un gain de 27%. De plus, en prenant en compte que chaque cellule SRAM comporte 7 vias, cet empilement 3D permettrait alors d'atteindre $10^8\text{vias}/\text{mm}^2$, soit la densité théorique indiquée dans le Chapitre 1. Cette densité est notamment permise par l'excellent alignement des composants des deux niveaux et donc de l'espacement minimal entre les vias ($\text{pitch} < 100\text{nm}$), comparé par exemple à la technologie 3D-TSV où l'espacement est de l'ordre du micromètre. A noter toutefois que ces évaluations ne sont faites qu'en termes de gain de surface, et qu'une étude plus poussée des performances de cette SRAM 3D reste à être réalisée. En particulier, l'asymétrie de la cellule en 3D et des deux inverseurs combinés à la présence de lignes métalliques de résistance différente entre les deux niveaux (BEOL intermédiaire en tungstène) et au-dessus du niveau supérieur (Metal 1 et 2 en cuivre) peuvent dégrader les performances de la SRAM.

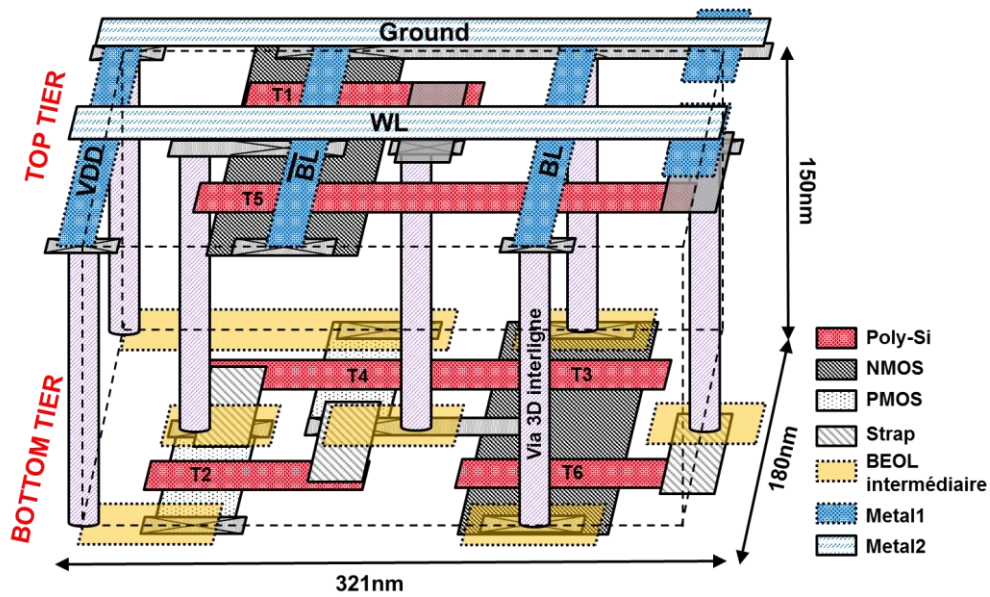


Figure 5.6 : Schéma 3D de la cellule SRAM 6T 3D proposée intégrant un niveau de BEOL intermédiaire et deux niveaux de métal. Le strap correspond à une couche conductrice au niveau des transistors.

Cet exemple montre donc les possibilités offertes par l'intégration 3D Séquentielle pour obtenir des circuits logiques denses sans réduire les dimensions des composants. Pour citer d'autres applications, l'intégration 3D Séquentielle peut également être utilisée pour des empilements hybrides, par exemple pour des applications d'imageurs où photodiode et transistors peuvent être séparés sur deux niveaux [Coudrain08], pour des capteurs NEMS (*Nano Electro-Mechanical Sensors*) très précis [Ouerghi15] voire pour le domaine du neuromorphique, où une grande capacité de calcul est nécessaire afin de simuler le fonctionnement de structures biologiques telles que le système nerveux ou le cerveau [An16]. Toutefois, il reste encore plusieurs autres questions à répondre avant d'arriver à une application tangible : routage des interconnexions 3D (i.e. nécessité d'un outil de design dédié) [Billoint16], évaluation de l'échauffement thermique (*self-heating*) sur deux voire plusieurs niveaux de composants [Brocard17] ou encore la gestion de la contamination (e.g. présence de siliciure lors de la fabrication des transistors du niveau supérieur en front-end-of-line) [Brunet16].

En conclusion, l'étude et le développement de briques technologiques « froides », s'ils ont été réalisés dans le cadre de l'intégration 3D Séquentielle de transistors sur transistors, peuvent également être utiles pour d'autres technologies pas forcément limitées en budget thermique : diélectriques low-k pour la réduction des délais, épitaxie basse température pour la photonique, alternative à l'intégration complexe du gate-last pour donner des exemples. Le développement de briques technologiques à bas budget thermique, bien que complexe car opposé à des phénomènes physiques immuables, peut donc avoir bien plus d'applications que celles considérées dans ce manuscrit.

Références

- [Acosta16] P. Acosta Alba, S. Kerdiles, B. Mathieu, M. Veillerot, R. Kachtouli, F. Mazzamuto, C. Fenouillet-Beranger, *Nanosecond laser annealing for phosphorus activation in ultra-thin implanted silicon-on-insulator substrates*, International Conference on Ion Implantation Technology, 2016.
- [Alam05] M.A. Alam, S. Mahapatra, *A comprehensive model of PMOS NBTI degradation*, Microelectronics Reliability, 45, p.71-81, 2005.
- [Alers98] G.B. Alers, D.J. Werder, Y. Chabal, H.C. Lu, E.P. Gusev, E. Garfunkel, T. Gustafsson, R.S. Urdahl, *Intermixing at the tantalum oxide/silicon interface in gate dielectric structures*, Applied Physics Letters, Vol73 Issue11, p.1517-1519, 1998.
- [An16] H. An, M. Amimul Ehsan, Z. Zhou, Y. Yi, *Electrical modeling and analysis of 3D Neuromorphic IC with Monolithic Inter-Tier Vias*, IEEE Conference on Electrical Performance Of Electronic Packaging and Systems, 2016.
- [Ando12] T. Ando, *Ultimate Scaling of High-k Gate Dielectrics: Higher-k or Interfacial Layer Scavenging?*, Materials, 5, p.478-500, 2012.
- [Ang06] D.S. Ang, S. Wang, *Recovery of the NBTI-Stressed Ultrathin Gate p-MOSFET: The Role of Deep-Level Hole Traps*, IEEE Electron Device Letters, Vol27 No11, p.914-916, 2006.
- [Ardigo14] M.R. Ardigo, M. Ahmed, A. Besnard, *Stoney formula: Investigation of curvature measurements by optical profilometer*, Advanced in Material research, 996 (IX), p361-366, 2014.
- [Bae16] D.I. Bae, G. Bae, K.K. Bhuwalka, S.H. Lee, M.G. Song, T.s. Jeon, C. Kim, W. Kim, J. Park, S. Kim, U. Kwon, J. Jeon, K.J. Nam, S. Lee, S. Lian, K.i. Seo, S.G. Lee, J.H. Park, Y.C. Heo, M.S. Rodder, J.A. Kittl, Y. Kim, K. Hwang, D.W. Kim, M.s. Liang, E.S. Jung, *A novel tensile Si (n) and compressive SiGe (p) dual-channel CMOS FinFET co-integration scheme for 5nm logic applications and beyond*, IEEE International Electron Device Meeting, 683, 2016.
- [Balobas15] D. Balobas, N. Konofaos, *Design and evaluation of 6T SRAM layout designs at modern nanoscale CMOS processes*, International Conference on Modern Circuits and System Technologies, 2015.
- [Batude15] P. Batude, C. Fenouillet-Beranger, L. Pasini, V. Lu, F. Deprat, L. Brunet, B. Sklenard, F. Piegas-Luce, M. Cassé, B. Mathieu, O. Billoint, G. Cibrario, O. Turkyilmaz, H. Sarhan, S. Thuries, L. Hutin, S. Sollier, J. Widiez, L. Hortemel, C. Tabone, M-P. Samson, B. Previtali, N. Rambal, F. Ponthenier, J. Mazurier, R. Beneyton, M. Bidaud, E. Josse, E. Petitprez, O. Rozeau, M. Rivoire, C. Euvard-Colnat, A. Seignard, F. Fournel, L. Benaissa, P. Coudrain, P. Leduc, J-M. Hartmann, P. Besson, S. Kerdiles, C. Bout, F. Nemouchi, A. Royer, C. Agrafeuil, G. Ghibaud, T. Signamarcheix, M. Haond, F. Clermidy, O. Faynot, M. Vinet, *3DVLSI with CoolCube process: An alternative path to scaling*, IEEE Symposium on VLSI Technology, 2015.

