



HAL
open science

Intégration de matériaux à forte permittivité électrique (High-k) dans les mémoires non-volatiles pour les générations sub-45nm

Marc Bocquet

► **To cite this version:**

Marc Bocquet. Intégration de matériaux à forte permittivité électrique (High-k) dans les mémoires non-volatiles pour les générations sub-45nm. Micro et nanotechnologies/Microélectronique. Institut National Polytechnique de Grenoble - INPG, 2009. Français. NNT: . tel-00559617

HAL Id: tel-00559617

<https://theses.hal.science/tel-00559617>

Submitted on 26 Jan 2011

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.

INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque :

|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|

THESE

pour obtenir le grade de
Docteur de L'Institut Polytechnique de Grenoble
Spécialité : Micro et Nano Electronique
préparée au laboratoire : **CEA-Léti Minatec & IMEP-LAHC**
dans le cadre de l'**Ecole Doctorale *Electronique, Electrotechnique, Automatique***
et Traitement du Signal

présentée et soutenue publiquement

par

Marc BOCQUET

le 24 Novembre 2009

**Intégration de matériaux à forte permittivité
électrique (*High- κ*) dans les mémoires
non-volatiles pour les générations sub-45nm**

Directeur de Thèse : Georges PANANAKAKIS

Encadrant de Thèse : Gabriel MOLAS

JURY

Pr. GHIBAUDO Gérard	Président
Pr. BOUCHAKOUR Rachid	Rapporteur
Pr. SOUIFI Abdelkader	Rapporteur
Dr. MAURELLI Alfonso	Examineur
Pr. PANANAKAKIS Georges	Directeur de Thèse
Dr. MOLAS Gabriel	Encadrant de Thèse

*Chercher à connaître n'est souvent qu'apprendre à douter.
Madame Deshoulières*

Remerciements

La thèse étant la fin du cursus scolaire, il me faudrait remercier l'ensemble des enseignants qui ont marqué ma scolarité et même les dames de la garderie en maternelle qui m'ont appris à faire mes premiers avions en papier. Comme vous l'imaginez la liste serait bien trop longue si je devais remercier l'ensemble des gens qui m'ont permis de me construire depuis 26 ans. Aussi pour éviter que ces remerciements deviennent un annuaire téléphonique, je me contenterai des principaux acteurs de ma thèse.

Dans un premier temps je voudrais remercier Monsieur Rachid Bouchakour et Monsieur Abdelkader Souifi de m'avoir fait l'honneur d'être les rapporteurs de ma thèse. J'étais très heureux d'avoir Monsieur Gérard Ghibaudo comme président de jury car ses nombreux conseils, son expertise, sa gentillesse, sa disponibilité et son immense savoir scientifique ont marqué des générations de thésards. Je remercie également Monsieur Alfonso Maurelli pour son expertise industrielle et l'intérêt qu'il a porté à ma thèse en tant qu'examinateur.

Je me dois aussi de remercier Olivier Demolliens, Simon Deleonibus et Barbara De Salvo pour m'avoir permis d'effectuer ma thèse au CEA-Léti, dans un environnement de recherche extraordinaire et un laboratoire dynamique. Je remercie tout particulièrement cette dernière pour ses conseils, son professionnalisme et son franc parler.

Je voudrais aussi remercier chaleureusement Georges Pananakakis, mon directeur de thèse. Nos rendez-vous hebdomadaires n'ont fait que souligner ses qualités : gentillesse, attention, rigueur scientifique inébranlable, théoricien averti, etc et ont été des moments de construction et de mise à plat de mon travail de recherche.

Ce sont des remerciements tout aussi chaleureux que j'adresse à Gabriel Molas, mon encadrant de thèse. J'ai eu beaucoup de chance d'avoir un encadrant avec ces qualités humaines et scientifiques qui a su me guider sur les chemins de la recherche et de mon sujet ; en étant à la fois présent dans les moments difficiles et en retrait quand il le fallait pour que j'avance et que j'apprenne par moi-même, à la fois cool et exigeant. Merci, Gab, je suis fier et heureux d'avoir été ton premier thésard.

Je souhaite maintenant remercier l'ensemble des personnes avec qui j'ai travaillé pendant 3 ans sur les matériaux *High- κ* et les mémoires non-volatiles. Tout d'abord, j'aimerais tirer mon chapeau à Marc Gély : merci de t'être occupé de tous ces lots (14 lots électriques uniquement sur les *High- κ* !), merci pour ta gentillesse, ta disponibilité, ton humour et tout le reste. Merci à François Martin, Helen Grampeix et à Jean-Philippe Colonna, nos

experts matériaux, capables de remonter toutes les pistes d'investigations, et dont la porte est toujours ouverte même pour les questions naïves d'un jeune thésard. Je remercie en particulier Jean-Philou pour nos nombreuses discussions *High- κ* et de son intérêt pour la "carac élec". Merci à Nevine Rochat et à tous les spécialistes de la caractérisation morphologique pour l'énorme travail qu'ils ont effectué.

La caractérisation, notamment électrique, a été très présente dans ma vie de thésard et elle n'aurait pas été la même sans toutes les personnes du LSCE, laboratoire de Fabien Boulanger. Tout d'abord, un grand merci aux professionnels du test systématique qui nous permettent d'analyser des centaines de dispositifs sous plusieurs dizaines de conditions : Alain Toffoli, Fabienne Allain, Rabah Kies et Vincent Vidal. Une petite note particulière pour ce dernier que je suis souvent venu embêter et pour ces excellents conseils voyages. Je remercie Jacque Cluzel et Patrick Grosgeorges pour leur incroyable doigté dans le dépannage des SIAMs. Un grand merci, également, à Denis Blachier et ses programmes de tests de folie sans lesquels la caractérisation des lots mémoires auraient été un enfer. Merci également à Xavier Garros et Matthieu Charbonnier chez qui je venais pleurer quand mes courbes C-V étaient quelques peu complexes à analyser. Merci à Philippe Blaise pour ses simulations atomistiques de folie qui m'ont permis d'y voir plus clair dans l'alumine et le nitrure. Merci également à François De Crecy pour son plan d'expérience et notre rêve à tous de fitting automatique ! J'en profite pour m'excuser auprès de Charles Leroux de l'avoir maintes et maintes fois dérangé lors de nos longues discussions "mémoire" avec Etienne. Je salue aussi Gilles Reibold pour les moments partagés au SISC. J'ai aussi une pensée émue pour mes homologues thésards et post-doc de tests, Virginie, Giovanni, Jean et Andrea.

Enfin, un merci tout particulier aux trois personnes du LSCE avec lesquelles j'ai le plus interagi : Etienne Nowak, Elisa Vianello et Luca Perniola. J'ai eu énormément de plaisir à travailler avec vous, vous m'avez beaucoup apporté à la fois sur le plan scientifique et personnel.

Pendant 3 ans, j'ai aussi passé beaucoup de temps à Phelma en tant que moniteur. Ainsi, j'aimerais remercier toute l'équipe enseignante qui m'a accueilli pour tout ce qu'ils apportent au quotidien aux étudiants. Donc merci à Lionel Bastard, Irina Ionica, Fanny Poinssotte, Luc Bouro, Simon Hemour, Anne Vilcot, Patrice Petitclair et Eric. Je remercie tout particulièrement Laurent Aubart et Nathalie Mathieu qui ont été très présents dans mon activité de monitorat et d'étudiant à l'ENSERG. Et aussi, un grand merci aux étudiants que j'ai pu côtoyer. Mes TP et ateliers étaient une petite bouffée d'air frais dans mon travail de recherche, ils ont grandement contribué à mon équilibre durant ma thèse.

Je tiens aussi à remercier les thésards et permanents de l'IMEP pour leur accueil toujours chaleureux malgré la rareté de nos contacts : Leily, Loan, Tuan, Wipa, Simon, Claudio, Quentin et Raphaël.

Ensuite, j'aimerais parler des gens que j'ai côtoyés pendant 3 ans sans travailler vraiment avec eux mais qui ont fait, que tous les matins, j'étais content de venir au travail. Entre deux mesures, j'ai pu partager de très bons moments avec eux : des combats de savate, des croquages de pommes, des tablettes de chocolat, des balistos, des montages photos, des LaserGames, des origamis, des road trip, des soirées "Dude, Where's My Car ?", des atsushows, des foots endiablés le mardi, des séances de piscine le mercredi, des Comboire roudoudou le jeudi, des pots le vendredi, des "workshop" sur les couches minces aux 7laux, les tommatons "made in 4123", les sorties vélos sous la neige, les roses des sables, les sorties rafting, les soirées carac en température, les valse viennoises... Bref, j'ai passé de très bons moments durant cette thèse avec les gens qui travaillaient sur ma thématique, mais aussi avec les autres.

Je commencerai avec les deux personnes qui m'ont accompagné et m'ont supporté au quotidien durant des pauses café ou des goûters mais aussi durant l'écriture de papiers ou de modèles et qui sont devenus des amis, Vince Barral et Carine Jahan. Merci à vous deux pour tous les moments passés ensemble et tout ce que vous m'avez apporté.

Ensuite, je souhaite remercier tous les gens du labo mémoires de Barbara De Salvo qui m'ont accueilli pendant trois ans : Alain qui court comme un lion, Alexandra qui m'aura supporté pendant 6 mois, Carine qui est dingue de sucre, Eric qui fournit tout le CEA en ... dots, Gab qui mixe et joue du pipo comme un dieu, Guillaume avec qui j'ai fait le tour de la Californie (5000km en 7jours), Jean-François qui est le gérant du café de 9h, Julien qui est toujours de bon conseil (notamment sur les *High-κ*), Ludovic qui a un humour à tout renverser, Marc qui adore l'expérimentation, Micael avec qui je ne suis toujours pas allé surfer, Sandrine qui n'est pas du labo mais avec qui j'ai couru et bu quelques cafés, Stéphanie qui est partie au Liten (Steph, tu nous manques), Venera qui est très discrète, Veronique qui a toujours le sourire, Christelle qui est arrivée il n'y a pas longtemps mais qui s'est très vite intégrée et Cyril qui m'a soufflé des idées farfelues mais toujours pertinentes. A vous tous, merci pour votre accueil, votre gentillesse, votre bonne humeur. La bonne ambiance qui règne dans ce laboratoire va beaucoup me manquer.

Au passage, j'en profite pour saluer nos secrétaires : Brigitte et Sabine. Ainsi que l'homme le plus déjanté de tout le 4123, voir même du Léti, merci Pompon pour tes délires, tes conseils "Thermaux" et ton sérieux (si si, je vous jure).

Je souhaite aussi saluer mes collègues du LDI laboratoire d'Olivier Faynot que j'ai côtoyés au sein du feu LNDE : Carlo, Cyrille, Maud, Nathalie, Thomas et Virginie, avec une mention particulière pour Corine qui était un peu notre maman à tous, et pour mes compagnons de glisse : Bernard et Laurent avec qui j'ai beaucoup progressé en surf. Un petit clin d'œil à Christel et Carole (qui n'est pas du LDI d'ailleurs mais tant pis) pour m'avoir tant taquiné avec Carine. Je m'excuse au près de Marie-Pierre, encore des remerciements de thèse où elle sera remerciée pour ses petits gâteaux si délicieux. Je remercie Georges, roi de la carac, pour m'avoir supporté durant notre tour d'Écosse express et partagé avec moi ces paysages et ces B&B absolument magnifiques. Je salue Claude pour sa générosité et son humour pince-sans-rire, sans oublier Bernard avec lequel ils forment un duo truculent ! Une note spéciale pour Thierry dont les éloges de son ex-

thésards sont telles que cela mérite d'être soulignées. Prix spécial "Tout pour la vanne" pour François et Olivier pour leur humour dévastateur. Je vous remercie tous pour ces repas à 30 le midi qui faisaient l'âme du labo et qui partaient bien souvent dans tous les sens.

Je félicite les thésards qui ont terminé, j'encourage chaudement ceux qui n'ont pas encore fini. Mais je vous remercie tous pour nos délires (surtout d'après 19h), nos sorties, vos conseils et tout ce que l'on a partagé. *Best regards* à Estelle, Jérôme, Kiichi, Milène, Paul-Henry, Romain, Sophie et Stéphane. Je remercie chaleureusement les organisateurs de soirées et sorties pour avoir créé une vie entre les thésards pour que l'on soit un peu plus que de simples collègues de travail. Notamment les matchs de foot le mardi – petite dédicace à Andres le dribbleur fou, les soirées "films de merde" – je m'incline devant l'immense culture de Techtolouis, les soirées Time's Up – prix spécial pour Fred et son mime de Truman Streckfus Persons, les soirées JdR – prix du meilleur meujeu à Alex, les soirées déguisées – prix de la plus LNDE des grenobloises à Anne-Claire, et les soirées patinoire – petit clin d'œil à Jyotshna pour nos tours de piste. J'aimerais aussi remercier Cécilia pour sa gentillesse et son attention envers tout le monde, et Atsushi, le plus extraordinaire des japonais, à qui j'ai une promesse à tenir. Enfin, un dernier prix à la plus jolie et la plus déjantée des filles du 4123 même si elle n'en a pas l'air, la reine du fou-rire : Émilie.

Une petite note particulière pour mes confrères de galère, les thésards de la même promo que moi : Jean-Paul, Michaël, Perrine. Je vous admire beaucoup tous les trois pour vos qualités scientifiques et humaines. Pendant 3 ans, j'ai essayé d'être à votre hauteur et heureux de vous côtoyer ; en vous souhaitant bon vent pour la suite et que nos routes se recroisent.

Je souhaite aussi associer mes amis de Chambéry dans ces remerciements car ils m'ont permis de me décompresser presque tous les week-end. Cédric, Christophe, Stéphane et Steve, je me demande comment vous arrivez à me supporter chaque semaine. Amédée et Bruno, votre humour, si particulier à chacun, nous manque quand vous n'êtes pas là. Et Julien et Loïc, j'aimerais vous voir plus souvent ! Merci à vous tous pour tous les bons moments que j'ai passés avec vous depuis bien longtemps.

Enfin, je remercie ma famille d'avoir toujours été près de moi pour me soutenir. En particulier mes parents et ma soeur, merci pour tout ce que vous avez fait pour moi et ce depuis toujours. J'ai eu beaucoup de chance de vous avoir. Et j'en profite pour embrasser ma douce Emma.