Références

- [Bauer05] M. Bauer, S. Zollner, N.D. Theodore, M. Canonico, P. Tomasini, B.Y. Nguyen, C. Arena, *Si₃H₈ based epitaxy of biaxially stressed silicon films doped with carbon and arsenic for CMOS applications*, Materials Research Society Symposium Proceedings, Vol 864, p.143-148, 2005.
- [Bauer06] M. Bauer, D. Weeks, Y. Zhang, V. Machkaoutsan, *Tensile Strained Selective Silicon Carbon Alloys for Recessed Source Drain Areas of Devices*, ECS Transactions, 3 (7), p.187-196, 2006.
- [Bauer07] M. Bauer, Y. Zhang, D. Weeks, V. Machkaoutsan, S.G. Thomas, *Defect free Embedded Silicon Carbon Stressor selectively grown into Recessed Source Drain areas of NMOS devices*, ECS Transactions, 6 (1), p.419-427, 2007.
- [Bauer10] M. Bauer, S.G. Thomas, *Selective epitaxial growth (SEG) of highly doped Si:P on Source/Drain areas of NMOS devices using Si₃H₈/PH₃/Cl₂ chemistry*, ECS Transactions, 33 (6), p.629-636, 2010.
- [Benoit15] D. Benoit, J. Mazurier, B. Varadarajan, S. Chhun, S. Lagrasta, C. Gaumer, D. Galpin, C. Fenouillet-Beranger, D. Barge, R. Duru, R. Beneyton, B. Gong, N. Sun, N. Chauvet, P. Ruault, D. Winandy, B. van Schravendijk, P. Meijer, O. Hinsinger, *Interest of SiCO Low k=4.5 Spacer Deposited at Low Temperature (400°C) in the perspective of 3D VLSI Integration*, IEEE International Electron Devices Meeting, 2015.
- [Besson08] P. Besson, V. Loup, T. Salvetat, N. Rochat, S. Lhostis, S. Favier, K. Dabertrand, V. Cosnier, *Critical Thickness Threshold in HfO₂ layers*, Solid State Phenomena, Vol134, p.67-70, 2008.
- [Bhuwalka04] K.K. Bhuwalka, S. Sedlmaier, A.K. Ludsteck, C. Tolskdorf, J. Schulze, I. Eisele, *Vertical tunnel first-effect transistor*, IEEE Transactions on Electron Devices, Vol51 Issue2, 2004.
- [Billoint16] O. Billoint, P. Batude, S. Thuries, C. Fenouillet-Beranger, G. Berhault, M. Brocard, M. Vinet, F. Clermidy, *3D Sequential Integration for Monolithic 3DIC Design*, Design, Automation & Test in Europe, 2016.
- [Black04] C.T. Black, O. Bezencenet, *Nanometer-scale pattern registration and alignment by directed deblock copolymer self-assembly*, IEEE Transactions on Nanotechnology, Vol3 Issue3, 2004.
- [Bogumilowicz05] Y. Bogumilowicz, J.M. Hartmann, R. Truche, Y. Campidelli, G. Rolland, T. Billon, *Chemical vapour etching of Si, SiGe and Ge with HCl; applications to the formation of thin relaxed SiGe buffers and to the revelation of threading dislocations*, Semiconductor Science and Technology, 20, p.127-134, 2005.
- [Brocard16] M. Brocard, R. Boumchedda, J.P. Noel, K.C. Akyel, B. Giraud, E. Beigne, D. Turgis, S. Thuries, G. Berhault, O. Billoint, *High Density SRAM Bitcell Architecture in 3D Sequential CoolCubeTM 14nm Technology*, IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference, 2016.

Références

- [Brocard17] M. Brocard, B. Mathieu, J-P. Colonna, C. Lopes Dos Santos, C. Fenouillet-Beranger, C-M. V. Lu, G. Cibrario, L. Brunet, P. Batude, F. Andrieu, S. Thuries, O. Billoint, *Transistor Temperature Deviation Analysis in Monolithic 3D Standard Cells*, papier soumis à IEEE Computer Society Annual Symposium on VLSI, 2017.
- [Brue195] M. Bruel, B. Aspar, B. Charlet, C. Maleville, T. Poumeyrol, A. Soubie, A.J. Auberton-Herve, J.M. Lamure, T. Barge, F. Metral, S. Trucchi, *"Smart cut": a promising new SOI material technology*, IEEE International SOI Conference, 1995.
- [Brunet14] L. Brunet, P. Batude, F. Fournel, L. Benaissa, C. Fenouillet-Beranger, L. Pasini, F. Deprat, B. Previtali, F. Ponthenier, A. Seignard, C. Euvrard-Colnat, M. Rivoire, P. Besson, C. Arvet, E. Beche, O. Rozeau, O. Billoint, O. Turkyilmaz, F. Clermidy, T. Signamarcheix, M. Vinet, *Direct Bonding: a Key Enabler for 3D Monolithic Integration*, ECS Transactions, Vol64 Issue5, p.381-390, 2014.
- [Brunet16] L. Brunet, P. Batude, C. Fenouillet-Beranger, P. Besombes, L. Hortemel, F. Ponthenier, B. Previtali, C. Tabone, A.Royer, C. Agraffeil, C. Euvrard-Colnat, A. Seignard, C. Morales, F. Fournel, L. Benaissa, T. Signamarcheix, P. Besson, M. Jourdan, R. Kachtouli, V. Benevent, J.-M. Hartmann, C. Comboroure, N. Allouti, N. Posseme, C. Vizioz, C. Arvet, S. Barnola, S. Kerdiles, L. Baud, L. Pasini, C.-M. V. Lu, F. Deprat, A. Toffoli, G. Romano, C. Guedj, V. Delaye, F. Boeuf, O. Faynot, M. Vinet, *First demonstration of a CMOS over CMOS 3D VLSI CoolCube™ integration on 300mm wafers*, IEEE Symposium on VLSI Technology, 2016.
- [Carron14] V. Carron, F. Nemouchi, F. Milesi, Y. Morand, P. Batude, L. Hutin, C. Le Royer, *Thermal stability enhancement of Ni-based silicides and germanides using W and F beam line ion implantation (BLII) for 3D CMOS sequential integration*, International Workshop on Junction Technology, 2014.
- [Casse11] M. Casse, X. Garros, O. Weber, F. Andrieu, G. Reimbold, F. Boulanger, *A study of N-induced traps due to a nitrated gate in high-k/metal gate nMOSFETs and their impact on electron mobility*, Solid-State Electronics, 65-66, p.139-145, 2011.
- [Chang08] K.M. Chang, B.N. Chen, S.M. Huang, *The effects of plasma treatment on the thermal stability of HfO₂ thin films*, Applied Surface Science, 254, p.6116-6118, 2008.
- [Charbonnier10] M. Charbonnier, C. Leroux, V. Cosnier, P. Besson, E. Martinez, N. Benedetto, C. Licitra, N. Rochat, C. Gaumer, K. Kaja, G. Ghibaudo, F. Martin, G. Reimbold, *Measurement of Dipoles/Roll-off/Work Functions by Coupling CV and IPE and Study of Their Dependence on Fabrication Process*, IEEE Transactions on Electron Devices, Vol 57 No8, p.1809-1819, 2010.
- [Charbonnier11] M. Charbonnier, C. Leroux, F. Allain, A. Toffoli, G. Ghibaudo, G. Reimbold, *Automatic full quantum analysis of CV measurements for bulk and SOI devices*, Microelectronics Engineering, 88, p.3404-3406, 2011.
- [Chau05] R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz, M. Radosavljevic, *Benchmarking nanotechnology for high-performance and low-power logic transistor applications*, IEEE Transactions on Nanotechnology, Vol4 Issue2, p.153-158, 2005.
- [Cheng01] K. Cheng, K. Hess, J.W. Lyding, *Deuterium Passivation of Interface Traps in MOS Devices*, IEEE Electron Device Letters, Vol22 No9, p.441-443, 2001.

Références

- [Cheng09] K. Cheng, A. Khakifirooz, P. Kulkarni, S. Kanakasabapathy, S. Schmitz, A. Reznicek, T. Adam, Y. Zhu, J. Li, J. Faltermeier, T. Furukawa, L. F. Edge, B. Haran, S.C. Seo, P. Jamison, J. Holt, X. Li, R. Loesing, Z. Zhu, R. Johnson, A. Upham, T. Levin, M. Smalley, J. Herman, M. Di, J. Wang, D. Sadana*, P. Kozlowski, H. Bu, B. Doris, J. O'Neill, *Fully Depleted Extremely Thin SOI Technology Fabricated by a Novel Integration Scheme Featuring Implant-Free, Zero-Silicon-Loss, and Faceted Raised Source/Drain*, *Symposium on VLSI Technology 2009*, p.212-213, 2009 .
- [Chung08] K.H. Chung, N. Yao, J. Benziger, J.C. Sturm, K.K. Singh, D. Carlson, S. Kuppurao, *Ultrahigh growth rate of epitaxial silicon by chemical vapor deposition at low temperature with neopentasilane*, *Applied Physics Letters*, 92, 113506, 2008.
- [Chung10] K.H. Chung, *Silicon-based epitaxy by chemical vapor deposition using novel precursor neopentasilane*, Phd thesis, 2010.
- [Clark00] D.E. Clark, D.C. Folz, J.K. West, *Processing materials with microwave energy*, *Materials Science and Engineering: A*, Vol287 Issue2, p.153-158, 2000.
- [Coudrain08] 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 with low temperature Fully Depleted SOI transistors*, *IEEE International Electron Devices Meeting*, 2008.
- [Das12] A. Das, A. Dorofeev, *Intel's 22nm process gives MOSFET switch a facelift*, disponible sur <http://www.embedded.com>, 2012.
- [Deal67] B.E. Deal, M. Sklar, A.S. Grove, E.H. Snow, *Characteristics of the Surface-State Charge (Q_{ss}) of Thermally Oxidized Silicon*, *Journal of The Electrochemical Society*, Vol114 No3, p.266-274, 1967.
- [Deprat15] F. Deprat, F. Nemouchi, C. Fenouillet-Beranger, P. Batude, B. Previtali, M. Danielou, P. Rodriguez, S. Favier, C. Fournier, P. Gergaud, M. Vinet, *NiCo 10 at%: A promising silicide alternative to NiPt 15 at% for thermal stability improvement in 3DVLSI integration*, *Advanced Metallization Conference*, 2015.
- [Deprat16] F. Deprat, C. Fenouillet-Beranger, C. Guerin, V. Jousseau, V. Beugin, C. Tabone, M. Billaud, N. Rochat, C. Licitra, D. Benoit, G. Imbert, V. Caubet-Hilloutou, P. Batude, M. Vinet, *Dielectrics Stability for Intermediate BEOL in 3D Sequential Integration*, *Materials for Advanced Metallization*, 2016.
- [Deshpande15] V. Deshpande, V. Djara, E. O'Connor, P. Hashemi, K. Balakrishnan, M. Sousa, D. Caimi, A. Olziersky, L. Czornomaz, J. Fompeyrine, *Advanced 3D Monolithic Hybrid CMOS with Sub-50 nm Gate Inverters Featuring Replacement Metal Gate (RMG)-InGaAs nFETs on SiGe-OI Fin pFETs*, *IEEE International Electron Devices Meeting*, 2015.
- [Destefanis08] V. Destefanis, J.M. Hartmann, M. Hopstaken, V. Delaye, D. Bensahel, *Low-thermal surface preparation, HCl etch and Si/SiGe selective epitaxy on (110) silicon surfaces*, *Semiconductor Science and Technology*, 23, 105018, 2008.

Références

- [Destefanis10] V. Destefanis, J.M. Hartmann, J.L. Huguenin, V. Delaye, M.P. Samson, P. Boulitreau, Y. Morand, P. Brianceau, C. Arvet, P. Gautier, *Fabrication, structural and electrical properties of (110) localized silicon-on-insulator devices*, Semiconductor Science and Technology, Vol 25, No4, p. 045014.1-045014.10, 2010.
- [Dupre08] C. Dupre, A. Hubert, S. Becu, M. Jublot, V. Maffini-Alvaro, C. Vizioz, F. Aussenac, C. Arvet, S. Barnola, J. -M. Hartmann, G. Garnier, F. Allain, J. -P. Colonna, M. Rivoire, L. Baud, S. Pauliac, V. Loup, T. Chevolleau, P. Rivallin, B. Guillaumot, G. Ghibaud, O. Faynot, T. Ernst, S. Deleonibus, *15nm-diameter 3D stacked nanowires with independent gates operation: Φ FET*, IEEE International Devices Meeting, 2008.
- [FabianPease08] R. Fabian Pease, S.Y. Chou, *Lithography and Other Patterning Techniques for Future Electronics*, Proceedings of the IEEE, Vol96 Issue2, 2008.
- [Fang68] F.F. Fang, A.B. Fowler, *Transport Properties of Electrons in Inverted Silicon Surfaces*, Physical Review, Vol169 No3, 1968.
- [Fenouillet10] C. Fenouillet-Beranger, O. Thomas, P. Perreau, J-P. Noel, A. Bajolet, S. Haendler, L. Tosti, S. Barnola, R. Beneyton, C. Perrot, C. de Buttet, F. Abbate, F. Baron, B. Pernet, Y. Campidelli, L. Pinzelli, P. Gouraud, M. Casse, C. Borowiak, O. Weber, F. Andrieu, K.K. Bourdelle, B.Y. Nguyen, F. Boedt, S. Denorme, F. Bœuf, O. Faynot, T. Skotnicki, *Efficient multi-VT FDSOI technology with UTBOX for low power circuit design*, Symposium on VLSI Technology, 2010.
- [Fenouillet12] C. Fenouillet-Beranger, P. Perreau, O. Weber, I. Ben-Akkez, A. Cros, A. Bajolet, S. Haendler, P. Fonteneau, P. Gouraud, E. Richard, F. Abbate, D. Barge, D. Pellissier-Tanon, B. Dumont, F. Andrieu, J. Passieux, R. Bon, V. Barral, D. Golanski, D. Petit, N. Planes, O. Bonin, W. Schwarzenbach, T. Poiroux, O. Faynot, M. Haond, F. Bœuf, *Enhancement of Devices Performance of hybrid FDSOI/Bulk Technology by using UTBOX sSOI substrates*, Symposium on VLSI Technology, 2012.
- [Fenouillet14a] C. Fenouillet-Beranger, B. Previtali, P. Batude, F. Nemouchi, M. Casse, X. Garros, L. Tosti, N. Rambal, D. Laffond, H. Dansas, L. Pasini, L. Brunet, F. Deprat, M. Gregoire, M. Mellier, M. Vinet, *FDSOI bottom MOSFETs stability versus top transistor thermal budget featuring 3D monolithic integration*, European Solid State Device Research Conference, 2014.
- [Fenouillet14b] C. Fenouillet-Beranger, B.Mathieu, B. Previtali, M-P. Samson, N.Rambal, V. Benevent, S. Kerdiles, J-P. Barnes, D. Barge, P. Besson, R. Kachtouli, P. Batude, M. Cassé, X. Garros, A. Laurent, F. Nemouchi, K. Huet, I. Toqué-Trésonne, D.Lafond, H. Dansas, F. Aussenac, G. Druais, P. Perreau, E. Richard, S. Chhun, E. Petitprez, N. Guillot, F. Deprat, L. Pasini, L. Brunet, V. Lu, C. Reita, M. Vinet, *New insights on bottom layer thermal stability and laser annealing promises for high performance 3D VLSI*, IEEE International Electron Devices Meeting, 2014.
- [Fenouillet15] C. Fenouillet-Beranger, S. Kerdilès, F. Deprat, P. Batude, M-P. Samson, B. Previtali, N. Rambal, V. Lapras, L. Emery, C. Euvrard-Colnat, A. Seignard, P. Besson, R. Kachtouli, A. Roman, C. Ribière, V. Lu, L. Brunet, E. Gourvest, G. Druais, Y. Loquet, L. Arnaud, Y. Le-Friec, O. Pollet, V. Benevent, F. Aussenac, H. Denis, V. Jousseau, S. Maitrejean, M. Vinet, *W and Copper interconnection stability for 3D VLSI CoolCube integration*, International Conference on Solid State Devices and Materials, 2015.

Références

- [Garros09] X. Garros, M. Casse, M. Rafik, C. Fenouillet-Beranger, G. Reimbold, F. Martin, C. Wiemer, F. Boulanger, *Process dependence of BTI reliability in advanced HK MG stacks*, Microelectronics Reliability, 49, p.982-988, 2009.
- [Goetzberger73] A. Goetzberger, A.D. Lopez, R.J. Strain, *On the Formation of Surface States during Stress Aging of Thermal Si-SiO₂ Interfaces*, Journal of Electrochemical Society, Vol120 No1, p.90-96, 1973.
- [Gouyé09] A.Gouye, O.Kermarrec, A.Halimaoui, Y.Campidelli, D.Rouchon, M. Burdin, P.Holliger, D.Bensahel, *Low-temperature RPCVD of Si, SiGe alloy, and Si_{1-y}C_y films on Si substrates using trisilane (Silcore®)*, Journal of Crystal Growth, 311, p.3522-3527, 2009.
- [Guha07] S. Guha, V.K. Paruchuri, M. Copel, V. Narayanan, Y.Y. Wang, P.E. Batson, N.A. Bojarczuk, B. Linder, B. Doris, *Examination of flatband and threshold voltage tuning of HfO₂/TiN field effect transistors by dielectric cap layers*, Applied Physics Letters, Vol90 Issue 9, 2007.
- [Gusev01] E.P. Gusev, D.A. Buchanan, E. Cartier, A. Kurnar, D. DiMaria, S. Guha, A. Callegari, S. Zafar, P.C. Jamison, D.A. Neumayer, M. Copel, M.A. Gribelyuk, H. Okom-Schmidt, C. D'Emic, P. Kozlowski, K. Chan, N. Bojarczuk, L-A. Ragnarsson, P. Ronsheim, K. Rim, R. J. Fleming, A. Mocuta, A. Ajmera, *Ultrathin high-k gate stacks for advanced CMOS devices*, IEEE Electron Devices Meeting, 2001.
- [Hartmann06] J.M. Hartmann, M. Burdin, G. Rolland, T. Billon, *Growth kinetics of Si and SiGe on Si(100), Si(110) and Si(111) surfaces*, Journal of Crystal Growth, Vol. 294, Issue 2, p.288-295, 2006.
- [Hartmann11] J.M. Hartmann, M. Py, J.P. Barnes, B. Prévitali, P. Batude, T. Billon, *Low temperature boron and phosphorous doped SiGe for recessed and raised sources and drains*, Journal of Crystal Growth, 327, p.68-77, 2011.
- [Hartmann12] J.M. Hartmann, V. Benevent, J.F. Damlencourt, T. Billon, *A benchmarking of silane, disilane and dichlorosilane for the low temperature growth of group IV layers*, Thin Solid Films, 520, p.3185-3189, 2012.
- [Hartmann13a] J.M. Hartmann, V. Benevent, J.P. Barnes, M. Veillerot, D. Lafond, J.F. Damlencourt, S. Morvan, B. Prévitali, F. Andrieu, N. Loubet, D. Dutartre, *Mushroom-free selective epitaxial growth of Si, SiGe, SiGe:B raised sources and drains*, Solid-State Electronics, 83, p10-17, 2013.
- [Hartmann13b] J.M. Hartmann, V. Benevent, J.P. Barnes, M. Veillerot, C. Deguet, *Disilane-based cyclic deposition/etch of Si, Si:P and Si_{1-y}C_y:P layers: I. The elementary process steps*, Semiconductor Science and Technology, 28, 025017, 2013.
- [Hartmann14] J.M. Hartmann, V. Benevent, A. André, C. Sirisopanaporn, M. Veillerot, M.P. Samson, S. Barraud, Z. Essay, B. Sermage, *Very Low Temperature (Cyclic) Deposition/Etch of In Situ Boron-Doped SiGe Raised Sources and Drains*, ECS Journal of Solid State Science and Technology, 3 (11), p.382-390, 2014.