Table des matières

1	Présentation des mémoires non-volatiles	25
1.1	Positionnement des mémoires dans l'industrie des SCs	27
1.1.1	Contexte économique	27
1.1.2	Classification des mémoires	28
1.2	Fonctionnement et architecture des mémoires Flash	31
1.2.1	Présentation et fonctionnement de la cellule mémoire	31
1.2.1.1	Écriture	33
1.2.1.2	Effacement	34
1.2.2	Architectures des mémoires Flash	35
1.2.3	Les cellules mémoires multi-bits	37
1.2.4	Évolution technologique	39
1.3	Limitations des mémoires Flash et solutions envisagées	40
1.3.1	Problèmes liés à la réduction de la distance entre les cellules	41
1.3.1.1	Suppression des extensions latérales autour des cellules	41
1.3.1.2	Couplage entre cellules adjacentes	42
1.3.2	Problèmes liés à la réduction des cellules mémoires	43
1.3.2.1	Effets de canaux courts	43
1.3.2.2	Dégradation de l'oxyde <i>tunnel</i>	44
1.3.2.3	Nombre d'électrons pour coder un état	45
1.3.3	Intégration 3D des plans mémoires	46
1.3.4	Intégration de matériaux <i>High-κ</i> comme diélectrique d'interpoly	47
1.4	Les mémoires TANOS	49
1.4.1	Présentation de la technologie TANOS	49
1.4.2	Difficultés des mémoires TANOS et leurs évolutions	52
1.4.2.1	L'ingénierie de bandes de l'empilement <i>tunnel</i>	52
1.4.2.2	L'ingénierie de la couche de l'empilement de grille	53
1.5	Conclusion	56
2	Modélisation du courant et du piégeage dans les diélectriques	57
2.1	Conduction électronique dans les diélectriques	59
2.1.1	Conduction <i>tunnel</i> direct et Fowler-Nordheim	60
2.1.1.1	Expressions générales	60
2.1.1.2	Etats électroniques liés et continuum d'états	62
2.1.1.3	Expression des courants 2D	64
2.1.1.4	Expression des courants 3D	65

2.1.1.5	Courant d'émission d'une couche de piégeage	66
2.1.1.6	Calcul de la transparence	67
2.1.2	Modes de conduction assistée par des défauts	69
2.1.2.1	Courant Poole-Frenkel	70
2.1.2.2	Courant Poole	72
2.1.2.3	Identification des courants Poole et Poole-Frenkel	73
2.1.2.4	Courant TAT : Trap Assisted Tunneling	73
2.2	Mécanismes de piégeage dans les diélectriques	75
2.2.1	Les différentes approches pour la modélisation d'un piège	75
2.2.2	Modèle d'états indépendants	76
2.2.3	Aspect amphotérique des pièges	77
2.2.4	Modélisation des coefficients de piégeage	78
2.2.4.1	Échanges entre la BC et le piège du diélectrique	79
2.2.4.2	Échanges entre la BC du silicium et le piège	81
2.3	Modèles de piégeage pour les mémoires de type SONOS	83
2.3.1	État de l'art de quelques modèles SONOS	83
2.3.2	Modèle 1 : prise en compte de pièges amphotériques	86
2.3.3	Modèle 2 : prise en compte du transport dans la couche de piégeage	88
2.4	Conclusion	91
3	Etude des matériaux <i>High-κ</i>	93
3.1	Présentation des matériaux <i>High-κ</i>	95
3.1.1	Introduction générale des matériaux <i>High-κ</i>	95
3.1.2	Procédés de fabrication	97
3.2	Alliages à base de HfO_2 et Al_2O_3 : Les aluminates d'hafnium	97
3.2.1	Généralités sur les $HfAlO$	97
3.2.1.1	Procédés de fabrication	97
3.2.1.2	Stoechiométrie, gap et cristallinité	99
3.2.2	Étude des charges fixes dans les aluminates d'hafnium	102
3.2.2.1	Effet de la concentration en Hf	102
3.2.2.2	Localisation des charges fixes dans les aluminates	103
3.2.3	Courant de fuite dans les $HfAlO$: épaisseurs et natures	105
3.2.3.1	Influence de l'épaisseur des HTO dans un tricouche	106
3.2.3.2	Problématique du piégeage et mesure de courant	108
3.2.3.3	Influence de la couche <i>High-κ</i> à faible et à fort champ	109
3.2.3.4	Modes de conduction et d'activation en température	111
3.2.4	Propriétés de piégeage des $HfAlO$	115
3.3	Impact de la nitruration sur les $HfAlO$	118
3.3.1	Impact sur les propriétés physico-chimiques	119
3.3.1.1	Incorporation de l'azote dans le film	119
3.3.1.2	Influence sur le gap optique	120
3.3.1.3	Cristallinité	121
3.3.2	Effets du recuit NH_3 sur les propriétés de piégeage	123
3.4	Impact du traitement thermique sur l'alumine	126
3.4.1	Les propriétés physico-chimiques	127

3.4.1.1	Étude du stress, de la densité et de l'épaisseur	127
3.4.1.2	Mesures FTIR-ATR	128
3.4.1.3	Concentration d'hydrogène dans l'empilement mémoire	130
3.4.1.4	Simulation atomistique	131
3.4.2	Étude des caractéristiques électriques	134
3.4.2.1	Étude du courant de fuite	134
3.4.2.2	Étude des propriétés de piégeage	135
3.5	Silicate d'hafnium : <i>HfSiON</i>	137
3.5.1	Analyses matériaux	137
3.5.2	Étude du courant de fuite	140
3.5.2.1	Structure bicouche <i>SiO₂/HfSiON</i>	140
3.5.2.2	Structure tricouche <i>SiO₂/HfSiON/HTO</i>	142
3.6	Évaluation des performances mémoires	144
3.6.1	Simulation de la fenêtre de programmation	144
3.6.2	Simulation de la rétention	148
3.7	Conclusion	151
4	Intégration des <i>High-κ</i> dans les mémoires non-volatiles	155
4.1	<i>High-κ</i> dans les mémoires à nanocristaux	157
4.1.1	Procédé de fabrication des mémoires à nanocristaux	157
4.1.2	Caractéristiques de programmation des mémoires à nanocristaux	159
4.1.3	Caractéristiques de rétention et d'endurance	160
4.2	Mémoire à base de nitrure avec oxyde bloquant <i>High-κ</i>	162
4.2.1	Impact de l'épaisseur de l'oxyde <i>tunnel</i>	163
4.2.2	Impact de l'oxyde bloquant	165
4.2.2.1	Caractéristiques d'écriture et d'effacement	165
4.2.2.2	Caractéristiques d'endurance	169
4.2.2.3	Caractéristiques de rétention	170
4.2.3	Impact de l'épaisseur de l'alumine et du nitrure	171
4.2.3.1	Caractéristiques d'écriture et d'effacement	172
4.2.3.2	Caractéristiques d'endurance	176
4.2.3.3	Caractéristiques de rétention	178
4.2.4	Impact des traitements thermiques sur l'alumine	184
4.2.4.1	Caractéristiques d'écriture et d'effacement	185
4.2.4.2	Caractéristiques de rétention	187
4.2.5	Impact de la grille de contrôle	188
4.2.5.1	Étude des grilles en polysilicium	188
4.2.5.2	Étude des grilles métalliques en TiN, TaN ou TaAlN	190
4.3	Mémoire à base de nitrure avec oxyde <i>tunnel</i> en <i>SiO₂/HfSiON</i>	195
4.3.1	Caractéristiques d'écriture et d'effacement	196
4.3.2	Caractéristiques de rétention	199
4.4	Conclusion	201
	Conclusions et perspectives	201

A Procédé de fabrication d'une cellule mémoire réalisée au Léti	207
--	------------

Table des figures

1.1	Place de l'industrie de la microélectronique dans le monde économique. . .	28
1.2	Classifications des mémoires à semi-conducteurs.	28
1.3	Évolution des revenus du marché des mémoires à technologie MOS	30
1.4	Schémas d'une cellule mémoire Flash	31
1.5	Caractéristiques et schéma électrique d'une cellule mémoire Flash.	32
1.6	Modes d'écriture des mémoires Flash.	33
1.7	Modes d'effacement des mémoires Flash.	34
1.8	Architecture NOR.	35
1.9	Architecture NAND.	36
1.10	Évolution du marché des mémoires Flash NAND et Flash NOR [ICI09]. . .	37
1.11	Multi-niveaux : 7 niveaux de programmation pour une mémoire NAND. . .	38
1.12	Bit miroir : localisation physique de la charge.	38
1.13	Évolution des technologies mémoires NAND.	39
1.14	Problématiques de la suppression des flancs latéraux de la grille de contrôle.	41
1.15	Schéma électrique du couplage entre cellules NAND.	42
1.16	ΔV_T induit par la charge d'une cellule adjacente	42
1.17	Schéma d'une cellule FinFlash [Jah08].	43
1.18	Amélioration de l'injection grâce à une géométrie cylindrique	43
1.19	Schémas de la perte d'électrons dans la zone de stockage	44
1.20	Mesures de courant à travers un oxyde de silicium ou un bicouche SiO_2/HfO_2 .	45
1.21	Nombre d'électrons pour coder un bit en fonction du nœud technologique.	46
1.22	Photographies MET de cellules Flash d'une intégration séquentielle.	46
1.23	Cellules Flash d'une approche monolithique.	47
1.24	Exemple d'intégration de <i>High-κ</i> dans une cellule mémoire à grille flottante.	48
1.25	Photographies TEM d'une mémoire TANOS	50
1.26	Caractéristiques d'écriture et d'effacement typiques d'une cellule TANOS. .	50
1.27	Schémas de bandes d'une mémoire BE-SONOS en effacement et en rétention.	52
1.28	Caractéristiques d'effacement de mémoires BE-SONOS, MANOS et MA- NONOS.	53
1.29	Exemple d'ingénierie de bandes d'une TANOS	55
2.1	Principaux modes de conduction à travers un diélectrique.	59
2.2	Mécanismes de conduction (a) <i>tunnel</i> direct et (b) Fowler-Nordheim. . . .	61
2.3	Diagrammes de bandes pour trois polarisations	62
2.4	Etats liés et quasi-continuum d'états électroniques.	63

2.5	Diagramme de bandes d'une structure ONO pouvant émettre des électrons.	66
2.6	Vitesse moyenne pondérée pour des courants partant du Si_3N_4 .	67
2.8	Schéma de la conduction Poole-Frenkel	70
2.9	Schéma de la conduction Poole	72
2.10	Schémas des conductions <i>tunnel</i> assistées par pièges : TAT.	74
2.11	Calcul ab initio de la densité d'états du Si_3N_4	75
2.12	Schéma de bandes des pièges.	76
2.13	Schémas des différents passages entre les trois états des pièges amphotériques.	77
2.14	Échanges entre le piège et le silicium ou le diélectrique.	78
2.15	Schéma des échanges entre la bande de conduction du silicium et les pièges.	81
2.16	Schémas des mécanismes considérés dans le modèle 1.	86
2.17	Schéma des mécanismes considérés dans le modèle 2.	89
3.1	Rappels sur les <i>High-κ</i> : Constante diélectrique, E_G et ΔE_C .	96
3.2	Résumé des étapes d'un cycle d'ALCVD	98
3.3	Concentration et gap optique des <i>HfAlO</i> , impact de la concentration en <i>Hf</i>	100
3.4	Photographies MET d' <i>HfAlO</i> , pour différentes compositions [Mol06].	100
3.5	Impact des charges fixes négatives dans l'oxyde bloquant.	102
3.6	Structure capacitive tricouche étudiée pour évaluer les charges fixes.	102
3.7	Caractéristiques $C_G - V_G$, charges fixes dans les <i>HfAlO</i>	103
3.8	Localisation possible des charges dans une structure capacitive tricouche.	103
3.9	Localisation des charges fixes dans des tricouches <i>High-κ</i>	104
3.10	Schémas de cellules mémoires à grille flottante et à couche de piégeage.	105
3.11	Structure capacitive tricouche étudiée pour évaluer les courants de fuite.	106
3.12	Caractéristiques $J_G - V_G/EOT$, impact de l'épaisseur du <i>HTO</i> côté cathode	106
3.13	Caractéristiques $J_G - V_G/EOT$, impact de l'épaisseur du <i>HTO</i> côté anode	107
3.14	Caractéristiques $J_G - V_G/EOT$, impact du piégeage sur la mesure	108
3.15	Schéma de bande d'un tricouche pour différentes permittivités du <i>High-κ</i> .	109
3.16	Caractéristiques $J_G - V_G/EOT$, impact de la nature du <i>HfAlO</i>	109
3.17	Caractéristiques $J_G - V_G/EOT$, impact de la nature du <i>High-κ</i>	110
3.18	Caractéristiques $J_G - V_G/EOT$, impact de l'épaisseur du <i>HfAlO</i>	111
3.19	Activation en température des tricouches à base de <i>HfAlO</i>	112
3.20	Énergies d'activation en fonction des différentes compositions de <i>HfAlO</i>	112
3.21	Diagrammes de Hill des empilements <i>HTO-4nm/ HfAlO-3nm/HTO-4nm</i> .	112
3.22	Structure étudiée pour la dépendance en épaisseur et en température.	113
3.23	Activation en température des <i>HfAlO</i> , impact de l'épaisseur.	114
3.24	Structure capacitive tricouche étudiée pour le piégeage des <i>HfAlO</i> .	115
3.25	Caractéristiques de piégeage dans les <i>High-κ</i> à $V_G = 9Vm$ et $12V$	115
3.26	Caractéristiques de piégeage dans les <i>High-κ</i> à $V_G/EOT = 10MV/cm$	116
3.27	Localisation de la charge piégée dans des tricouches <i>High-κ</i>	116
3.28	Simulations du piégeage des empilements <i>HTO/ HfAlO/ HTO</i> .	117
3.29	Rétention de la charge piégée, impact de la concentration en <i>Hf</i>	117
3.30	Profile Auger de la concentration en nitrure dans les <i>HfAlO</i>	119
3.31	Gap optique des <i>HfAlO</i> , impact de la nitruration	120
3.32	Spectres ATR des <i>HfAlO</i> (a) après dépôt et (b) après nitruration.	121

3.33	Photographies MET d' HfO_2 et d' $HfAlO$ -1:2 nitrurés	122
3.34	Incorporation en azote en fonction de la concentration en Hf	123
3.35	Structure d'étude du piégeage des $HfAlO$ nitruré.	123
3.36	Caractéristiques de piégeage dans l' $HfAlO$ et l' HfO_2 nitrurés	124
3.37	Rétention des charges piégées dans un $High-\kappa$ nitruré	124
3.38	Localisation de la charge dans des tricouches $High-\kappa$ nitrurés	125
3.39	Stress, épaisseur et densité de l' Al_2O_3 en fonction du recuit post-déposition	128
3.40	Phase et nature cristalline de l' Al_2O_3 , impact du recuit post-déposition . .	129
3.41	Mesures MIR réalisées sur différents échantillons Al_2O_3 pour différents recuits	130
3.42	Mesures SIMS d'hydrogène dans un empilement SANOS suivant les PDA .	130
3.43	Diagramme d'états de l'alumine [San09].	131
3.44	Al_2O_3 en phase γ obtenue à partir d'une structure spinelle modifiée	132
3.45	Densité d'états dans l'alumine obtenu par simulation atomistique	133
3.46	Caractéristiques $J_G - V_G$ d'alumine cristalline, conduction Pool-Frenkel . .	135
3.47	Caractéristiques de piégeage d'alumine ayant subi différents recuits	136
3.48	Gap optique et permittivité des $HfSiO$, impact de la concentration en Hf	138
3.49	Gap optique des $HfSiON$, impact de la nitruration	138
3.50	Gap optique et concentration en N des $HfSiON$	139
3.51	Caractéristiques $J_G - V_G/EOT$ d'un bicouche $SiO_2/HfSiON$	141
3.52	Courant traversant l'empilement <i>tunnel</i> , intégré dans une cellule mémoire .	142
3.53	Courant à travers des empilements $SiO_2/HfSiON$ et $SiO_2/HfSiON/HTO$	143
3.54	Diagrammes de bandes d'une mémoire pour un <i>tunnel</i> tricouche et bicouche	144
3.55	Schéma électrique d'une cellule mémoire élémentaire.	145
3.56	Simulations de l'écriture pour des interpolys $High-\kappa$	146
3.57	Simulations de la tension de programmation pour des interpolys en $High-\kappa$	147
3.58	Simulations de la rétention pour des interpolys en $High-\kappa$	149
3.59	Simulations de l'activation en température pour des interpolys en $High-\kappa$. .	150
3.60	Simulations du temps de rétention en fonction de $V_{G_{Prog}}$	151
4.1	Schémas montrant l'intérêt des mémoires à nanocristaux face au SILC. . . .	157
4.2	Schémas des courants dans une mémoire à nanocristaux.	157
4.3	Photographies MET et MEB de mémoires à nanocristaux.	158
4.4	Schémas des variantes technologiques des mémoires à nanocristaux	158
4.5	Caractéristiques d'écriture et d'effacement des mémoires à nanocristaux . .	159
4.6	Diagrammes de bandes $V_G = 0V$ pour une mémoire SONOS et à nanocristaux.	160
4.7	Caractéristiques de rétention des mémoires à nanocristaux de Si.	161
4.8	Caractéristiques d'endurance de mémoires à nanocristaux de Si.	162
4.9	Schéma d'une mémoire à couche de piégeage.	162
4.10	Diagrammes de bandes d'une SONOS pour différents V_G et t_{tun}	163
4.11	Réseaux d'écriture de SONOS intégrant différentes épaisseurs d'oxyde <i>tunnel</i> .	164
4.12	Programmation d'une SONOS en fonction de l'épaisseur d'oxyde <i>tunnel</i> . .	164
4.13	Schémas des dispositifs étudiés pour évaluer l'impact de l'oxyde bloquant. .	165
4.14	Réseaux des caractéristiques de W/E des SONOS, SANOS et SAONOS . .	166
4.15	Caractéristiques W/E, impact de l'oxyde bloquant d'une SANOS	166

4.16	Diagramme de bandes et des mécanismes pris en compte pour la simulation.	167
4.17	Diagramme de bandes côté grille pour un diélectrique Al_2O_3 ou HTO/Al_2O_3 .	167
4.18	Champs dans l'oxyde lors de l'effacement pour différents oxydes bloquants.	168
4.19	Caractéristiques d'endurance, impact de l'oxyde bloquant d'une SANOS . .	169
4.20	Caractéristiques de rétention, impact de l'oxyde bloquant d'une SANOS . .	170
4.21	Caractéristiques de rétention, impact de l'oxyde bloquant d'une SANOS . .	171
4.22	Photographies MET et variantes technologiques des SANOS. [Boc09] . . .	172
4.23	Réseaux des caractéristiques W/E des SANOS, impact des épaisseurs . . .	172
4.24	Caractéristiques W/E, impact de l'épaisseur de Si_3N_4	173
4.25	Caractéristiques d'écriture/effacement pour différentes épaisseurs d'alumine.	174
4.26	Caractéristiques W/E, $V_T - V_G/EOT$	175
4.28	Caractéristiques d'endurance pour différentes épaisseurs d'alumine	177
4.29	Caractéristiques d'endurance pour différentes épaisseurs de nitrure	177
4.30	Caractéristiques de rétention de SANOS en fonction de l'épaisseur du Si_3N_4	179
4.31	Densité de charges piégées après programmation et durant la rétention . .	180
4.32	Séquence de simulation pour la programmation et l'effacement [Via09b]. . .	180
4.33	Schéma de bande lors de la rétention en température pour les SANOS . . .	181
4.34	Courant sortant du nitrure durant la rétention des SANOS	182
4.35	Rétention de mémoires SANOS en fonction de l'épaisseur d' Al_2O_3	183
4.36	TEM d'empilements $Al_2O_3/Si_3N_4/SiO_2$ pour différents PDA sur Al_2O_3 . .	184
4.37	EOT des empilements $Al_2O_3/Si_3N_4/SiO_2$ pour différents PDA sur Al_2O_3 .	185
4.38	Impact du PDA de Al_2O_3 sur l'écriture des SANOS	186
4.39	Impact du PDA de Al_2O_3 sur l'effacement des SANOS	186
4.40	Caractéristiques de rétention des SANOS ayant subi différents recuits. . . .	187
4.41	Diagrammes de bandes pour différentes grilles de contrôle.	188
4.42	Caractéristiques d'effacement pour SONOS avec une grille N+ ou P+. . . .	189
4.43	Caractéristiques d'effacement pour SAONOS avec une grille N+ ou P+. . .	190
4.44	$I_G - V_G$ d' Al_2O_3 -16nm pour différents recuits et différents types de grilles .	192
4.45	Caractéristiques d'écriture et d'effacement pour différents types de grilles. .	193
4.46	Caractéristiques de rétention pour différents types de grilles.	193
4.47	Caractéristiques d'endurance pour différents types de grilles.	194
4.48	Photographies MET de $TiN/Al_2O_3/HTO/Si_3N_4/HfSiON/SiO_2$	196
4.49	Comparaison de la programmation d'une SAONOS et d'une SAONHOS . .	197
4.50	Comparaison de l'endurance d'une SAONOS et d'une SAONHOS	197
4.51	Impacts des différents éléments d'une SAONHOS sur la programmation . .	198
4.52	Rétention pour une SAONHOS et une SAONOS	199
4.53	Diagrammes de bandes d'une SAONHOS et d'une SAONOS à $V_G = 0V$. .	200

Liste des tableaux

1.1	Tableau de l'ITRS 2007 pour les mémoires Flash NOR et NAND.	40
1.2	Tableau de l'ITRS 2007 pour les mémoires SONOS.	49
1.3	Résumé des propriétés des mémoires à grille flottante et à couche de piégeage.	51
1.4	Description d'empilements intégrant des <i>High-κ</i>	53
1.5	Comparaisons des caractéristiques mémoires de différents dispositifs.	54
2.1	Résumé des propriétés des deux modèles.	90
3.1	Propriétés des oxydes métalliques : <i>High-κ</i>	96
3.2	Concentration en <i>Hf</i> , gap, permittivité relative, ΔE_C et ΔE_V des <i>HfAlO</i>	101
3.3	Paramètres physiques des empilements pour la localisation de la charge.	104
3.4	Profondeurs des pièges d'une conduction Poole-Frenkel dans les <i>HfAlO</i>	113
3.5	Gap de l'alumine en fonction du budget thermique.	129
3.6	Résumé des propriétés des <i>HfAlO</i>	151
3.7	Résumé de l'impact de la nitruration sur les <i>HfAlO</i>	152
3.8	Résumé des propriétés de l' Al_2O_3	152
3.9	Résumé des propriétés des <i>HfSiON</i>	153
4.1	Paramètres du modèle utilisés dans la figure 4.15.	168
4.2	Paramètres du modèle utilisés dans les figures 4.30 et 4.31.	179
4.3	Travail de sortie des grilles métalliques étudiées [Als06, Gil09].	191
4.4	Résumé des propriétés des mémoires de type SONOS étudiées	202

Liste des abréviations

Pour des raisons de lisibilité, la signification d'une abréviation ou d'un acronyme n'est rappelé qu'à sa première apparition dans le texte d'un chapitre. Par ailleurs, puisque nous utilisons toujours l'abréviation la plus usuelle, il est fréquent que ce soit le terme anglais qui soit employé, auquel cas nous présentons une traduction.

ALCVD Atomic Layer Chemical Vapor Deposition

ALD Atomic Layer Deposition

Al₂O₃ Alumine

AR-XPS Angle Resolved X-ray Photo-electron Spectroscopy

ATR-FTIR Attenuated Total Reflection Fourier Transform Infrared spectroscopy

BC Bande de conduction

BV Bande de valence

EBT Electron Back Tunnelling

EELS Electron Energy Loss Spectroscopy

EEPROM Electrical Erase Programmable Read Only Memory

IRTF Spectroscopie Infrarouge à Transformée de Fourier

HTO High thermal oxides (oxyde dpos haute temprature)

HfAlO Aluminates d'hafnium

HfO₂ Oxyde d'hafnium

HfSiO Silicate d'hafnium

HfSiON Silicate d'hafnium nitruré

MET Microscopie électronique en transmission

MOCVD Metal-Organic Chemical Vapor Deposition

OHO Oxyde/*High-κ*/Oxyde

ONO Oxyde/Nitruure/Oxyde

PDA Post-Deposition Annealling

PVD Physical Vapor Deposition

resp. respectivement

SAHOS Silicium/Alumine/*High-κ*/Oxyde/Silicium

SANOS Silicium/Alumine/Nitrure/Oxyde/Silicium

SiH₂Cl₂ Dichlorosilane

SILC Stress Induced Leakage Current

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

Si₃N₄ Nitrure de silicium

SiO₂ Oxyde de silicium

SONOS Silicium/Oxyde/Nitrure/Oxyde/Silicium

TEM Transmission Electron Microscopy

TMA Tri-Méthyle Aluminium : *Al(CH₃)₃*

VPD-ICPMS Vapor Phase Chemical Decomposition Inductively Coupled Plasma Mass Spectroscopy

WDXRF Wavelength Dispersive X-Ray Fluorescence

Liste des notations

Nous avons regroupé ci-dessous les principales notations employées dans les différents chapitres du document. Dans la mesure du possible, nous avons tenté de conserver les mêmes notations d'un chapitre à l'autre. Nous présentons tout d'abord une liste générale, puis des listes relatives aux différents chapitres. On notera que seules les notations qui diffèrent de celles précédemment définies seront données dans ces listes. Enfin, certaines notations, apparaissant uniquement de manière ponctuelle, ont été omises.

E	Énergie [J]
F	Champ électrique en [V/m]
J	Densité de courant [A/m ²]
Q_{GF}	Charge dans la grille flottante
V_{GF}	Tension de la grille flottante
V_T	Tension de seuil d'un dispositif MOS
T	Température en [K]
k	Constant de bolzmann [$J \cdot K^{-1}$]
ε_0	Permittivité du vide : $8,85 \cdot 10^{-12}$ [F/m]
ε_r	Permittivité relative [-]
β_{PF}	la constante de Poole-Frenkel
m_e^{Si}	Masse des électron dans le silicium [kg]
m_e^{ox}	Masse des électron dans l'oxyde [kg]
Φ_b	Hauteur de barrière entre la bande de conduction du <i>Si</i> et de l'oxyde [J]
h	Constante de Planck : $6,626 \cdot 10^{-34}$ [$J \cdot s$]
\hbar	Constante réduite de Planck : $\frac{h}{2\pi} = 1,055 \cdot 10^{-34}$ [$J \cdot s$]
f_T	remplissage de l'état d'un piège
f	fonction de distribution des électrons ou des trous.

Introduction Générale

Contexte

Il existe actuellement un très fort engouement pour les applications électroniques portables telles que les netbook, téléphones cellulaires, appareils photos numériques, PDA, clés USB... Ces différents dispositifs sont en train de fusionner pour donner naissance un appareil unique permettant de recevoir des messages électroniques, réaliser des photographies, écouter de la musique, passer des appels téléphoniques, s'orienter grâce à une puce GPS, naviguer sur le web...

Ces nouvelles applications nécessitent de pouvoir traiter rapidement l'information mais avec une faible consommation. Cela est réalisable grâce à l'amélioration continue des performances des dispositifs CMOS logiques et à l'augmentation des capacités mémoire. En effet, la quantité de données à stocker ne cesse croître au fil des années avec l'apparition de nouveaux formats et services, comme les vidéos Haute Définition. Dans ce contexte, les nouveaux produits, en plus de fournir une grande quantité de service, doivent annoncer une bonne ergonomie et une grande autonomie à des prix attractifs.

La solution qui a été largement choisie par l'industrie est d'utiliser les mémoires Flash non-volatiles, qui grâce à des innovations continues permettent de stocker de plus en plus de données. En plus de leur forte densité d'intégration, un intérêt supplémentaire des mémoires Flash est qu'elles permettent de stocker l'information pendant plus de 10 ans, sans aucun apport extérieur d'énergie. Cela constitue un critère majeur pour des applications portables à basse consommation. On voit ainsi apparaître sur la marché de nouveaux produits alliant haute performance et faible consommation, par exemple l'arrivée des produits mémoire SSD réalisés en technologie $32nm$.

Face à la demande croissante du marché pour stocker davantage de données, au plus bas coût et avec une consommation minimale, il est essentiel que les mémoires Flash puissent continuer leur course à la miniaturisation malgré des difficultés physiques et technologiques grandissantes. Cet état de fait explique donc l'engouement actuel de la recherche appliquée pour ce domaine.

Présentation de la thèse

Cette thèse porte sur l'étude de l'intégration des diélectriques à forte permittivité électrique¹, couramment appelé, *High- κ* , dans les mémoires non-volatiles. Nous avons donc étudié leurs effets sur l'ensemble des couches de l'empilement mémoire.

Nous présentons dans ce manuscrit les points majeurs de notre étude selon 4 chapitres.

Le **Chapitre 1** présente tout d'abord une brève description du contexte actuel et les perspectives du marché des mémoires Flash. La partie suivante expose leur fonctionnement et les limites physiques à la poursuite de la réduction des dimensions avant de montrer les solutions qui semblent les plus prometteuses pour les repousser.

Dans le **Chapitre 2** sont d'abord rappelés les modes de transport dans les diélectriques. Puis les différentes approches pour modéliser le piégeage de charge dans ces couches seront présentées. Enfin, en s'appuyant sur ces notions, deux modèles de piégeage pour les mémoires à couche de piégeage en nitrure seront décrits. Ces modèles permettront une meilleure compréhension des mécanismes d'écriture, d'effacement et de rétention, dans les chapitres suivants.

Le **Chapitre 3** consiste en une présentation et une analyse des matériaux *High- κ* . En s'appuyant sur les analyses physico-chimiques réalisées au CEA-Léti, une étude des propriétés électriques (charge fixe, courant de fuite, piégeage...) a été réalisée pour différents *High- κ* (alumine, oxyde d'hafnium, aluminates d'hafnium, silicate d'hafnium). Enfin, en se basant sur ces mesures, le potentiel de ces matériaux a été évalué pour leur intégration en tant que diélectrique d'interpoly dans une cellule mémoires Flash.

L'objet du **Chapitre 4** concerne l'intégration des matériaux *High- κ* dans des cellules mémoires à nanocristaux de silicium ou à couche de piégeage nitrure. Ces matériaux *High- κ* sont utilisés comme oxyde bloquant ou diélectrique d'interpoly et comme oxyde tunnel. L'objectif est de relier les propriétés matériaux et électriques obtenues dans le chapitre précédant avec les performances ces dispositifs mémoires et ainsi de comprendre les mécanismes mis en jeu en vue de l'optimisation de la cellule mémoire. Les résultats sont analysés à partir de modèles de dispositifs mémoire.

Le manuscrit se termine enfin par une conclusion générale résumant les principaux résultats obtenus au cours de l'étude, avant de proposer des perspectives à ce travail.

¹Forte permittivité électrique sous entendue plus forte que celle de la silice qui vaut $3,9\epsilon_0$.

Chapitre 1

Présentation des mémoires non-volatiles

Ce chapitre présente, dans un premier temps, le contexte économique, l'environnement technique et le principe de fonctionnement des mémoires EEPROM-Flash. Ensuite nous verrons les limites physiques à la poursuite de la réduction de la taille des cellules mémoires ainsi que les principales solutions envisagées. Enfin, nous concluons sur les approches abordées dans cette thèse.