Références

- [Hartmann16] J.M. Hartmann, *Current Trends in Group-IV Semiconductor Epitaxy for Nanoelectronics*, Thin Films on Silicon: Electronic and Photonic Applications, Materials and Energy Vol.8, Chapter 1, 2016.
- [Hatton06] B.D. Hatton, K. Landskron, W.J. Hunks, M.R. Bennett, D. Shukaris, D.D. Perovic, G.A. Ozin, *Materials chemistry for low-k materials*, Materials Today, Volume 9 Issue 3, p.22-31, 2006.
- [Havemann02] R.H. Havemann, J.A. Hutchby, *High-performance interconnects: an integration overview*, Proceedings of the IEEE, Vol89 Issue5, p.586-601, 2002.
- [Hisamoto00] D. Hisamoto, W-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T-J. King, J. Bokor, C. Hu, *FinFET – A Self-Aligned Double-Gate MOSFET Scalable to 20nm*, IEEE Transactions on Electron Devices, Vol47 No12, p.2320-2325, 2000.
- [Hobbs04] C.C. Hobbs, L.R.C. Fonseca, A. Knizhnik, V. Dhandapani, S.B. Samavedam, W.J. Taylor, J.M. Grant, L.G. Dip, D.H. Triyoso, R.I. Hegde, D.C. Gilmer, R. Garcia, D. Roan, M.L. Lovejoy, R.S. Rai, E.A. Hebert, H.H. Tseng, S.G.H. Anderson, B.E. White, P.J. Tobin, *Fermi-Level Pinning at the Polysilicon/Metal-Oxide Interface-Part II*, IEEE Transactions on Electron Devices, Vol51 No6, p.978-984, 2004.
- [Hook05] T.B. Hook, R. Bolam, W. Clark, J. Burnham, N. Rovedo, L. Schutz, *Negative bias temperature instability of three oxide thicknesses (1.4/2.2/5.2 nm) with nitridation variations and deuteration*, Microelectronics Reliability, 45, p.47-59, 2005.
- [Horstmann05] M. Horstmann, A. Wei, T. Kammler, J. Höntschel, H. Bierstedt, T. Feudel, K. Froberg, M. Gerhardt, A. Hellmich, K. Hempel, J. Hohage, P. Javorka, J. Klais, G. Koerner, M. Lenski, A. Neu, R. Otterbach, P. Press, C. Reichel, M. Trentsch, B. Trui, H. Salz, M. Schaller, H.-J. Engelmann, O. Herzog, H. Ruelke, P. Hübler, R. Stephan, D. Greenlaw, M. Raab, N. Kepler, *Integration and Optimization of Embedded-SiGe, Compressive and Tensile Stressed Liner Films, and Stress Memorization in Advanced SOI CMOS Technologies*, IEEE International Electron Devices Meeting, 2005.
- [Huang08] E. Huang, E. Joseph, H. Bu, X. Wang, N. Fuller, C. Ouyang, E. Simonyi, H. Shobha, T. Cheng, A. Mallikarjunan, I. Lauer, S. Fang, W. Haensch, C.Y. Sung, S. Purushothaman, G. Shahidi, *Low-k Spacers for Advanced Low Power CMOS Devices with Reduced Parasitic Capacitances*, IEEE International SOI Conference Proceedings, 2008.
- [Huard03] V. Huard, F. Monsieur, G. Ribes, S. Bruyere, *Evidence for hydrogen-related defects during NBTI stress in p-MOSFETs*, International Reliability Physics Symposium Proceedings, p.178-182, 2003.
- [Hutin16] L. Hutin, R. Maurand, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, X. Jehl, S. Barraud, S. De Franceschi, M. Sanquer, M. Vinet, *Si CMOS Platform for Quantum Information Processing*, IEEE Symposium on VLSI Technology, 2016.
- [Hykavyy16] A. Hikavyy, E. Rosseel, S. Kubicek, G. Mannaert, P. Favia, H. Bender, R. Loo, N. Horiguchi, *Properties and growth peculiarities of Si_{0.30}Ge_{0.70} stressor integrated in 14 nm fin-based p-type metal-oxide-semiconductor field-effect transistors*, Thin Solid Films 602, p.72–77, 2016.

Références

- [Irisawa14] T. Irisawa, K. Ikeda, Y. Moriyama, M. Oda, E. Mieda, T. Maeda, T. Tezuka, *Demonstration of Ultimate CMOS based on 3D Stacked InGaAs-OI/SGOI Wire Channel MOSFETs with Independent Back Gate*, IEEE Symposium on VLSI Technology, 2014.
- [Ishida98] M. Ishida, T. Kawakami, A. Tsuji, N. Kawamoto, M. Motoyoshi, N. Ouchi, *Novel 6T-SRAM cell technology designed with rectangular patterns scalable beyond 0.18 μ m generation and desirable for ultra high speed operation*, IEEE International Electron Devices Meeting, p.201-204, 1998.
- [Iyer92] S.S. Iyer, K. Eberl, M.S. Goorsky, F.K. LeGoues, J.C. Tsang, F. Cardone, *Synthesis of Si_{1-y}C_y alloys by molecular beam epitaxy*, Applied Physics Letters, 60, p.356-358, 1992.
- [Jahan05] C. Jahan, O. Faynot, L. Tosti, J.M. Hartmann, *Agglomeration control during the selective epitaxial growth of Si raised sources and drains on ultra-thin silicon-on-insulator substrates*, Journal of Crystal Growth, Volume 280, Issues 3-4, p.530-538, 2005.
- [Jung07] S-M. Jung, H. Lim, C. Yeo, K. Kwak, B. Son, H. Park, J. Na, J-J. Shim, C. Hong, 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*, IEEE Symposium on VLSI Technology, 2007.
- [Kanoh93] H. Kanoh, O. Sugiura, M. Matsumura, *Chemical Vapor Deposition of Amorphous Silicon Using Tetrasilane*, Japanese Journal of Applied Physics, Volume 32, p.2613-2619, 1993.
- [Kerdiles16] S. Kerdiles, P. Acosta Alba, B. Mathieu, M. Veillerot, R. Kachtouli, P. Besson, H. Denis, F. Mazzamuto, I. Toque-Tresonne, K. Huet, C. Fenouillet-Beranger, *Dopant activation and crystal recovery in arsenic-implanted ultra-thin silicon-on-insulator structures using 308nm nanosecond laser annealing*, International Workshop on Junction Technology, 2016.
- [Kern90] W. Kern, *The Evolution of Silicon Wafer Cleaning Technology*, Journal of Electrochemical Society, Vol 137 No6, p1887-1892, 1990.
- [Kim04] G. Kim, P.C. McIntyre, C.O. Chui, K.C. Saraswat, S. Stemmer, *Engineering chemically abrupt high-k metal oxide/silicon interfaces using an oxygen-gettering metal overlayer*, Journal of Applied Physics, Vol96 No6, p.3467-3472, 2004.
- [Kimizuka00] N. Kimizuka, K. Yamaguchi, K. Imai, T. Iizuka, C.T. Liu, R.C. Keller, T. Horiuchi, *NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10-/spl mu/m gate CMOS generation*, Symposium on VLSI Technology, 2000.
- [Ko08] C.H. Ko, T.M. Kuan, K. Zhang, G. Tsai, S.M. Seutter, C.H. Wu, T.J. Wang, C.N. Ye, H.W. Chen, C.H. Ge, K.H. Wu, W.C. Lee, *A Novel CVD-SiBCN Low-K Spacer Technology for High-Speed Applications*, Symposium on VLSI Technology, 2008
- [Kuhn12] K.J. Kuhn, *Considerations for Ultimate CMOS Scaling*, Vol59, Issue7, IEEE Transactions on Electron Devices, Vol59 Issue7, p.1813-1828, 2012.
- [Kuo06] Y. Kuo, J. Lu, S. Chatterjee, J. Yan, H.C. Kim, T. Yuan, W. Luo, J. Peterson, M. Gardner, *Sub 2nm Thick Zirconium Doped Hafnium Oxide High-K Gate Dielectrics*, ECS Transactions, Vol1 Issue5, p.447-454, 2006.

Références

- [Labrot16] M. Labrot, F. Cheynis, D. Barge, P. Müller, M. Juhel, *Low thermal budget for Si and SiGe surface preparation for FD-SOI technology*, Applied Surface Science, Volume 371, p.436-446, 2016.
- [Lee14] Y.-J. Lee, T.-C. Cho, S.-S. Chuang, F.-K. Hsueh, Y.-L. Lu, P.-J. Sung, H.-C. Chen, M.-I. Current, T.-Y. Tseng, T.-S. Chao, C. Hu, F.-L. Yang, *Low-Temperature Microwave Annealing Processes for Future IC Fabrication – A Review*, IEEE Transactions on Electron Devices, Vol61 No3, p.651-665, 2014.
- [Lei06] J. Lei, S.E. Phan, X. Lu, C.T. Kao, K. Lavu, K. Moraes, K. Tanaka, B. Wood, B. Ninan, S. Gandikota, *Advantage of SiconiTM Pre-clean over Wet Clean for Pre Salicide Applications Beyond 65nm Node*, IEEE International Symposium on Semiconductor Manufacturing, 2006
- [Leroux07] C. Leroux, F. Allain, A. Toffoli, G. Ghibaudo, G. Reimbold, *Automatic statistical full quantum analysis of C-V and I-V characteristics for advanced MOS gate stacks*, Microelectronic Engineering, 84, p.2408-2411, 2007.
- [Leroyer11] C. Le Royer, A. Villalon, M. Casse, D. Cooper, J. Mazurier, B. Prévitali, C. Tabone, P. Perreau, J.M. Hartmann, P. Scheiblein, F. Allain, F. Andrieu, O. Weber, P. Batude, O. Faynot, T. Poiroux, *First demonstration of ultrathin body c-SiGe channel FDSOI pMOSFETs combined with SiGe(:B) RSD: drastic improvement of electrostatics (V_{th} , p Tuning, DIBL) and transport (μ_0 , I_{sat}) properties down to 23nm gate length*, Proceedings Of 2011 IEDM Conference, p. 16.5.1-16.5.4, 2011.
- [LeRoyer12] C. Le Royer, A. Villalon, D. Cooper, F. Andrieu, J.-M. Hartmann, P. Perreau, B. Prévitali, *High Performance FDSOI MOSFETs and TFETs using SiGe Channels and Raised Source and Drain*, International Silicon-Germanium Technology and Device Meeting, 2012.
- [Liu05] X. Liu, Q. Tang, J.S. Harris, T.I. Kamins, *Arsenic surface segregation during in situ doped silicon and $Si_{1-x}Ge_x$ molecular beam epitaxy*, Journal of Crystal Growth, 281, p334-343, 2005.
- [Liu12] C. Liu, S.K. Lim, *Ultra-High Density 3D SRAM Cell Designs for Monolithic 3D Integration*, IEEE International Interconnect Technology Conference, 2012.
- [Loubet12] N. Loubet, T. Adam, M. Raymond, Q. Liu, K. Cheng, R. Sreenivasan, A. Reznicek, P. Khare, W. Kleemeier, V. Paruchuri, B. Doris, R. Sampson, *Ultra-low resistivity in-situ phosphorus doped Si and SiC epitaxy for source/drain formation in advanced 20 nm n-type field effect transistor devices*, Thin Solid Films, 520, p.3140-3154, 2012.
- [Lu15] C.-M. V. Lu, C. Bout, C. Fenouillet-Beranger, A. Roule, M.P. Samson, B. Prévitali, C. Arvet, A. Michallet, N. Rochat, S. Favier, R. Kachtouli, V. Loup, P. Besson, M. Garcia-Barros, N. Posseme, F. Pierre, P. Maury, D. Benoit, P. Batude, M. Vinet, T. Skotnicki, *A Novel Low Temperature (<500°C) and Low-k (3.8) Boron Nitride PECVD Offset Spacer Featuring 3D VLSI Integration*, International Conference on Solid State Devices and Materials, 2015.
- [Lu16a] C.-M. V. Lu, C. Fenouillet-Beranger, C. Bout, A. Roule, V. Beugin, J. Fort, C. Arvet, N. Posseme, V. Loup, P. Besson, M.P. Samson, B. Prévitali, C. Tabone, A. Michallet, N. Rochat, D. Benoit, F. Pierre, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, *Integration of Low Temperature 480°C SiOCN as Offset Spacer in view of 3D Sequential Integration*, International Conference on Solid State Devices and Materials, 2016.

Références

- [Lu16b] C.-M. V. Lu, C. Fenouillet-Béranger, J.M. Hartmann, Ph. Rodriguez, V. Benevent, M.-P. Samson, B. Prévitali, C. Tabone, M. Cassé, F. Allain, G. Romano, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, *Integration of Low Temperature SiGe:B Raised Sources and Drains in p-type FDSOI Field Effect Transistors*, ECS Transactions, Vol 75 Issue 8, p51-58, 2016.
- [Lu16c] C.-M. V. Lu, C. Fenouillet-Béranger, R. Gassilloud, H. Graoui, D. Larmagnac, S. Sharma, X. Garros, C. Leroux, A. Toffoli, G. Romano, R. Kies, N. Rambal, M.-P. Samson, B. Previtalli, C. Guedj, N. Bernier, S. Kerdiles, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, *Influence of Low Thermal Budget Plasma Oxidation and Millisecond Laser Anneal on Gate Stack Reliability in view of 3D Sequential*, International Conference on Solid State Devices and Materials, 2016.
- [Lu17a] C.-M. V. Lu, F. Deprat, C. Fenouillet-Beranger, P. Batude, X. Garros, A. Tsiara, C. Leroux, R. Gassilloud, D. Nouguier, D. Ney, X. Federspiel, P. Besombes, A. Toffoli, G. Romano, N. Rambal, V. Delaye, D. Barge, M.-P. Samson, B. Previtalli, C. Tabone, L. Pasini, L. Brunet, F. Andrieu, J. Micoud, T. Skotnicki, M. Vinet, *Key process steps for high performance and reliable 3D Sequential Integration*, IEEE Symposium on VLSI Technology, 2017.
- [Lu17b] C.-M. V. Lu, C. Fenouillet-Beranger, M. Brocard, O. Billoint, G. Cibrario, L. Brunet, X. Garros, C. Leroux, M. Casse, A. Laurent, A. Toffoli, G. Romano, R. Kies, R. Gassilloud, N. Rambal, V. Lapras, M.-P. Samson, C. Tallaron, C. Tabone, B. Previtalli, D. Barge, A. Ayres, L. Pasini, P. Besombes, F. Andrieu, P. Batude, T. Skotnicki, M. Vinet, *Dense N over CMOS 6T SRAM Cells using 3D Sequential Integration*, International Symposium on VLSI Technology, Systems and Applications, 2017.
- [Magee81] T.J. Magee, C. Leung, H. Kawayoshi, L.J. Palkuti, B.K. Furman, C.A. Evans Jr., L.A. Christel, J.F. Gibbons, D.S. Day, *Recoil oxygen implants and thermal redistribution of oxygen in through-oxide arsenic-implanted Si*, Applied Physics Letters, Vol39 Issue 7, p.564-566, 1981.
- [Mahapatra13] S. Mahapatra, N. Goel, S. Desai, S. Gupta, B. Jose, S. Mukhopadhyay, K. Joshi, A. Jain, A.E. Islam, M.A. Alam, *A Comparative Study of Different Physics-Based NBTI Models*, IEEE Transactions on Electron Devices, Vol60 No3, 2013.
- [Mathieu15] B. Mathieu, C. Fenouillet-Beranger, S. Kerdiles, J.C. Barbe, *Thermal Simulation of Nanosecond Laser Annealing of 3D Sequential VLSI*, International Conference on Simulation of Semiconductor Processes and Devices, 2015.
- [Meyerson90] B.S. Meyerson, F.J. Himpsel, K.J. Uram, *Bistable conditions for low-temperature silicon epitaxy*, Applied Physics Letters, 57, p1034-1036, 1990.
- [Mistry07] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neiryneck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, *A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging*, IEEE Electron Devices Meeting, 2007.

Références

- [Nicollian65] E.H. Nicollian, A. Goetzberger, *MOS Conductance Technique for Measuring Surface State Parameters*, Applied Physics Letters, 7(8), p.216-219, 1965.
- [Ouerghi15] I. Ouerghi, M. Sansa, W. Ludurczak, L. Duraffourg, K. Benedetto, P. Besombes, T. Moffitt, B. Adams, D. Larmagnac, P. Gergaud, C. Poulain, A. I. Vidaña, C. Ladner, J. M. Fabbri, D. Muyard, G. Rodriguez, G. Rabille, O. Pollet, P. Brianceau, S. Kerdiles, S. Hentz, T. Ernst, *Polysilicon Nanowire NEMS fabricated at low temperature for above IC NEMS mass sensing applications*, IEEE International Electron Devices Meeting, 2015.
- [Ouerghi16] I. Ouerghi, M. Sansa, L. Duraffourg, K. Benedetto, P. Besombes, K. Huet, F. Mazzamuto, I. Toque-Tresonne, C. Ladner, S. Kerdiles, W. Ludurczak, T. Ernst, *High performance polysilicon nems fabricated at low temperature with UV laser annealing*, International Conference on Micro Electro Mechanical Systems, 2016.
- [Pasini14] L. Pasini, B. Sklenard, P. Batude, M. Casse, P. Rivallin, B. Previtali, C. Fenouillet-Beranger, M. Haond, G. Ghibaudo, M. Vinet, *Insights in Accesses Optimization for nFET Low Temperature Fully Depleted Silicon on Insulator Devices*, International Workshop on Junction Technology, 2014.
- [Pasini15] L. Pasini, P. Batude M. Cassé, B. Mathieu, B. Sklenard, F. Piegas Luce, S. Reboh, N. Bernier, C. Tabone, O. Rozeau, S. Martini, C. Fenouillet-Beranger, L. Brunet, G. Audoit, D. Lafond, F. Aussenac, F. Allain, G. Romano, S. Barraud, N. Rambal, V. Barral, L. Hutin, J-M. Hartmann, P. Besson, S. Kerdiles, M. Haond, G. Ghibaudo, and M. Vinet, *High performance low temperature activated devices and optimization guidelines for 3D VLSI integration of FD, TriGate, FinFET on insulator*, Symposium on VLSI Technology, 2015.
- [Pasini16a] L. Pasini, *Optimisation des jonctions de dispositifs (FDSOI, TriGate) fabriqués à faible temperature pour l'intégration 3D séquentielle*, Thèse de l'Université Grenoble Alpes.
- [Pasini16b] L. Pasini, P. Batude, J. Lacord, M. Casse, B. Mathieu, B. Sklenard, F. Piegas Luce, J. Micout, A. Payet, F. Mazen, P. Besson, E. Ghegin, J. Borrel, R. Daubriac, L. Hutin, D. Blachier, D. Barge, S. Chhun, V. Mazzocchi, A. Cros, J-P. Barnes, Z. Saghi, V. Delaye, N. Rambal, V. Lapras, J. Mazurier, O. Weber, F. Andrieu, L. Brunet, C. Fenouillet-Beranger, Q. Rafhay, G. Ghibaudo, F. Cristiano, M. Haond, F. Bœuf, M. Vinet, *High performance CMOS FDSOI devices activated at low temperature*, Symposium on VLSI Technology, 2016.
- [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, Vol22 Issue1, p.297 (2004).
- [Posseme14] N. Posseme, O. Pollet, S. Barnola, *Alternative process for thin layer etching: Application to nitride spacer etching stopping on silicon germanium*, Applied Physics Letters, 105, 051605, 2014.
- [Rajendran00] K. Rajendran, W. Schoenmaker, S. Decoutere, M. Caymax, *Simulation of boron diffusion in strained $Si_{1-x}Ge_x$ epitaxial layers*, International Conference on Simulation of Semiconductor Processes and Devices, 2000.