Sommaire

1.1	Positionnement des mémoires dans l'industrie des SCs	27
1.1.1	Contexte économique	27
1.1.2	Classification des mémoires	28
1.2	Fonctionnement et architecture des mémoires Flash	31
1.2.1	Présentation et fonctionnement de la cellule mémoire	31
1.2.1.1	Écriture	33
1.2.1.2	Effacement	34
1.2.2	Architectures des mémoires Flash	35
1.2.3	Les cellules mémoires multi-bits	37
1.2.4	Évolution technologique	39
1.3	Limitations des mémoires Flash et solutions envisagées	40
1.3.1	Problèmes liés à la réduction de la distance entre les cellules	41
1.3.1.1	Suppression des extensions latérales autour des cellules	41
1.3.1.2	Couplage entre cellules adjacentes	42
1.3.2	Problèmes liés à la réduction des cellules mémoires	43
1.3.2.1	Effets de canaux courts	43
1.3.2.2	Dégradation de l'oxyde <i>tunnel</i>	44
1.3.2.3	Nombre d'électrons pour coder un état	45
1.3.3	Intégration 3D des plans mémoires	46
1.3.4	Intégration de matériaux <i>High-κ</i> comme diélectrique d'interpoly	47
1.4	Les mémoires TANOS	49

1.4.1	Présentation de la technologie TANOS	49
1.4.2	Difficultés des mémoires TANOS et leurs évolutions	52
1.4.2.1	L'ingénierie de bandes de l'empilement <i>tunnel</i>	52
1.4.2.2	L'ingénierie de la couche de l'empilement de grille	53
1.5	Conclusion	56

1.1 Positionnement des mémoires non-volatiles dans l'industrie des semi-conducteurs

1.1.1 Contexte économique

Depuis le milieu des années 60, le marché des circuits intégrés (CI) et en particulier celui des mémoires, a connu un essor exceptionnel. Les revenus du marché des circuits intégrés pour l'année 2006 sont d'environ 150000 millions de dollars. La microélectronique est à la base des technologies de l'information et de la communication comme les ordinateurs personnels¹, les stations de travail², les serveurs³, les ordinateurs portables⁴, l'infrastructure d'Internet⁵, les téléphones portables⁶, etc. D'autres applications utilisant des circuits intégrés sont présentes dans des secteurs comme l'automobile, l'énergie, le transport, la médecine, etc. Ainsi l'industrie des semi-conducteurs joue un rôle important dans l'économie mondiale et sur les moyens de production depuis les dernières décennies (cf figure 1.1a).

Or le marché des mémoires, dont les revenus représentent presque 30% du marché des CI, est un moteur essentiel de l'industrie des semi-conducteurs. En effet, tous les produits électroniques qui connaissent actuellement un grand succès auprès du grand public (téléphones portables, ordinateurs, clés USB, lecteurs MP3, assistants personnels) contiennent des mémoires.

La solution idéale serait une mémoire qui retiendrait l'information sans alimentation électrique extérieure, avec un accès en lecture et une programmation rapide et tout cela avec une haute densité d'intégration et une basse consommation en énergie. Cependant, la mémoire idéale, regroupant tous ces avantages à la fois n'existe pas, même si les mémoires Flash en présentent plusieurs d'entre eux (cf figure 1.2a). Il existe donc plusieurs catégories de mémoires selon les applications visées.

¹PC : Personal computers, présenté par IBM le 12 Aout 1981.

²Les stations de travail sont généralement des ordinateurs puissants mis en réseau à la disposition d'utilisation. Ce terme date du milieu des années 80. Les stations représentaient le maximum de capacité informatique que l'on pouvait mettre à la disposition d'un utilisateur unique et étaient dédiées à des applications professionnelles : CAO (Conception Assistée par Ordinateur).

³Les premiers serveurs de données, utilisés pour faciliter l'échange de fichiers, sont apparus en 1980. En 2009, d'après le cabinet Netcraft, il y a plus de 220 millions de serveurs web dans le monde.

⁴Le premier ordinateur portable fut créé par Adam Osborne en 1981, le 1^{er} modèle s'appelait l'OsborneI. Il pesait 11kg et avait 64ko de RAM

⁵L'Internet tel que nous le connaissons aujourd'hui, l'interface *World Wide Web*, a été mis au point par Tim Berners-Lee en 1991, même si les premières briques de ce réseau avaient été pensées dès 1961 par Leonard Kleinrock

⁶Téléphone portable, téléphone mobile ou téléphone cellulaire se basent sur les technologies de radio communication par ondes électromagnétiques. L'invention, tout comme le 1^{er} appel, est généralement attribué à Martin Cooper directeur R&D chez Motorola en avril 1973.

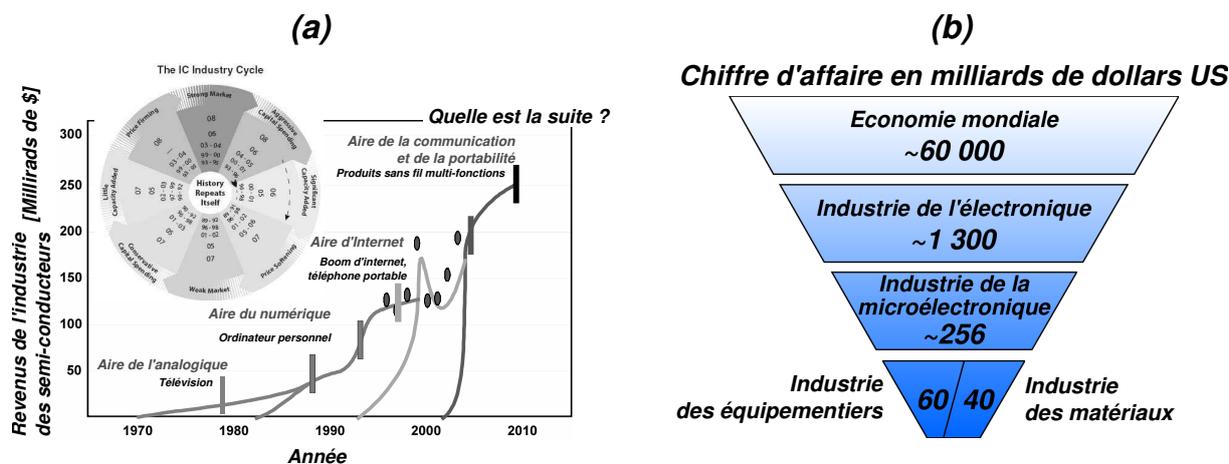


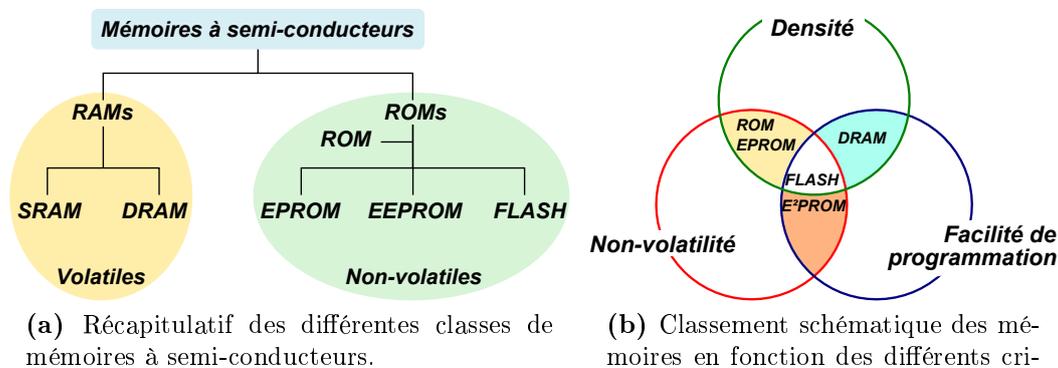
FIG. 1.1: a) Aire des différentes applications des semi-conducteurs (source : Semico Research Corporation) et cycle de l'industrie des circuits intégrés [ICI07]. b) Place de l'activité de la microélectronique et de son chiffre d'affaire par rapport aux activités économiques dans le monde (valeur 2007)[DeS07].

1.1.2 Classification des mémoires

Il existe plusieurs façons de classer les mémoires en fonction de l'angle sous lequel on regarde leurs caractéristiques ou leurs utilisations. Il est possible de classer les mémoires d'un point de vue architecture système : mémoire système, centrale, vidéo, stockage de masse, niveau de caches, registre, etc. Nous avons choisi comme principal critère de classement les caractéristiques électriques des mémoires sur semi-conducteur (cf figure 1.2a).

Les mémoires volatiles sont des mémoires qui perdent l'information lorsqu'elles ne sont plus alimentées. Il en existe deux types :

Les mémoires statiques sont des mémoires qui conservent l'information tant qu'elles sont alimentées. Elles n'ont donc pas besoin d'être rafraîchies. La plus répandue est la SRAM, *Static Random Access Memory*. Elle est constituée de



(a) Récapitulatif des différentes classes de mémoires à semi-conducteurs.

(b) Classement schématisé des mémoires en fonction des différents critères de performance.

FIG. 1.2: Classifications des mémoires à semi-conducteurs.

six transistors⁷. Comme son nom l'indique, il est possible d'accéder à n'importe quelle donnée de la mémoire de façon aléatoire⁸. La taille de chaque cellule est assez grande mais grâce à leur grande vitesse d'accès elles sont souvent intégrées au plus proche des microprocesseurs et comme 1^{er} niveau de cache.

Les mémoires dynamiques sont des mémoires qui ne conservent pas l'information dans le temps. Elles ont donc besoin d'être rafraîchies régulièrement. La plus connue est la DRAM, *Dynamic Random Access Memory*. Le point mémoire est constitué d'un transistor et d'une capacité servant à stocker une charge (information stockée). Ce sont les courants de fuite autour de la capacité (exemple : courant à l'état bas du transistor) qui la vident progressivement, d'où l'obligation de rafraîchir l'information. De plus, la lecture est destructive. Aussi d'un point de vue pratique, la lecture se fait par page. Ces données sont chargées et accessibles dans une mémoire tampon pendant que l'ensemble du plan mémoire est rafraîchi périodiquement. Les DRAM, de par leur petite taille et leur coût de fabrication faible, 1T1C⁹, peuvent être utilisées dans les applications avec une haute densité d'intégration, comme la mémoire centrale des ordinateurs couramment appelée *RAM*. C'est pourquoi, le marché des DRAM est largement supérieur à celui des SRAM, comme le montre la figure 1.3

Les mémoires non-volatiles conservent l'information lorsqu'elles ne sont plus alimentées. Il existe différentes technologies qui permettent cette propriété (bandes magnétiques, disques durs, CD/DVD, FeRAM, PCRAM, etc.). Néanmoins, nous ne détaillerons que les mémoires réalisées sur silicium qui sont couramment utilisées :

ROM : *Read Only Memory* est la première technologie commerciale de mémoires non-volatiles. Les données sont écrites par gravure, de façon permanente, lors de la fabrication. Elle est donc uniquement lisible mais garantit un temps de rétention de l'information quasi infini (à l'échelle d'une vie humaine). Son coût de fabrication est relativement élevé car son développement nécessite la réalisation d'une nouvelle série de masques de photolithographie dédiés. Elle répond encore aux applications nécessitant robustesse et vitesse de fonctionnement. Ses premiers succès commerciaux eurent lieu à partir des années 70 où elle était utilisée comme support de stockage pour les jeux vidéo. La ROM est maintenant fortement employée dans les microcontrôleurs [Bro97].

PROM : *Programmable ROM* est une mémoire proche des ROMs, mais programmable par l'utilisateur¹⁰. Elle fut inventée en 1956 par Wen Tsing Chow, à la demande de la United States Air Force¹¹. Elle constitue une alternative moins coûteuse que la ROM car elle est électriquement programmable par le client¹² mais de façon permanente, en faisant claquer des fusibles.

⁷2 transistors de sélection et 2 cellules inverseuses tête - bêche

⁸Il est courant d'utiliser les initiales *RAM* pour indiquer des mémoires à accès aléatoire en opposition aux mémoires à bande où il faut lire toutes les données précédentes pour obtenir l'information désirée.

⁹Notation pour indiquer une mémoire à 1 transistor et 1 capacité

¹⁰Les ROMs étaient programmées durant le procédé de fabrication.

¹¹Avant la première apparition des circuits intégrés dont la fabrication industrielle démarra en 1959 chez Texas Instruments

¹²La PROM ne nécessite pas un nouveau masque de photolithographie à chaque programmation

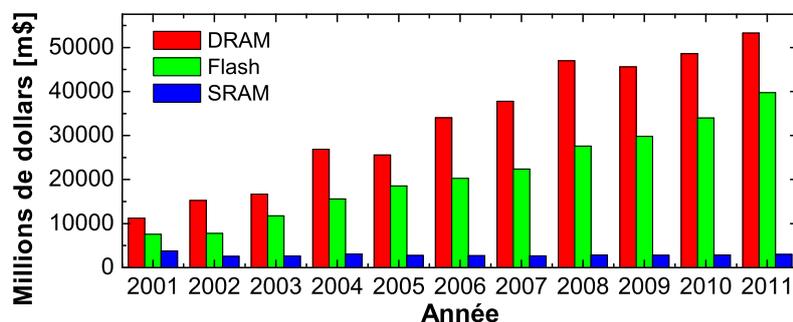


FIG. 1.3: Évolution des revenus du marché des mémoires à technologie MOS : les DRAM dominent mais les mémoires Flash gagnent du terrain : WSTS, IC Insights.

EPROM : *Erasable PROM* est une mémoire qui peut-être effacée et programmée par l'utilisateur. Mais l'effacement doit être effectué par ultra-violets, ce qui nécessite un démontage du boîtier et un passage sous rayons UV. La particularité de ces mémoires est qu'elles stockent l'information dans une grille flottante. Ce type de dispositif a été décrit pour la première fois par [Kah67] dans une structure MIMIS¹³. C'est la base des mémoires à grille flottante continue sur lesquelles reposent actuellement tous les produits EEPROM et Flash. [Weg67] proposa une alternative en remplaçant la grille flottante par une couche de piégeage en silicium nitruré (appelé couramment nitrure), c'était la première SONOS, *Silicium/Oxyde/Nitrure/Oxyde/Silicium*.

EEPROM : *Electrically EPROM* peut s'écrire et s'effacer électriquement, sans démontage de son boîtier. Le point mémoire est constitué de deux transistors : un transistor de sélection et une cellule mémoire. Elle présente donc une surface importante mais chaque cellule mémoire peut être programmé de manière indépendante. La première mémoire EEPROM (inspirée des structures MNOS¹⁴ et à grille flottante) a été introduite sur le marché en 1983 au format de 16Kbit.

Flash : ou mémoire Flash EEPROM est semblable à la mémoire EEPROM mais sans le transistor de sélection. La cellule mémoire pouvant avoir le rôle de transistor de sélection. La mémoire Flash EEPROM a été mise en production vers 1990. Son nom provient de son effacement ultra-rapide permettant l'effacement par secteur. Comme le point mémoire est constitué d'un seul transistor, la surface occupée est faible. Ainsi comme le montre la figure 1.2b, la mémoire Flash est un compromis entre les différentes mémoires citées précédemment. Pour toutes ces raisons, les mémoires Flash sont actuellement les mémoires non-volatiles les plus utilisées. La figure 1.3 montre la croissance remarquable du marché des mémoires Flash, qui, bien qu'inférieur à celui des DRAM, tend à le rejoindre d'ici les prochaines années.

¹³MIMIS : Metal Insulator Metal Insulator Semiconductor

¹⁴MNOS ou NROM sont des mémoires à couche de piégeage discret intégrant un oxyde *tunnel* épais, $\sim 5nm$, et une couche de nitrure jouant le rôle de couche de piégeage et d'oxyde bloquant, $\sim 60nm$.

1.2 Fonctionnement et architecture des mémoires Flash

1.2.1 Présentation et fonctionnement de la cellule mémoire

La cellule mémoire Flash à grille flottante peut être décrite selon la figure 1.4. Cette structure est basée sur celle d'un transistor MOS avec une zone de stockage de charges (usuellement grille en polysilicium) isolées entre deux diélectriques : l'oxyde *tunnel* et le diélectrique d'interpoly. En injectant une charge dans cette grille flottante, il est possible d'induire un décalage de la tension de seuil du transistor MOS, ΔV_T qui peut être exprimé de la façon suivante :

$$\Delta V_T = -\frac{Q_{GF}}{C_{IPD}} \quad (1.1)$$

Q_{GF} : charge stockée dans la grille flottante.

C_{IPD} : capacité entre grille de contrôle et grille flottante.

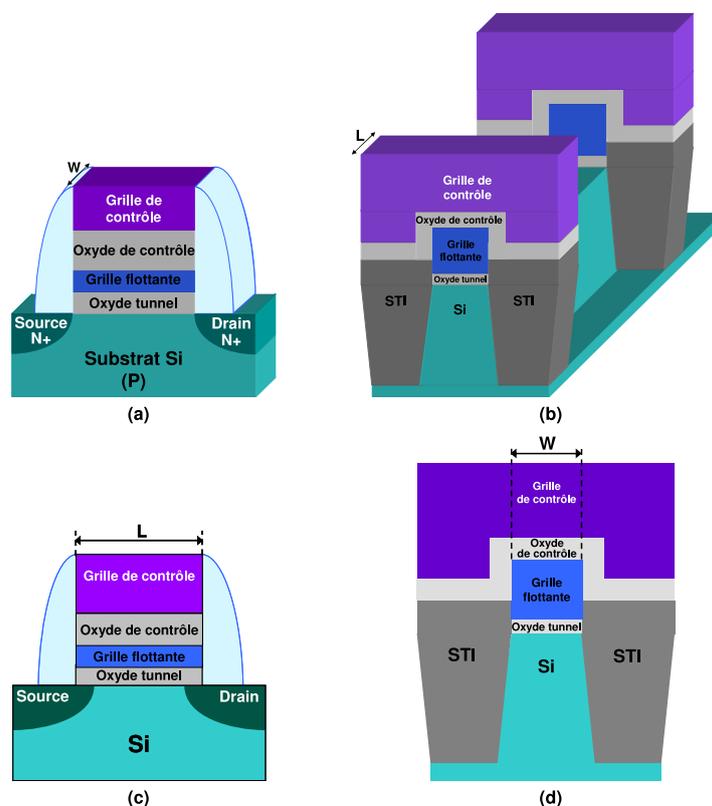


FIG. 1.4: Schéma d'une cellule mémoire Flash : (a) vue générale de la cellule mémoire, (b) vue schématique de deux cellules mémoires Flash dans un plan mémoire, (c) coupe longitudinale de la cellule, (d) coupe latérale de la cellule [Jac08].

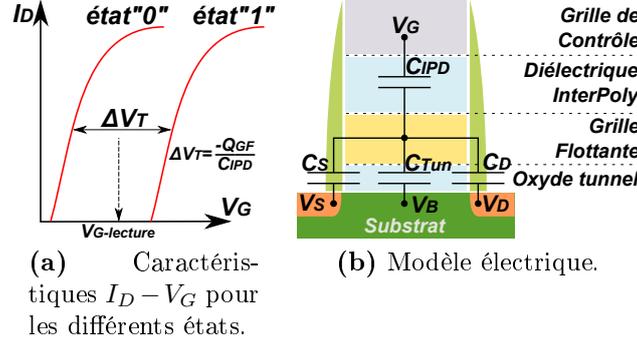


FIG. 1.5: Caractéristiques et schéma électrique d'une cellule mémoire Flash.

La lecture de la cellule s'effectue en appliquant une tension de grille comprise entre la tension de seuil du dispositif écrit et effacé (cf figure 1.5a). Cela détermine deux états logiques de la mémoire :

L'état 0 : la mémoire est effacée. La grille flottante est déchargée, et à $V_{G-lecture}$, le transistor est dit *passant* (le courant de drain est important).

L'état 1 : la mémoire est écrite, la grille flottante est chargée, et à $V_{G-lecture}$, le transistor est dit *bloqué* (le courant de drain est presque nul).

Le schéma électrique d'une mémoire Flash est donné à la figure 1.5b. L'application du théorème de Gauss à la grille flottante permet de relier le potentiel de la grille flottante aux autres potentiels mis en jeu en fonction des capacités du système (en négligeant le potentiel de bandes plates) :

$$V_{GF} = \frac{Q_{GF}}{C_{tot}} + \alpha_G \cdot V_G + \alpha_D \cdot V_D + \alpha_S \cdot V_S + \alpha_B \cdot V_B \quad (1.2)$$

$$\alpha_G = \frac{C_{IPD}}{C_{tot}}$$

$$\alpha_D = \frac{C_D}{C_{tot}}$$

$$\alpha_S = \frac{C_S}{C_{tot}}$$

$$\alpha_B = \frac{C_{Tun}}{C_{tot}}$$

$$C_{tot} = C_{IPD} + C_{Tun} + C_D + C_S$$

C_{Tun} : capacité grille flottante / canal.

C_D : capacité grille flottante / drain.

C_S : capacité grille flottante / source.

V_G, V_D, V_S et V_B : respectivement les tensions de grille, drain, source et substrat (*bulk*).

Si la source et le substrat sont à la masse et si la tension de drain est faible, l'expression se simplifie en :

$$V_{GF} = \frac{Q_{GF}}{C_{tot}} + \alpha_G \cdot V_G \quad (1.3)$$

Ainsi, plus le couplage de grille α_G est élevé, plus le potentiel de la grille flottante sera proche de la tension de grille appliquée, et plus le champ électrique dans l'interpoly sera réduit. Ce couplage traduit donc la manière dont le champ électrique va se répartir entre les isolants *tunnel* et interpoly. Ainsi, tout élément d'architecture de la cellule permettant d'augmenter α_G est favorable à sa programmation. On comprend maintenant l'intérêt des extensions latérales (figure 1.4d) et des diélectriques d'interpoly ONO (Oxyde/Nitride/Oxyde) qui permettent d'augmenter ce couplage. Néanmoins, un couplage trop fort masquerait la charge piégée et réduirait le ΔV_T , ainsi d'après [ITR07], $\alpha_G \simeq 0.6 \sim 0.7$.

1.2.1.1 Écriture

La figure 1.6 regroupe les différents modes d'écriture d'une cellule mémoire Flash.

Écriture Fowler-Nordheim : elle est effectuée en appliquant une tension élevée positive sur la grille de contrôle. Le fort champ électrique dans l'oxyde *tunnel* permet aux électrons du canal de franchir la barrière de potentiel de l'oxyde de grille par courant *tunnel*. Ce mode de programmation est généralement employé dans les architectures NAND. Il est plus lent que l'écriture par porteurs chauds mais il est moins destructif [Fow28].

Écriture par porteurs chauds : elle se fait en appliquant une tension positive à la fois sur le drain et la grille de contrôle. Du fait de l'application d'une tension de drain fort, les électrons du canal sont accélérés et acquièrent une énergie importante, supérieure à $\Delta E_{C_{SiO_2}}$. Ils sont ensuite déviés par la tension de grille et injectés dans la grille flottante. L'écriture par porteurs chauds nécessite de plus une jonction de drain abrupte pour générer des électrons à forte énergie et favoriser l'injection. Elle est beaucoup plus rapide que l'écriture par Fowler-Nordheim et nécessite des tensions de programmation plus faibles. Cependant, sa consommation est plus importante du fait du courant de drain élevé. Elle est aussi plus destructrice de par la forte énergie des électrons injectés [Tam84].

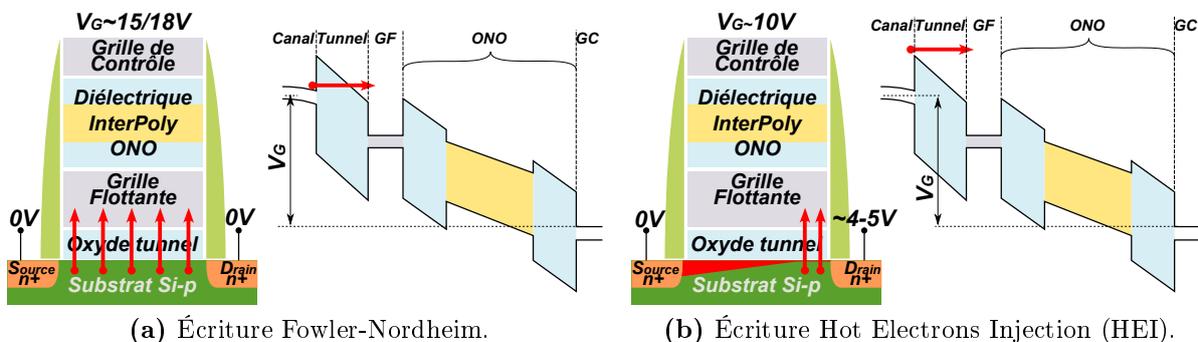


FIG. 1.6: Modes d'écriture des mémoires Flash.

1.2.1.2 Effacement

La figure 1.7 regroupe les différents modes d'effacement d'une cellule mémoire Flash.

Effacement par le canal : permet aux électrons de la grille flottante d'être émis dans le canal, en traversant l'oxyde *tunnel* par Fowler-Nordheim, dû à l'application d'une tension fortement négative entre la grille de contrôle et le substrat. La vitesse d'effacement est contrôlée par la tension de grille, la chute de potentiel dans l'interpoly restant faible (comme dans le cas de l'écriture FN). On notera pour les mémoires de type SONOS, que l'effacement s'effectue par l'injection de trous vers la couche de stockage de charges.

Effacement par la source : permet d'évacuer les électrons de la grille flottante dans la source par Fowler-Nordheim. Cet effacement nécessite un large recouvrement entre la jonction de source et la grille flottante. Le drain est de plus laissé flottant afin de limiter le courant de fuite drain-source. Les cellules NOR sont généralement effacées par la source (et donc par bloc de cellules étant donné que la source est commune).

Effacement mixte grille-source : permet d'effacer la cellule en partie par le canal et en partie par la source. De plus il permet l'effacement sélectif d'une cellule située à l'intersection d'une *bit line* et d'une *word line*.

Effacement par trous chauds : est principalement utilisé pour les mémoires NROM¹⁵. De la même façon que l'écriture par porteurs chauds, il s'agit d'injecter des trous très énergétiques dans la couche de piégeage [Lar04].

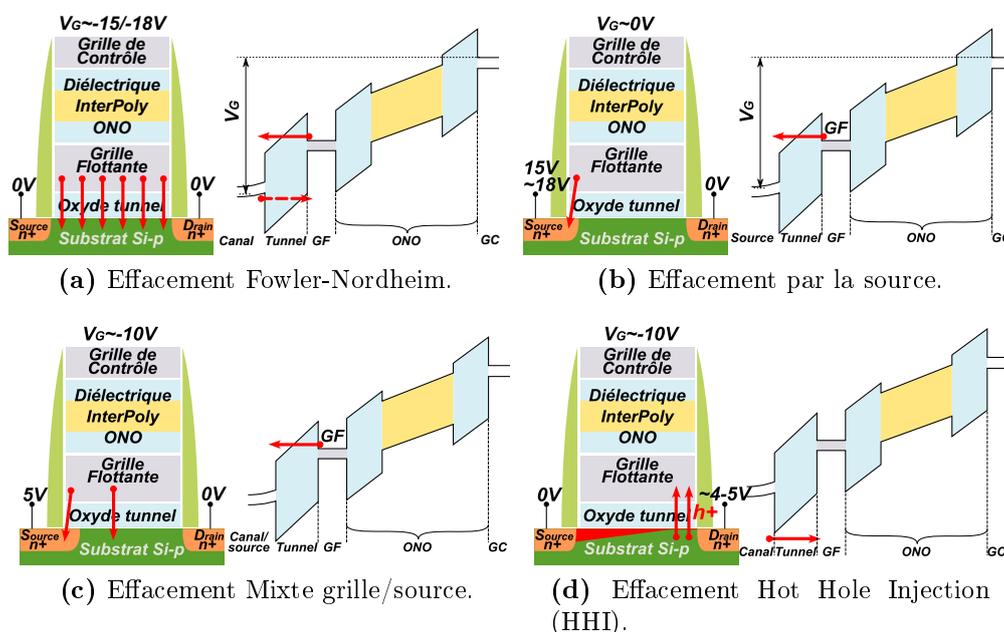


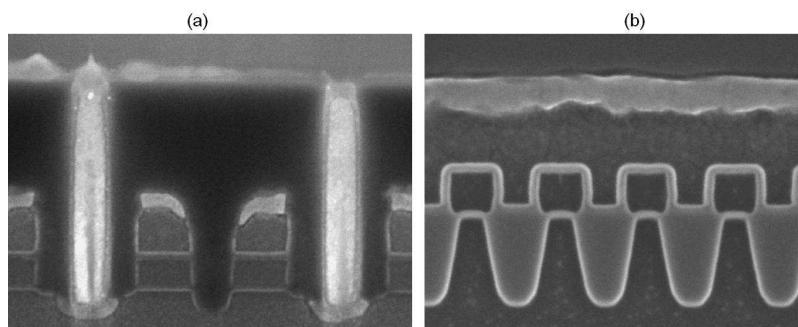
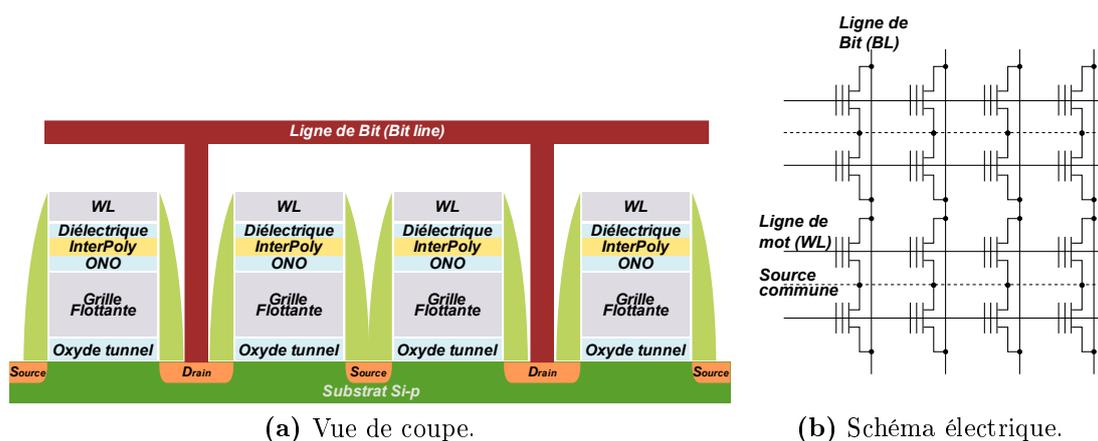
FIG. 1.7: Modes d'effacement des mémoires Flash.

¹⁵Les mémoires NROM sont généralement constituées d'un oxyde bloquant ($\sim 5nm$), d'une couche de nitrure épais et d'une grille. Le nitrure jouant le rôle de couche de piégeage et d'oxyde bloquant. Ces mémoires sont écrites et effacées par porteurs chauds

1.2.2 Architectures des mémoires Flash

Les deux principale architectures de mémoires Flash sont : NOR¹⁶ et NAND¹⁷. Cela correspond aux différentes architectures adoptées dans les matrices de cellules (cf figure 1.9) : NOR sont connectées en parallèle, les NAND sont connectées en série¹⁸. Voyons maintenant les caractéristiques spécifiques des NOR et des NAND :

La flash NOR : La grille de contrôle est commune à une ligne de cellules appelée ligne de mots (*word line*) et le drain est commun à une ligne de bits (*bit line*). Le drain de chaque cellule étant accessible, l'écriture peut s'effectuer par porteurs chauds. Les cellules sont effacées par Fowler-Nordheim. La programmation par porteurs chauds permet une écriture plus rapide d'une cellule isolée, mais réduit la durée de vie en terme de nombre de cycles. L'interface d'adressage permet un accès aléatoire à n'importe quelle position. La Flash NOR est essentiellement utilisée pour le stockage de codes d'instruction, ce sont les mémoires programmes des applications embarquées (applications *embedded*), téléphones portables, assistants personnels, ...