Références

- [Regis97] J.M. Regis, A.M. Joshi, T. Lill, M. Yu, *Reactive Ion Etch of Silicon Nitride Spacer With High Selectivity to Oxide*, IEEE/SEMI Advanced Semiconductor Manufacturing Conference, 1997.
- [Reznicek12] A. Reznicek, T.N. Adam, J. Li, Z. Zhu, R. Murphy, S.W. Bedell, V. Paruchuri, D.K. Sadana, *Microstructure development in epitaxially grown in-situ Boron and Carbon co-doped strained 60% Silicon-Germanium layers*, ECS Transactions, 50(9), p.1013-1024, 2012.
- [Ribes05] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, G. Ghibaudo, *Review on High-k Dielectrics Reliability Issues*, IEEE Transactions on Device and Materials Reliability, Vol5 No1, 2005.
- [Robertson00] J. Robertson, *Band offsets of wide-band-gap oxides and implications for future electronic devices*, Journal of Vacuum Science & Technology B, Vol18 Issue3, 2000.
- [Robertson06] J. Robertson, *High dielectric constant gate oxides for metal oxide Si transistors*, Reports on Progress in Physics, 69, p.327-396, 2006.
- [Robertson15] J. Robertson, R.M. Wallace, *High-k materials and metal gates for CMOS applications*, Materials Science and Engineering, 88, p.1-41, 2015.
- [Shao13] Q. Shao, F. Chen, L. Jiang, M. Li, Q. Wang, P. Zhu, J. Cheng, S. Xiong, H. Liu, *High-k metal gate CMP process development for 32nm & beyond gate last approach*, ECS Transactions, Vol52 Issue1, p.523-528, 2013.
- [Sharia08] O. Sharia, A.A. Demkov, G. Bersuker, B.H. Lee, *Effects of aluminum incorporation on band alignment at the SiO₂/HfO₂ interface*, Physical Review B – Condensed Matter and Materials Physics, Vol77 Issue8, 2008.
- [Shen14] C-H. Shen, J-M. Shieh, W-H. Huang, T-T. Wu, C-F. Chen, M-H. Kao, C-C. Yang, C-D. Lin, H-H. Wang, T-Y. Hsieh, B-Y. Chen, G-W. Huang, M-F. Chang, F-L. Yang, *Heterogeneously integrated sub-40nm low-power epi-like Ge/Si monolithic 3D-IC with stacked SiGeC ambient light harvester*, IEEE International Electron Devices Meeting, 2014.
- [Shulaker14] M.M. Shulaker, K. Saraswat, H-S.P. Wong, S. Mitra, *Monolithic Three-Dimensional Integration of Carbon Nanotube FETs with Silicon CMOS*, IEEE Symposium on VLSI Technology, 2014.
- [Skotnicki03] T. Skotnicki, F. Boeuf, *Introduction à la Physique du Transistor MOS (Physique des dispositifs pour circuits intégrés silicium)*, Lavoisier, 2003
- [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, 2007.
- [Song06] S.C. Song, Z. Zhang, C. Huffman, J.H. Sim, S.H. Bae, P.D. Kirsch, P. Majhi, R. Choi, N. Moumen, B.H. Lee, *Highly Manufacturable Advanced Gate-Stack Technology for Sub-45-nm Self-Aligned Gate-First CMOSFETs*, IEEE Transactions on Electron Devices, Vol53 No5, p.979-988, 2006.

Références

- [Tamura08] N. Tamura, Y. Shimamune, *45nm CMOS technology with low temperature selective epitaxy of SiGe*, Applied Surface Science, 254, p.6067-6071, 2008.
- [Tang08] S. Tang, R.M. Wallace, A. Seabaugh, D. King-Smith, *Evaluating the minimum thickness of gate oxide on silicon using first-principles method*, Applied Surface Science, 135, p.137-142, 1998.
- [Taylor02] P. Taylor, D. Sinha, *Preventing Cu Contamination from BEOL Processing*, Controlled Environments Magazine, 2002.
- [Tsukijawa03] S. Tsukijawa, T. Mine, K. Watanabe, Y. Shimamoto, R. Tsuchiya, K. Ohnishi, T. Onai, J. Yugami, S. Kimura, *Negative Bias Temperature Instability of pMOSFETs with Ultra-thin SiON Gate Dielectrics*, IEEE International Reliability Physics Symposium Proceedings, p.183-188, 2003.
- [Tsutsui16] G. Tsutsui, R. Bao, K.y. Lim, R.R. Robison, R.A. Vega, J. Yang, Z. Liu, M. Wang, O. Gluschenkov, C.W. Yeung, K. Watanabe, S. Bentley, H. Niimi, D. Liu, H. Zhou, S. Siddiqui, H. Kim, R. Galatage, R. Venigalla, M. Raymond, P. Adusumilli, S. Mochizuki, T.S. Devarajan, B. Miao, B. Liu, A. Greene, J. Shearer, P. Montanini, J.W. Strane, C. Prindle, E.R. Miller, J. Fronheiser, C.C. Niu, K. Chung, J.J. Kelly, H. Jagannathan, S. Kanakasabapathy, G. Karve, F.L. Lie, P. Oldiges, V. Narayanan, T.B. Hook, A. Knorr, D. Gupta, D. Guo, R. Divakaruni, H. Bu, M. Khare, *Technology viable DC performance elements for Si/SiGe channel CMOS FinFET*, IEEE International Electron Device Meeting, 456, 2016
- [Vandervorst16] W. Vandervorst, C. Fleischmann, J. Bogdanowicz, A. Franquet, U. Celano, K. Paredis, A. Budrevich, *Dopant, composition and carrier profiling for 3D structures*, Materials Science in Semiconductor Processing, 2016.
- [Velo14] A. Veloso, G. Boccardi, L.A. Ragnarsson, Y. Higuchi, H. Arimura, J.W. Lee, E. Simoen, M.J. Cho, P.J. Roussel, V. Paraschiv, X. Shi, T. Schram, S.A. Chew, S. Brus, A. Dangol, E. Vecchio, F. Sebaai, K. Kellens, N. Heylen, K. Devriendt, H. Dekkers, A. Van Ammel, T. Witters, T. Conard, I. Vaesen, O. Richard, H. Bender, R. Athimulam, T. Chiarella, A. Thean, N. Horiguchi, *Thermal and plasma treatments for improved (sub-)1nm equivalent oxide thickness planar and FinFET-based replacement metal gate high-k last devices and enabling a simplified scalable CMOS integration scheme*, Japanese Journal of Applied Physics, 53, 2014.
- [Verheyen08] P. Verheyen, V. Machkaoutsan, M. Bauer, D. Weeks, C. Kerner, F. Clemente, H. Bender, D. Shamiryan, R. Loo, T. Hoffmann, P. Absil, S. Biesemans, S.G. Thomas, *Strain Enhanced nMOS Using In Situ Doped Embedded Si_{1-x}C_x S/D Stressors With up to 1.5% Substitutional Carbon Content Grown Using a Novel Deposition Process*, IEEE Electron Device Letters, Vol. 29, No. 11, 2008.
- [Watanabe05] H. Watanabe, S. Kamiyama, N. Umezawa, K. Shiraishi, S. Yoshida, Y. Watanabe, T. Arikado, T. Chikyow, K. Yamada, K. Yasutake, *Role of Nitrogen Incorporation into Hf-Based High-k Gate Dielectrics for Termination of Local Current Leakage Paths*, Japanese Journal of Applied Physics, Vol 44 No 43, p.1333-1336, 2005.
- [Weber15a] O. Weber, E. Josse, M. Haond, *14nm FDSOI Technology for High-Speed and Energy-Efficient CMOS*, ECS Transactions, 66(5), p.37-42, 2015.