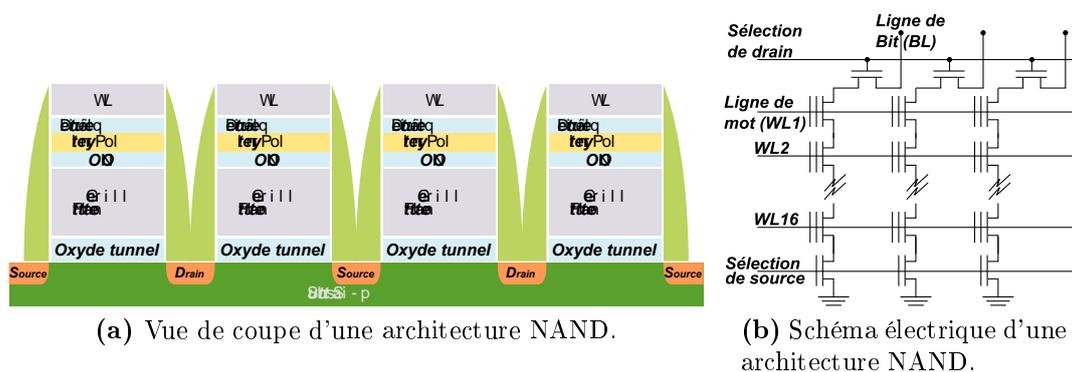
(c) Photographies TEM, a) parallèlement à la ligne de bits et b) parallèlement à la ligne de mots [Ser05].

FIG. 1.8: Architecture NOR.

¹⁶L'architecture NOR est la première à avoir été développée et mise en production en 1988 par Intel.

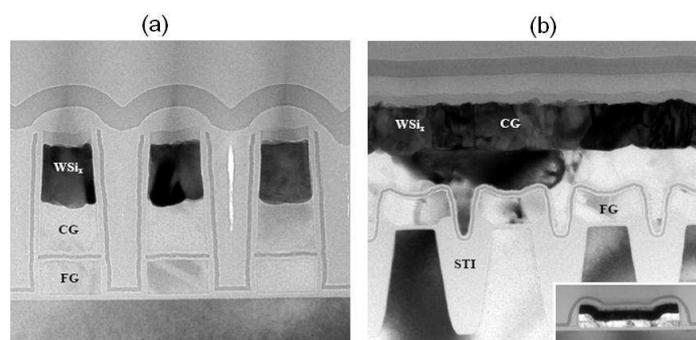
¹⁷L'architecture NAND a été développée par Toshiba en 1989.

¹⁸Ces architectures rappellent les portes logiques NAND et NOR, d'où leur nom.



(a) Vue de coupe d'une architecture NAND.

(b) Schéma électrique d'une architecture NAND.



(c) Photographies TEM, a) parallèlement à la ligne de bits et b) parallèlement à la ligne de mots [Jam07].

FIG. 1.9: Architecture NAND.

La flash NAND : Les cellules sont organisées en série (string). Le drain des cellules n'étant pas adressable, les cellules ne peuvent donc être écrites et effacées qu'en Fowler-Nordheim. L'accès en lecture à une cellule de la *bit line* s'effectue en polarisant la grille des autres cellules de façon à ce quelles soient passantes. C'est pourquoi le temps d'accès est plus long que celui des NOR. Mais de par la connexion en série des cellules, la programmation est réalisée par blocs mémoires, d'où des temps de programmation globaux plus rapides que la NOR. Enfin, comme on peut le voir sur la figure 1.9a, l'absence de contact de drain permet une diminution de la taille de la Flash NAND (environ 40% par rapport à une Flash NOR). La densité des Flash NAND est donc plus élevée pour un coût par bit plus faible. Ainsi, elle est donc plutôt utilisée pour le stockage de données, aussi appelé *stand alone* (clés USB, cartes mémoires, baladeur MP3 ...). En fonction des technologies réalisées, le nombre de cellules le long de la ligne de bits augmente : 16, 32 pour les technologies $\sim 63nm$, voir 44 pour le $\sim 55nm$ [Kim08].

Le marché des Flashs NOR a dominé longtemps celui des mémoires non-volatiles. Mais depuis 2004, le marché des NAND a presque doublé en taille (cf figure 1.10) grâce notamment, au succès des supports de stockage de données portables (lecteurs MP3, clés USB, appareils photos numériques...). Les flash NAND devraient encore accroître leur domination sur les NOR dans les prochaines années : le marché des NOR devrait croître de 6% d'ici 2011 et celui des NAND de 18%.

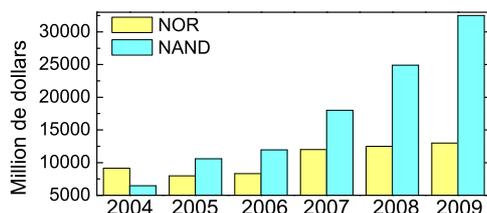


FIG. 1.10: Évolution du marché des mémoires Flash NAND et Flash NOR [ICI09].

On notera qu'avec des mémoires Flash NAND, la volonté de remplacer les disques durs, appelés aussi HDD (*hard disk drives*), est devenue de plus en plus forte. L'objectif est d'avoir une mémoire sans partie mécanique, sans latence¹⁹ et avec une consommation faible. L'augmentation des capacités des Flash NAND rend cela possible avec l'apparition de mémoires dit "SDD, *Solid-State Drives*". Leurs principaux avantages sont leur latence faible, leur robustesse aux chocs et leur faible consommation. Leurs principaux défauts sont leur capacité de stockage inférieure aux HDD (elles ne peuvent pas encore remplacer les disques durs classiques) et leur endurance²⁰ limitée (elles ne peuvent pas remplacer la DRAM), elles sont envisagées comme niveau de cache supplémentaire pour les serveurs de données ou comme mémoires systèmes pour les applications personnelles.

1.2.3 Les cellules mémoires multi-bits

Stocker plusieurs bits dans une seule cellule permet d'augmenter considérablement la densité de stockage, sans augmenter le nombre de cellules du plan mémoire. Cela permet une réduction du coût de développement²¹. Deux solutions ont été proposées afin de réaliser des mémoires permettant de stocker plus d'un bit par cellule.

La technologie multi-niveaux ou MLC²² a été pour la première fois démontrée en 1995 par Intel puis mise en œuvre dans sa famille StrataFlash [Atw97]. Elle tire profit de la nature analogique de la cellule Flash, en assignant une séquence de bits à une plage de tension spécifique. Ainsi pour coder 2 bits dans une cellule, il faudra 4 niveaux de tension différents, soit quatre V_T distincts. Les problèmes de cette technique sont liés au calibrage de la charge injectée, au maintien de la charge quelles que soient les conditions et la lecture précise des différents niveaux de tension. Les mémoires MLC sont donc plus lentes que les mémoires SLC, *Simple Layer Cell*, n'intégrant qu'un bit par cellule²³.

¹⁹La latence est le temps d'attente pour accéder à la première donnée d'un paquet, d'une banque ou d'une page mémoire.

²⁰L'endurance est le nombre de cycles d'écriture/effacement que l'on peut réaliser sur la mémoire avant qu'elle ne soit inutilisable.

²¹Dû au passage à la lithographie de prochaine génération.

²²MLC : *Multi Layer Cell*

²³ On notera que pour une mémoire multi-niveaux 2-bits (exemple : 0V, 2V, 4V et 6V) de surface

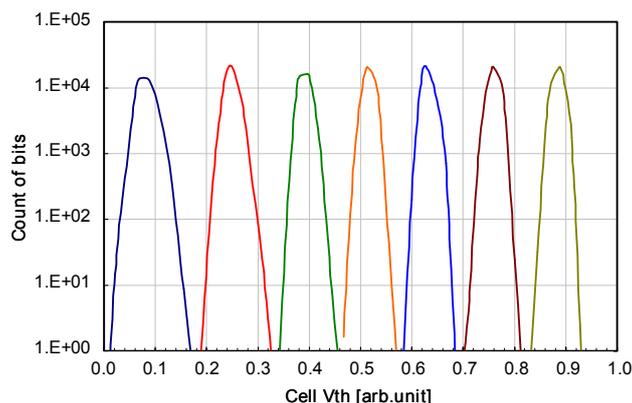


FIG. 1.11: Multi-niveaux : 7 niveaux de programmation pour une mémoire NAND 16Gbit [Kim08].

La technologie **Bit miroir** est une approche retenue par AMD et Fujitsu. La couche de piégeage est partagée en deux zones physiques distinctes aux extrémités desquelles sont piégées deux charges. Ainsi, cela nécessite une couche de piégeage discret (nitruure, nanocristaux, ...) et la possibilité d'inverser les rôles du drain et de la source du transistor. Cette approche est donc réservée aux architectures NOR et à une écriture par porteurs chauds.

Il est donc possible d'utiliser ces deux méthodes pour augmenter le nombre de bits par cellule : en utilisant la localisation de la charge par porteurs chauds (Bit miroir) et en contrôlant la charge injectée (multi-niveaux). Dans tous les cas, une grande fenêtre mémoire parfaitement contrôlée contribue fortement à la bonne réalisation des mémoires multi-niveaux.

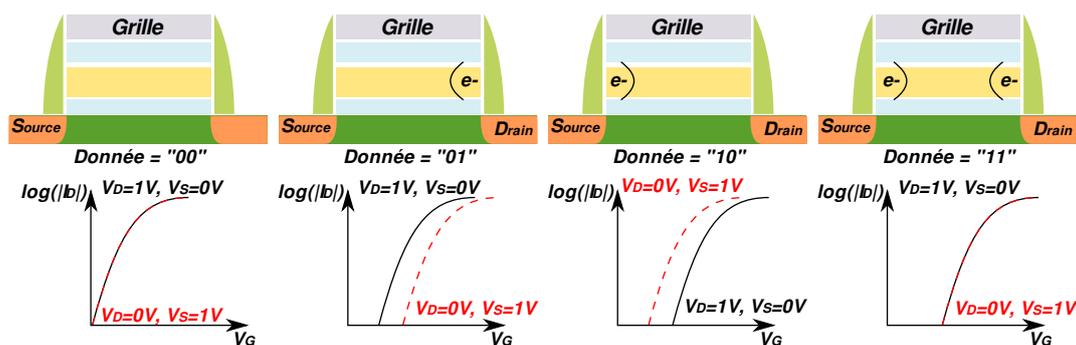


FIG. 1.12: Bit miroir : la localisation physique de la charge permet de coder deux bits dans une cellule mémoire

$100\text{nm} \times 100\text{nm}$, la différence entre chaque niveau est de ~ 400 électrons dans la grille flottante. Pour une mémoire 1-bit de surface $70\text{nm} \times 70\text{nm}$ (donnant la même densité d'information) avec un ΔV_T de 6V, il faudra ~ 600 électrons pour coder ce bit. De manière générale, le nombre d'électrons séparant les niveaux d'une mémoire multi-niveaux divisé par le nombre d'électrons utilisés pour coder une mémoire 1bit (plus petite, mais donnant la même densité mémoire) est de $\frac{n}{2^n - 1}$.

1.2.4 Évolution technologique

La technologie des mémoires non-volatiles, en particulier celle des architectures Flash NAND, est le secteur où l'évolution technologique est la plus rapide des technologies silicium. Nous avons résumé l'évolution des technologies Flash NAND stand-alone²⁴ de 90nm à 43nm sur la figure 1.13. On notera que le nœud technologique 32nm est déjà rentré en production courant 2009 [Fab08].

On peut observer, que les approches sont restées les mêmes d'une génération à une autre : grille flottante en polysilicium et diélectrique d'interpoly en ONO. Le coefficient de couplage est maintenu à 0,6 ~ 0,7 par des débordements latéraux de la grille de contrôle sur la grille flottante.

Néanmoins, il semble que les limites de cette technologie (grille flottante et diélectrique d'interpoly ONO) vont être atteintes à l'approche du nœud technologique 32nm. Le marché des mémoires Flash est donc face à de nombreux challenges et devra peut-être faire face à plusieurs ruptures technologiques pour continuer la miniaturisation et l'augmentation des densités mémoires.

Le marché des mémoires embarquées est quant à lui un élément clé de l'augmentation des performances des systèmes sur puce : SOC (System On Chip). La technologie embarquée doit donc arriver à s'adapter aux contraintes des technologies MOS en minimisant les coûts de développement et de fabrication.

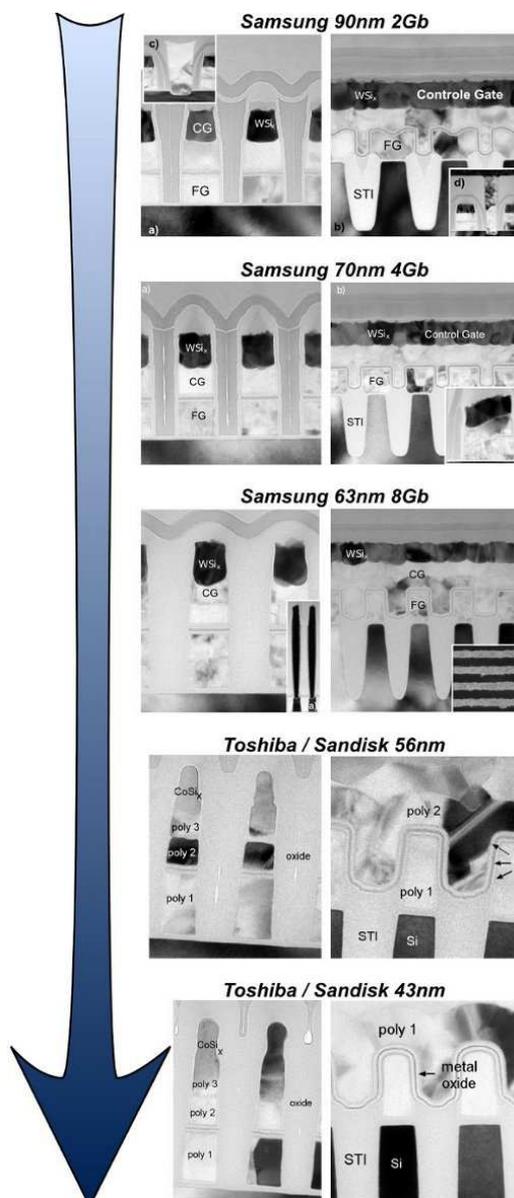


FIG. 1.13: Photographies MET prises parallèlement à (a) la ligne de bits et (b) la ligne de mots pour différentes technologies de produits NAND stand-alone [Jam07, Ice08].

²⁴Le terme *stand-alone* regroupe les produits n'offrant que la fonctionnalité mémoire. Cela est en opposition au terme *embedded* désignant les applications embarquées.

1.3 Limitations des mémoires Flash et solutions envisagées

Le tableau 1.1 résume les principales spécificités que doivent suivre les mémoires Flash en fonction des différents nœuds technologiques. Il représente un plan de route²⁵ des mémoires non-volatiles donné par l'ITRS, mis à jour tous les deux ans [ITR07], cf tableau 1.1. Les zones en jaune représentent des points qui posent problèmes mais pour lesquels les solutions sont connues. Les parties en rouge sont les points pour lesquels aucune solution n'est connue.

Il est intéressant d'observer que pour les mémoires à grille flottante continue le diélectrique d'interpoly ONO²⁶ est prévu jusqu'au nœud technologique 32nm. Pour les générations plus agressives, l'intégration devient très délicate et demande des modifications importantes de l'empilement mémoire pour résoudre les problèmes physiques, électriques, etc [Pra06].