Références

- [Weber15b] O. Weber, E. Josse, J. Mazurier, N. Degors, S. Chhun, P. Maury, S. Lagrasta, D. Barge, J-P. Manceau, M. Haond, *14nm FDSOI upgraded device performance for ultra-low voltage operation*, IEEE Symposium on VLSI Technology, 2015.
- [Wilk01] G.D. Wilk, R.M. Wallace, J.M. Anthony, *High-k gate dielectrics: Current status and materials properties considerations*, Journal of Applied Physics, Vol89 No10, p.5243- 5275, 2001.
- [Wong08] H.S. Wong, K.W. Ang, L. Chan, K.M. Hoe, C.H. Tung, N. Balasubramanian, D. Weeks, M. Bauer, J. Spear, S.G. Thomas, G. Samudra, Y.C. Yeo, *Silicon-Carbon Stressors With High Substitutional Carbon Concentration and In Situ Doping Formed in Source/Drain Extensions of n-Channel Transistors*, IEEE Electron Device Letters, Vol 29, No 5, 2008.
- [Wostyn14] K. Wostyn, K. Kenis, D. Rondas, R. Loo, A. Hikavyy, S. Dhayalan, B. Douhard, P. Mertens, F. Holsteyns, S. De Gendta, G. Simpson, G. Bast, K. Swaminathan, *Evaluation of the Si_{0.8}Ge_{0.2}-on-Si Epitaxial Quality By Inline Surface Light Scattering: A Case Study On The Impact of Interfacial Oxygen*, ECS Transactions, 64 (6) p.989-995, 2014.
- [Wu13] Y. Wu, D.B. Farmer, F. Xia, P. Avouris, *Graphene Electronics: Materials, Devices, and Circuits*, Proceedings of IEEE, Vol101 Issue7, 2013.
- [Yaegashi15] H. Yaegashi, K. Oyama, A. Hara, S. Natori, S. Yamauchi, M. Yamato, N. Okabe, K. Koike, *Enabling capability of multi-patterning towards 10nm and beyond*, China Semiconductor Technology International Conference, 2015.
- [Yamashita15] T. Yamashita, S. Mehta, V.S. Basker, R. Southwick, A. Kumar, R. Kambhampati, R. Sathiyarayanan, J. Johnson, T. Hook, S. Cohen, J. Li, A. Madan, Z. Zhu, L. Tai, Y. Yao, P. Chinthamanipeta, M. Hopstaken, Z. Liu, D. Lu, F. Chen, S. Khan, D. Canaperi, B. Haran, J. Stathis, P. Oldiges, C.H. Lin, S. Narasimha, A. Bryant, W.K. Henson, S. Kanakasabapathy, K.V.R.M. Murali, T. Gow, D. McHerron, H. Bu, M. Khare, *A Novel ALD SIBCN Low-k Spacer for Parasitic Capacitance Reduction in FinFETs*, Symposium on VLSI Technology, 2015.
- [Yang15] C-C. Yang, J-M. Shieh, T-Y Hsieh, W-H. Huang, H-H. Wang, C-H. Shen, T-T. Wu, Y-F. Hou, Y-J. Chen, Y-J. Lee, M-C. Chen, F-L. Yang, Y-H Chen, M-C. Wu, W-K. Yeh, *Enabling low power BEOL compatible monolithic 3D+ nanoelectronics for IoTs using local and selective far-infrared ray laser anneal technology*, IEEE International Electron Devices Meeting, 2016.
- [Yeap13] G. Yeap, *Smart mobile SoCs driving the semiconductor industry: Technology trend, challenges and opportunities*, IEEE International Electron Device Meeting, 2013
- [Yeo01] Y.C. Yeo, Q. Lu, P. Ranade, H. Takeuchi, K.J. Yang, I. Polishchuk, T.J King, C. Hu, S.C. Song, H.F. Luan, D.L. Kwong, *Dual-Metal Gate CMOS Technology with Ultrathin Silicon Nitride Gate Dielectric*, IEEE Electron Device Letters, Vol22, No5, 2001.
- [Yeo06] Y.C. Yeo, *Silicon-Carbon Source/Drain: Selective Epitaxy, Process Integration and Transistor Strain Engineering*, ECS Transactions, 3 (7), p.1143-1150, 2006.
- [Yeow75] Y.T. Yeow, D.R. Lamb, S.D. Brotherton, *An investigation of the influence of low-temperature annealing treatments on the interface state density at the Si-SiO₂ interface*, Journal of Physics D, Applied Physics, Vol8 No13, p.1495-1506, 1975.

Références

[Yu04] H.Y. Yu, C. Ren, Y.C. Yeo, J.F. Kang, X.P. Wang, H.H.H. Ma, M.F. Li, D.S.H. Chan, D.L. Kwong, *Fermi Pinning-Induced Thermal Instability of Metal-Gate Work Functions*, IEEE Electron Device Letters, Vol25 No5, 2004.

Liste de publications

- Auteur principal

- **A Novel Low Temperature (<500°C) and Low-k (3.8) Boron Nitride PECVD Offset Spacer Featuring 3D VLSI Integration**, C.-M. V. Lu, C. Bout, C. Fenouillet-Beranger, A. Roule, M.P. Samson, B. Prévitali, C. Arvet, A. Michallet, N. Rochat, S. Favier, R. Kachtouli, V. Loup, P. Besson, M. Garcia-Barros, N. Posseme, F. Pierre, P. Maury, D. Benoit, P. Batude, M. Vinet, T. Skotnicki, **International Conference on Solid State Devices and Materials**, 2015, poster.
- **Influence of Low Thermal Budget Processing and Post-Processing on Performance and Reliability of HfO₂/TiN Gate-First for 3D VLSI Integration**, C.-M. V. Lu, R. Gassilloud, C. Fenouillet-Beranger, C. Leroux, X. Garros, A. Toffoli, N. Rambal, V. Lapras, J.-M. Pedini, M.-P. Samson, B. Prévitali, N. Bernier, G. Audoit, V. Delaye, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, **IEEE Semiconductor Interface Specialists Conference**, 2015, présentation orale (papier non publié).
- **Integration of Low Temperature 480°C SiOCN as Offset Spacer in view of 3D Sequential Integration**, C.-M. V. Lu, C. Fenouillet-Beranger, C. Bout, A. Roule, V. Beugin, J. Fort, C. Arvet, N. Posseme, V. Loup, P. Besson, M.P. Samson, B. Prévitali, C. Tabone, A. Michallet, N. Rochat, D. Benoit, F. Pierre, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, **International Conference on Solid State Devices and Materials**, 2016, poster.
- **Influence of Low Thermal Budget Plasma Oxidation and Millisecond Laser Anneal on Gate Stack Reliability in view of 3D Sequential**, C.-M. V. Lu, C. Fenouillet-Béranger, R. Gassilloud, H. Graoui, D. Larmagnac, S. Sharma, X. Garros, C. Leroux, A. Toffoli, G. Romano, R. Kies, N. Rambal, M.-P. Samson, B. Prévitali, C. Guedj, N. Bernier, S. Kerdiles, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, **International Conference on Solid State Devices and Materials**, 2016, poster.
- **Integration of Low Temperature SiGe:B Raised Sources and Drains in p-type FDSOI Field Effect Transistors**, C.-M. V. Lu, C. Fenouillet-Béranger, J.M. Hartmann, Ph. Rodriguez, V. Benevent, M.-P. Samson, B. Prévitali, C. Tabone, M. Cassé, F. Allain, G. Romano, L. Brunet, P. Batude, T. Skotnicki, M. Vinet, **ECS Transactions**, Vol 75 Issue 8, p51-58, 2016, présentation orale & papier étendu.
- **Dense N over CMOS 6T SRAM Cells using 3D Sequential Integration**, C.-M. V. Lu, C. Fenouillet-Beranger, M. Brocard, O. Billoint, G. Cibrario, L. Brunet, X. Garros, C. Leroux, M. Casse, A. Laurent, A. Toffoli, G. Romano, R. Kies, R. Gassilloud, N. Rambal, V. Lapras, M.-P. Samson, C. Tallaron, C. Tabone, B. Prévitali, D. Barge, A. Ayres, L. Pasini, P. Besombes, F. Andrieu, P. Batude, T. Skotnicki, M. Vinet, **International Symposium on VLSI Technology, Systems and Applications**, 2017, présentation orale.
- **Key process steps for high performance and reliable 3D Sequential Integration**, C.-M. V. Lu, F. Deprat, C. Fenouillet-Beranger, P. Batude, X. Garros, A. Tsiara, C. Leroux, R. Gassilloud, D. Nouguié, D. Ney, X. Federspiel, P. Besombes, A. Toffoli, G. Romano, N. Rambal, V. Delaye, D. Barge, M.-P. Samson, B. Prévitali, C. Tabone, L. Pasini, L. Brunet, F. Andrieu, J. Micoud, T. Skotnicki, M. Vinet, **IEEE Symposium on VLSI Technology**, 2017, présentation orale.