Année	2009	2010	2011	2012	2013	2014	2015
Nœud technologique F							
DRAM [nm]	50	45	40	36	32	30	25
Flash NAND [nm]	40	36	32	28	25	22	20
Flash NOR [nm]	50	45	40	35	32	28	25
Taille des cellules F^2							
NAND	4						
NOR grille flottante	9 – 11						
NOR SONOS/NROM	6-7	6-7	7-8	7-8	7-8	7-8	8-9
Coefficient de couplage	0,6 – 0,7						
Rétention [années]	10 – 20						
1— Flash NAND							
Endurance [cycles]	10 ⁵						
Nb max de bits par cellule	3	4					
EOT de l'oxyde <i>tunnel</i> [nm]	6 – 7						
Composition de l'oxyde de contrôle	ONO			High- κ			
EOT de l'oxyde de contrôle [nm]	10-13	10-13	10-13	9-10			
2— Flash NOR							
EOT de l'oxyde tunnel [nm]	8 – 9	8				7 – 8	
Composition de l'oxyde de contrôle	ONO				High- κ		
EOT de l'oxyde de contrôle [nm]	13 – 15				8 – 10		
Tension de programmation [V]	7 – 9				6 – 8		
Endurance [cycles]	10 ⁵	10 ⁶					
Nb max de bits par cellule	2						

TAB. 1.1: Résumé des objectifs technologiques pour les mémoires non-volatiles de type Flash-NOR et Flash-NAND [ITR07].

²⁵Couramment appelé *road map*.

²⁶ONO : Oxyde/Nitride/Oxyde désigne les différents matériaux qui composent ce diélectrique d'interpoly. Cet empilement est couramment utilisé dans les mémoires non-volatiles.

1.3.1 Problèmes liés à la réduction de la distance entre les cellules

1.3.1.1 Suppression des extensions latérales autour des cellules

D'après l'ITRS 2007 [ITR07], pour les nœuds technologiques entre $32\text{nm} - 28\text{nm}$, l'espace entre deux cellules NAND adjacentes (suivant la ligne de mots) n'est plus suffisant pour permettre la réalisation d'extensions²⁷ ou de débordements latéraux de la grille de contrôle sur la grille flottante (cf figure 1.14a). Or, pour une structure plane, ce sont ces ajustements technologiques qui permettent d'assurer un couplage suffisant entre ces deux grilles (cf figure 1.14b). Ainsi, sans le couplage par les flancs, le coefficient de couplage chute à 0,3 pour le nœud 32nm , au lieu des 0,6 prévu par l'ITRS (cf 1.14c). De plus, le diélectrique d'interpoly couramment utilisé (ONO : Oxyde-Nitride-Oxyde) atteint ses limites en terme d'épaisseur. En effet, pour des raisons de fiabilité, son EOT²⁸ ne peut pas être réduite en dessous de 15nm [Kim05].

L'une des solutions est l'intégration de matériaux à forte permittivité, appelés *High- κ* , pour remplacer l'empilement ONO en diélectrique d'interpoly pour les générations sub- 28nm [ITR07]. Ces matériaux *High- κ* permettent une augmentation importante du couplage grâce à leur forte constante diélectrique. Ainsi, ils permettraient de maintenir un coefficient de couplage de 0,6 – 0,7 même pour les plans mémoires de grande densité.

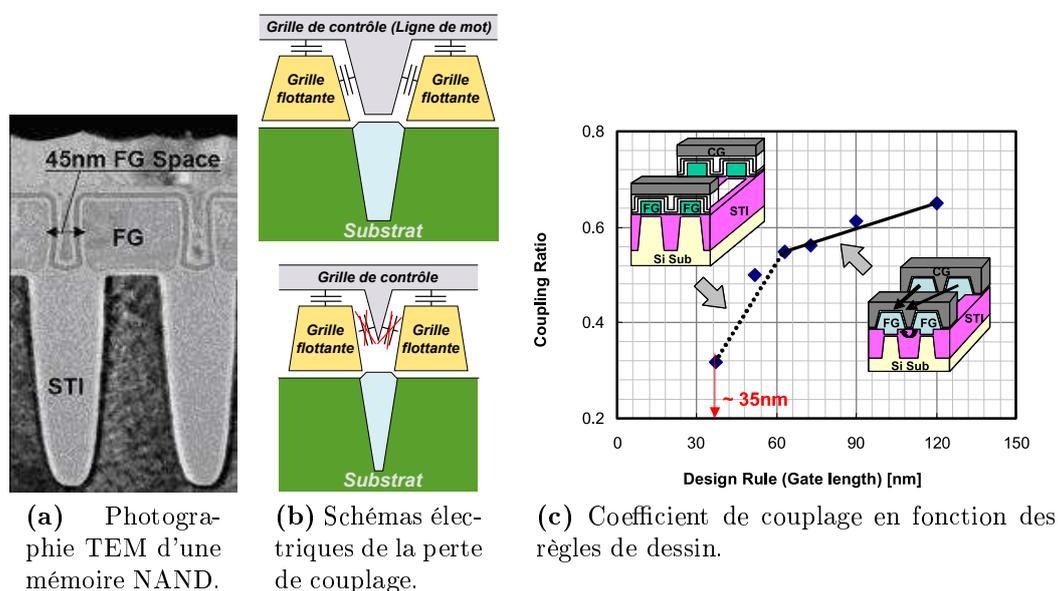


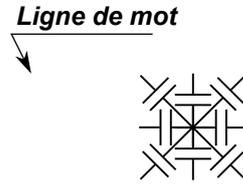
FIG. 1.14: a-b) Impossibilité de réduire la distance entre cellules adjacentes en gardant les flancs latéraux de la grille de contrôle. c) Réduction dramatique du coefficient de couplage pour les technologies avancées due à la suppression des flancs latéraux [Kim05].

²⁷ Appellées *wings*.

²⁸ EOT : Equivalence Oxide Thickness, il s'agit de l'épaisseur équivalente d'oxyde de silicium qui aurait le même couplage. $EOT = \frac{3,9}{\epsilon_r} \text{ épaisseur}$.

1.3.1.2 Couplage entre cellules adjacentes

Une autre limitation de la miniaturisation est l'augmentation des interférences entre cellules. En effet, lorsque la densité d'intégration augmente, les cellules se rapprochent et le couplage capacitif entre grilles flottantes augmente (cf figure 1.15). Ainsi, la charge stockée dans une grille flottante peut modifier la tension de seuil de la cellule voisine [Huo09, Blo09b].



Pour palier à cela, le remplacement des espaceurs en nitrure par du SiO_2 et l'utilisation de matériaux à très basse permittivité, dits *low- κ* , entre les cellules sont envisagés [Kim07]. L'objectif de ces développements est de diminuer la permittivité du diélectrique qui assure l'isolation entre les cellules. Néanmoins, ces techniques sont très vite limitées par la maturité des matériaux poreux et par la permittivité du vide.

Ligne de bit ↗

FIG. 1.15: Schéma électrique du couplage entre cellules NAND.

Une autre solution est l'utilisation de mémoires moins sensibles au couplage inter-cellules. Les mémoires à nanocristaux [Mol07] et à couche de piégeage nitrure présentent une plus grande immunité à ces couplages parasites (cf figure 1.16). En effet, ces couches de piégeage sont fines (entre $4nm$ et $10nm$) comparées à celles des mémoires à grille flottante continue (une centaine de nanomètres) et diminuent la surface en regard entre les cellules, réduisant ainsi grandement le couplage. De plus, des algorithmes de découplage, similaires à ceux utilisés dans les disques durs²⁹, peuvent aussi être utilisés [Li 09].

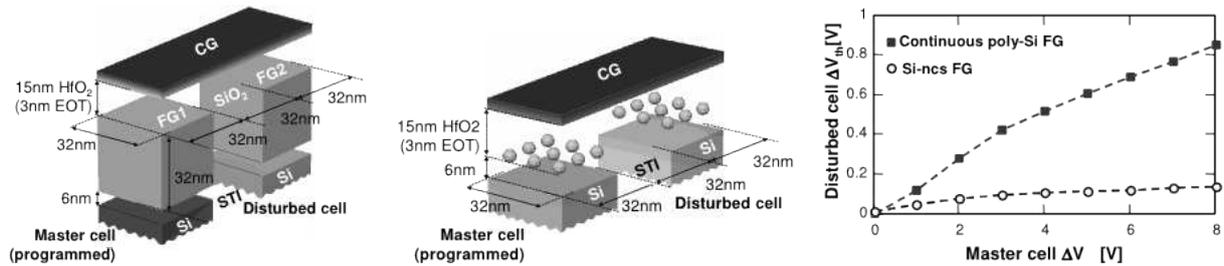


FIG. 1.16: ΔV_T induit par la charge d'une cellule adjacente pour une mémoire à nanocristaux ou à grille flottante continue [Mol07].

²⁹En effet, la tête de lecture étant trop grosse pour la densité de l'information, on lit plusieurs bits à la fois avec une plus ou moins grande importance du bit central. Il est donc nécessaire de décorrélérer les informations lues.

1.3.2 Problèmes liés à la réduction des cellules mémoires

1.3.2.1 Effets de canaux courts

La limitation principale de la miniaturisation des mémoires Flash est la réduction de l'épaisseur de l'oxyde *tunnel*. En effet, il est nécessaire de diminuer cette épaisseur lorsque les dimensions de la cellule diminuent si l'on veut garder un bon contrôle électrostatique du canal. Il s'agit des mêmes difficultés que l'on rencontre pour les transistors MOS. Mais la diminution de l'épaisseur du diélectrique *tunnel* affecte la rétention de la charge. La limite intrinsèque de l'épaisseur permettant de respecter le critère de dix ans est d'environ 6nm pour une mémoire Flash à grille flottante continue [Lai98]. En dessous de cette limite, le courant de fuite à travers l'oxyde *tunnel* est trop important.

L'une des solutions est d'utiliser des mémoires Flash de type FinFET [Kim05]. Comme la structure FinFET [Jah05], les FinFlash utilisent une grille débordante le long du canal permettant un meilleur contrôle électrostatique, améliorant le courant de sortie, sans réduction de l'épaisseur de l'oxyde *tunnel* (cf figure 1.17). Cette augmentation du courant I_{ON} permet aussi une réduction du temps d'accès pour les structures NAND.

Néanmoins, malgré différentes variantes des structures FinFET, sur substrat SOI ou sur substrat massif, il n'y a, à ce jour, aucune technologie FinFLASH commercialisée. De nombreux challenges restent encore à résoudre, comme par exemple, l'espace entre les différents doigts de Si le long de la ligne de mot qui doit être suffisant pour intégrer l'empilement de grille.

De plus, il est important de noter que les structures cylindriques permettent une augmentation du couplage entre la grille de contrôle et la grille flottante avec la diminution du rayon de courbure [Now08]. Ainsi, cela permet une augmentation du coefficient de couplage (*coupling ratio*) et des performances mémoires par effet géométrique [Cha08] en utilisant les propriétés électrostatiques des capacités cylindriques (cf figure 1.18).

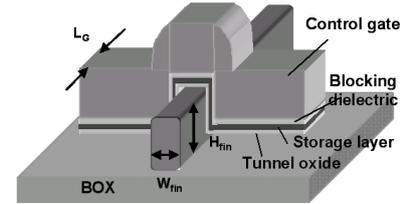
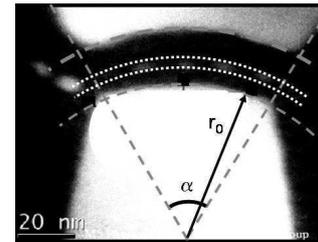


FIG. 1.17: Schéma d'une cellule FinFlash [Jah08].



C_i : [F/m] J_i : cylindricale [A/m] Q_{cyl} : [C/m]

$$J_i^{cyl} = 2\pi r_{i-1} J_i^{planar}, \quad (i=1, 2) \left. \begin{array}{l} dQ \\ dQ_{cyl} = J_2^{cyl} - J_1^{cyl} \end{array} \right\} \Delta V_{th} = \frac{Q_{cyl}}{C_2}$$

$$C_1 = \frac{2\pi\epsilon_{ox}}{\ln(r_1/r_0)} \quad C_2 = \frac{2\pi}{\sum_{i=1}^n \frac{1}{\epsilon_i} \ln(r_{i+1}/r_i)}$$

FIG. 1.18: Vue de coupe d'une mémoire à nanocristaux améliorant l'injection grâce à une géométrie cylindrique [Ger08].

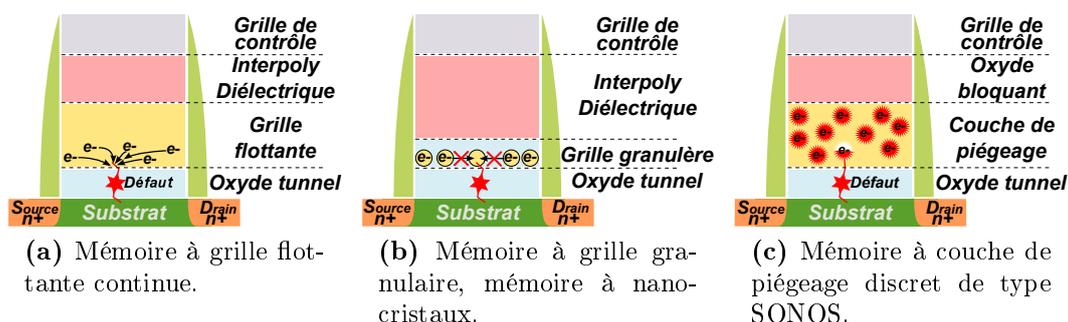


FIG. 1.19: Schémas de la perte d'électrons dans la zone de stockage due à un chemin de conduction dans l'oxyde *tunnel*. La mémoire à grille flottante continue perd la totalité de sa charge alors que la mémoire à couche de stockage discret perd seulement la charge située au-dessus du défaut.

1.3.2.2 Dégradation de l'oxyde *tunnel*

Outre la limite intrinsèque de l'épaisseur de l'oxyde *tunnel*, due à la nécessité de maintenir la rétention, il existe une autre limite due au vieillissement causé par la répétition des cycles d'écriture/effacement. Durant la programmation, l'oxyde *tunnel* est soumis à de forts champs électriques ($\sim 10MV/cm$). Ce stress excessif peut induire un courant de fuite important dû à la formation de défauts dans le diélectrique. Ce phénomène de courant induit par un stress est appelé SILC, *Stress Induced Leakage Current* [Mas82, Nar88, Oli88]. Etant donné que le SILC augmente lorsque l'épaisseur de l'oxyde *tunnel* diminue, l'épaisseur de l'oxyde *tunnel* est limitée à $6nm$ [ITR07].

L'approche la plus répandue pour se prémunir du SILC est l'utilisation de mémoires à couche de piégeage discret, telles que les mémoires nitrures (SONOS³⁰, NROM³¹, ...) ou les mémoires à nanocristaux. En effet, les pertes de charge par SILC sont dues à des ponts de conduction dans l'oxyde. Dans le cas d'une grille flottante continue, la présence de défauts dans l'oxyde *tunnel* peut provoquer la fuite de la totalité de la charge stockée dans la grille flottante. En revanche, avec des sites de stockages discrets, isolés les uns des autres, seuls les électrons proches du défauts seront perdus (figure 1.19). Ce concept permet la réduction de l'épaisseur de l'oxyde *tunnel* sans compromettre la rétention et l'endurance de la cellule mémoire.

La problématique globale de l'oxyde *tunnel* est un compromis entre vitesse de programmation et rétention. La diminution de son épaisseur en favorise un, mais dégrade l'autre, et vis-versa. C'est d'ailleurs, l'une des principales difficultés des mémoires SONOS, d'allier effacement et rétention [Pra06].