- **Co-auteur**

- **New insights on bottom layer thermal stability and laser annealing promises for high performance 3D VLSI**, C. Fenouillet-Beranger, B.Mathieu, B. Previtali, M-P. Samson, N.Rambal, V. Benevent, S. Kerdiles, J-P. Barnes, D. Barge, P. Besson, R. Kachtouli, P. Batude, M. Cassé, X. Garros, A. Laurent, F. Ne-mouchi, K. Huet, I. Toqué-Trésonne, D.Lafond, H. Dansas, F. Aussenac, G. Druais, P. Perreau, E. Richard, S. Chhun, E. Petitprez, N. Guillot, F. Deprat, L. Pasini, L. Brunet, V. Lu, C. Reita, M. Vinet, **IEEE International Electron Devices Meeting**, 2014, papier contributeur.
- **Monolithic 3D Integration: a powerful alternative to classical 2D Scaling**, M. Vinet, P. Batude, C. Fenouillet-Beranger, F. Clermidy, L. Brunet, O. Rozeau, JM Hartmann, O. Billoint, G. Cibrario, B. Previtali, C. Tabone, B. Sklenard, O. Turkyilmaz, F Ponthenier, N. Rambal, MP. Samson, F. Deprat, V. Lu, L. Pasini, S. Thuries, H. Sarhan, J-E. Michallet, O. Faynot, **IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference**, 2014, papier invité.
- **3DVLSI with CoolCube process: An alternative path to scaling**, P. Batude, C. Fenouillet-Beranger, L. Pasini, V. Lu, F. Deprat, L. Brunet, B. Sklenard, F. Piegas-Luce, M. Cassé, B. Mathieu, O. Billoint, G. Cibrario, O. Turkyilmaz, H. Sarhan, S. Thuries, L. Hutin, S. Sollier, J. Widiez, L. Hortemel C. Tabone, M-P Samson, B. Previtali, N. Rambal, F. Ponthenier, J. Mazurier, R. Beneyton, M. Bidaud, E. Josse, E. Petitprez, O. Rozeau, M. Rivoire, C. Euvard-Colnat, A. Seignard, F. Fournel, L. Benaissa, P. Coudrain, P. Leduc, J-M. Hartmann, P. Besson, S. Kerdiles, C. Bout, F. Nemouchi, A. Royer, C. Agraffeil, G. Ghibaud, T. Signamarcheix, M. Haond, F. Clermidy, O. Faynot, M. Vinet, **IEEE Symposium on VLSI Technology**, 2015, papier invité.
- **Recent advances In 3D VLSI integration**, C. Fenouillet-Beranger, P. Batude, L. Brunet, V. Mazzocchi, C-M. V. Lu, F. Deprat, J. Micout, M-P. Samson, B. Previtali, P. Besombes, N. Rambal, F. Andrieu, O. Billoint, M. Brocard, S. Thuries, G. Cibrario, M. Vinet, **International Conference on IC Design and Technology**, 2016, papier invite.
- **Impact of low thermal processes on reliability of high-k/metal gate stacks**, A. Tsiara, X. Garros, C-M. V. Lu, C. Fenouillet-Beranger, **Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena**, 35, 01A114, 2017, papier journal.
- **Recent advances in low temperature process In view of 3D VLSI integration**, C. Fenouillet-Beranger, P. Batude, L. Brunet, V. Mazzocchi, C-M.V. Lu, F. Deprat, J. Micout, M-P. Samson, B. Previtali, P. Besombes, N. Rambal, V. Lapras, F. Andrieu, O. Billoint, M. Brocard, S. Thuries, G. Cibrario, P. Acosta-Alba, B. Mathieu, S. Kerdilès, F. Nemouchi, C. Arvet, P. Besson, V. Loup, R. Gassilloud, X. Garros, C. Leroux, V. Beugin, C. Guerin, D. Benoit, L. Pasini, J-M. Hartmann, M. Vinet, **IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference**, 2016, papier invité.
- **First demonstration of a CMOS over CMOS 3D VLSI CoolCube™ integration on 300mm wafers**, L. Brunet, P. Batude, C. Fenouillet-Beranger, P. Besombes, L. Hortemel, F. Ponthenier, B. Previtali, C. Tabone, A.Royer, C. Agraffeil, C. Euvard-Colnat, A. Seignard, C. Morales, F. Fournel, L. Benaissa, T. Signamarcheix, P. Besson, M. Jourdan, R. Kachtouli, V. Benevent, J.-M.

Hartmann, C. Comboroure, N. Allouti, N. Posseme, C. Vizioz, C. Arvet, S. Barnola, S. Kerdiles, L. Baud, L. Pasini, C.-M. V. Lu, F. Deprat, A. Toffoli, G. Romano, C. Guedj, V. Delaye, F. Boeuf, O. Faynot, M. Vinet, **IEEE Symposium on VLSI Technology**, 2016, papier contributeur.

◦ **Opportunities brought by sequential 3D CoolCube™ integration**, M. Vinet, P. Batude, C. Fenouillet-Beranger, L. Brunet, V. Mazzocchi, C.-M. V. Lu, F. Deprat, J. Micout, B. Previtali, P. Besombes, N. Rambal, F. Andrieu, O. Billoint, M. Brocard, S. Thuries, G. Berhault, C. Lopes Dos Santos, G. Cibrario, F. Clermidy, D. Gitlin, O. Faynot, **European Solid-State Device Research Conference**, 2016, papier invité.

◦ **First SOI Tunnel FETs with Low-Temperature Process**, C. Diaz Llorente, C. Le Royer, C.-M. V. Lu, P. Batude, C. Fenouillet-Beranger, S. Martinie, F. Allain, S. Cristoloveanu, G. Ghibaud, M. Vinet, **Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon**, 2017, papier contributeur.

◦ **Transistor Temperature Deviation Analysis in Monolithic 3D Standard Cells**, M. Brocard, B. Mathieu, J-P. Colonna, C. Lopes Dos Santos, C. Fenouillet-Beranger, C.-M. V. Lu, G. Cibrario, L. Brunet, P. Batude, F. Andrieu, S. Thuries, O. Billoint, **IEEE Computer Society Annual Symposium on VLSI**, 2017, papier contributeur soumis.

Fabrication de CMOS à basse température pour l'intégration 3D séquentielle

Alors que la miniaturisation des transistors suivant la loi de Moore semble ralentir dû à des limites physique, technologique et économique, il devient essentiel de trouver des alternatives afin de répondre à la demande croissante en électronique : informatique et télécommunication, objets intelligents et interconnectés, domaine médical et biologique... En cela, l'utilisation de la troisième dimension, par opposition à la fabrication planaire de composants électrique, semble être une option prometteuse. L'intégration 3D permet en effet d'incorporer plus de composants sur une même surface en les empilant à un coût technologique et économique plus faibles que celui de la miniaturisation. En particulier, l'intégration séquentielle ou CoolCube™ au CEA-Leti permet de profiter pleinement de la troisième dimension en fabriquant successivement les uns sur les autres chaque étage d'une puce, permettant un alignement optimal des transistors unitaires à chaque niveau. Néanmoins, plusieurs verrous technologiques particuliers à l'intégration 3D Séquentielle doivent alors être levés.

Dans ce manuscrit, nous nous intéresserons à la réduction du budget thermique pour la fabrication des transistors supérieurs, nécessaire afin de ne pas endommager les étages inférieurs lors de la réalisation des composants sus-jacents. Nous commencerons par définir le budget thermique maximal afin de ne pas dégrader les couches inférieures avant d'identifier les briques technologiques impactées lors de la fabrication d'un transistor. Nous verrons alors dans ce manuscrit qu'il sera non seulement nécessaire d'étudier de nouveaux matériaux, mais aussi de nouveaux procédés voire de nouvelles techniques de recuit. Plus particulièrement, nous évaluerons tout d'abord l'utilisation des diélectriques low-k comme espaceurs de grille permettant notamment d'améliorer les performances dynamiques des composants. Ensuite, nous présenterons différentes stratégies de préparation de surface et de croissance épitaxiale à basse température pour la réalisation des sources et drains surélevés. Enfin, nous étudierons l'impact d'un budget thermique faible ainsi que de nouvelles techniques de recuits micro-onde et laser sur les propriétés de l'empilement de grille. Nous verrons en particulier que la difficulté principale d'une intégration à bas budget thermique est l'obtention d'une bonne fiabilité des transistors. Toute cette étude nous permettra alors de proposer des solutions à l'intégration d'un transistor à un bas budget thermique compatible avec l'intégration 3D Séquentielle.

Low thermal budget CMOS processing for 3D Sequential Integration

As the scaling of transistors following Moore's law seems to slow down due to physical, technological and economical barriers, it becomes mandatory to find alternatives to cope with the increasing demand in electronics: computing and telecommunication, smart and interconnected objects, medical and biological fields... To that end, the use of the third dimension, in opposition to the planar processing of electronical devices, appears to be a promising option. Indeed, 3D integration allows incorporating more devices per area by stacking them at a lower technological and economical cost than scaling. More specifically, 3D sequential or CoolCube™ at CEA-Leti allows benefiting fully from the third dimension by processing successively one on top of each other each level of a die, allowing an optimal alignment of single transistors at each layer. However, several technological barriers specific to 3D Sequential Integration need then to be alleviated.

In this work, we will study the reduction of thermal budget for the transistors fabrication, which is required to not damage bottom levels during the processing of top devices. First, we will define the maximal thermal budget in order not to degrade bottom layers prior to identifying the technological modules impacted during the fabrication of a transistor. We will then see in this work that not only new materials need to be studied, but also new processes and new annealing techniques. Specifically, we will first evaluate the use of low-k dielectrics as gate offset spacers, allowing the improvement of devices dynamic performance. Then we will present different strategies of surface preparation and epitaxial growth at low temperature for the formation of raised sources and drains. Finally, we will study the impact of a low thermal budget process flow along with novel microwaves and laser annealing techniques on the gate stack properties. In particular, we will see that the biggest challenge in a low thermal budget integration is to get a good reliability of transistors. This study leads to a proposed low thermal budget process flow for transistor fabrication compatible with 3D Sequential Integration.