³⁰Le nom SONOS vient des initiales de l'empilement mémoire : Silicium / Oxyde / Nitrure / Oxyde / Silicium. Ces mémoires sont principalement utilisées pour des programmations Fowler-Nordheim. Les épaisseurs typiques sont $8nm$ pour l'oxyde bloquant, $6nm$ pour le nitrure et $2 \sim 3nm$ pour l'oxyde *tunnel*.

³¹La mémoire NROM ne possède pas d'oxyde bloquant. La couche de nitrure est épaisse et joue le rôle de couche de piégeage et d'isolation. L'oxyde *tunnel* est épais $> 5nm$. Elles sont programmées par porteurs chauds.

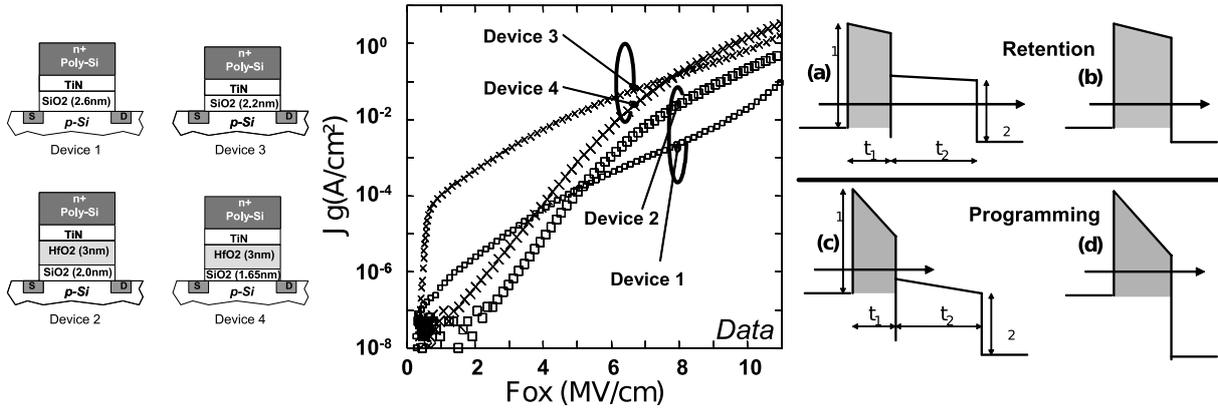


FIG. 1.20: Mesures de courant à travers un oxyde de silicium ou un bicouche SiO_2/HfO_2 . Ces caractéristiques mettent en évidence les meilleures propriétés d'injection de l'empilement bicouche permettant d'améliorer la rétention et la programmation [Buc05].

L'ingénierie de bande est présentée comme une approche intéressante pour améliorer programmation et rétention, comme le montre la figure 1.20. L'idée est de remplacer l'oxyde de silicium par un empilement de plusieurs diélectriques. En travaillant sur les constantes diélectriques et les hauteurs de barrière, il est possible d'améliorer la sensibilité en champ du courant à travers l'oxyde *tunnel*, et ainsi obtenir, à faible champ un courant faible³² et à fort champ un courant fort³³. La principale difficulté de cette approche est l'intégration de nouveaux matériaux dans l'empilement *tunnel*. En effet, ces diélectriques n'étant pas des oxydes parfaits, ils peuvent introduire un piégeage parasite ou des conductions assistées par pièges, perturbant le fonctionnement de la mémoire. Toutefois, cette approche semble prometteuse pour les mémoires SONOS [Lai07].

1.3.2.3 Nombre d'électrons pour coder un état

Outre les limites technologiques, il existe des limites physiques intrinsèques. La réduction des dimensions des dispositifs implique que la quantité d'électrons utilisés pour coder un bit³⁴ devient de plus en plus faible (figure 1.21a). Cette réduction, qui peut aller jusqu'à une dizaine d'électrons pour les générations décananométriques, peut induire des problèmes de fiabilité [Mol04]. Pour les dimensions ultimes, les phénomènes de chargement/déchargement ne peuvent plus être décrits comme continus. Il faut considérer le chargement/déchargement, électron par électron. Ces phénomènes stochastiques induisent une dispersion sur le temps de rétention et la fenêtre de programmation (cf figure 1.21b). Ces effets ont été observés pour les nœuds technologiques $90nm$ et $60nm$ [Com03], mettant en évidence que ces nouvelles contraintes devraient être prises en compte, notamment pour les intégrations ultimes de Flash NAND multi-niveaux.

³²En rétention le courant de fuite est faible, la charge est conservée.

³³En programmation le courant d'injection est fort, la cellule s'écrit/s'efface rapidement.

³⁴Nombre d'électrons qui séparent le V_T écrit et le V_T effacé.

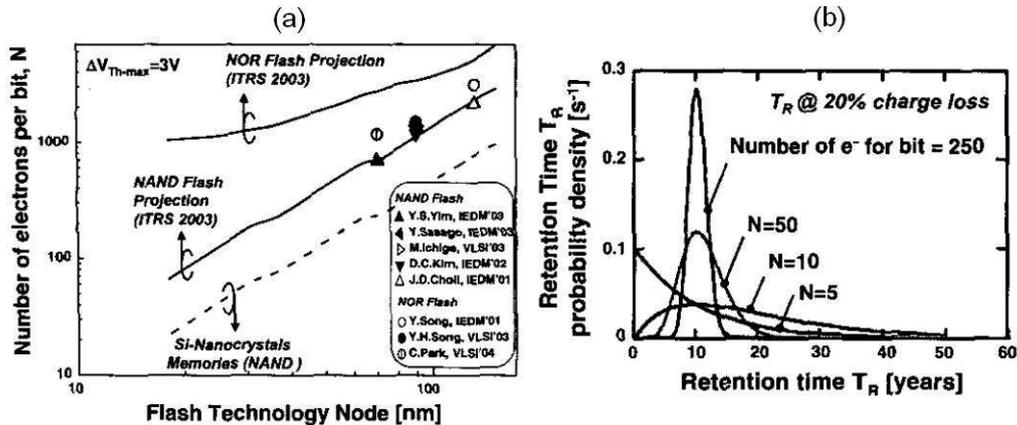


FIG. 1.21: (a) Nombre d'électrons pour coder un bit en fonction du nœud technologique Flash. (b) Densité de probabilité du temps de rétention pour différentes quantités d'électrons stockés [Mol04].

1.3.3 Intégration 3D des plans mémoires

Une solution, pour augmenter la densité mémoire (bit/m^2) sans diminuer la taille des cellules et ainsi d'éviter les problèmes électriques, physiques et de fiabilité, consiste à empiler verticalement les plans mémoires. Plusieurs approches existent :

Travail sur le conditionnement : il s'agit du *Packaging 3D*. Cette technique consiste à assembler des puces mémoires fabriquées séparément, puis à les coller les unes sur les autres pour réaliser le produit fini. Elle est appelée *bonding*. C'est une approche simple mais coûteuse : le prix du bit restant le même que pour une technologie classique.

Intégration séquentielle : elle concerne les approches dites *buildup*. Elle a été démontrée par Samsung [Jun06]. Elle consiste à fabriquer un premier plan mémoire, puis à réaliser un autre plan mémoire par-dessus. Cette approche permet de regrouper

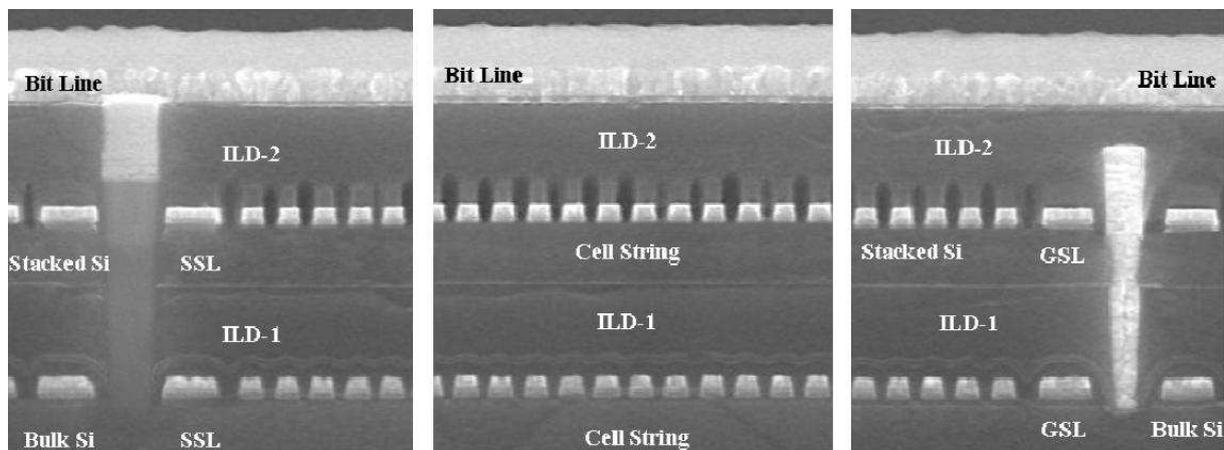


FIG. 1.22: Photographies MET d'une vue en coupe des cellules Flash d'une intégration séquentielle, démontrées par Samsung dans [Jun06].

les étapes de fabrication des interconnexions, *Back-end*. Ses inconvénients sont la difficulté à réaliser le deuxième niveau (substrat, budget thermique, détérioration du premier, ...) et d'assurer un bon alignement.

Intégration monolithique : cette approche a été présentée par Toshiba [Fuk07]. Les cellules mémoires sont intégrées verticalement. Elles sont auto-alignées, avec un canal en poly-silicium, ce qui permet de réaliser l'ensemble des différents niveaux 3D, les uns à la suite des autres sans procédé de photo-lithographie ou gravure supplémentaire. Cette approche permet une réduction importante du coût de fabrication d'un bit, appelé BiCS (Bit-Cost Scalable Flash Memory). Le principal inconvénient est son canal en poly-silicium et la difficulté de développement d'une structure aussi complexe.

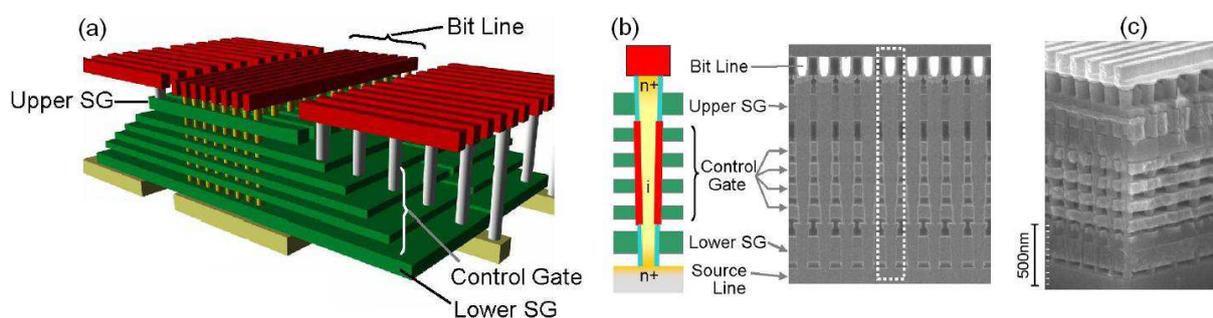


FIG. 1.23: (a) Schéma 3D, (b) vues en coupe, (c) photographie 3D de cellules Flash de l'approche monolithique démontrée par Toshiba dans [Fuk07].

1.3.4 Intégration de matériaux *High- κ* comme diélectrique d'interpoly

L'intégration de matériaux à forte permittivité comme diélectrique d'interpoly est une solution intéressante pour compenser la perte de couplage due à la suppression des débordements latéraux et ainsi maintenir un coefficient de couplage de 0,6 – 0,7, même pour les plans mémoires de grande densité. Leur intégration est aussi une piste intéressante pour les applications embarquées en vue de la réduction des tensions de programmation. En effet, [Pow08] présente un dispositif mémoire intégrant de l'alumine et possédant une bonne fiabilité sur un produit de 2Mb pour les applications embarquées.

La figure 1.24 présente un exemple de réalisation de cellule mémoire à grille flottante sans débordements latéraux [Van06a]. Le diélectrique d'interpoly est réalisé en HfO_2 ou en $HfSiON$. La figure 1.24d met en évidence la réduction des tensions de programmation avec l'intégration de matériaux *High- κ* comme diélectrique d'interpoly et la réduction de son EOT.

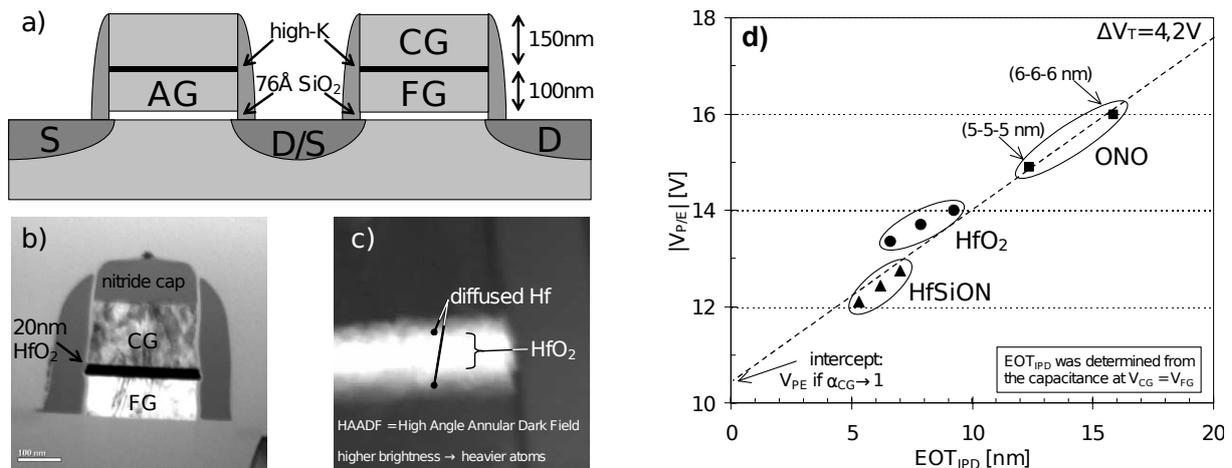


FIG. 1.24: Exemple d'intégration de *High- κ* dans une cellule mémoire à grille flottante. Photographie MET (b) de la cellule mémoire, on remarque l'absence de débordements latéraux, et (c) de l'interface Poly-Si/ HfO_2 . (d) Évolution de la tension de programmation en fonction de l'EOT du diélectrique d'interpoly. Les matériaux *High- κ* permettent une réduction notable des tensions d'écriture/effacement [Van06a].

Les matériaux les plus étudiés en vu de leur intégration sont :

- L'alumine [Wel06, Pow08].
- L'oxyde d'hafnium [Van06a].
- Les aluminates d'hafnium [Mol06].
- Les silicates d'hafnium [Mir06].

Néanmoins, les matériaux *High- κ* sont encore mal connus et mal maîtrisés en comparaison à l'oxyde de silicium. Ils souffrent de charges fixes, de piègeages parasites et de courants de fuite assistés par pièges [Ler04, Wil01, Tsa08]. Ils nécessitent donc une étude approfondie pour comprendre les mécanismes mis en jeu et pour une utilisation optimale dans les dispositifs. Ainsi, malgré leur fort potentiel, les matériaux *High- κ* ne sont pas encore intégrés dans des produits mémoires³⁵. Ils sont malgré tout très attractifs pour les applications embarquées car permettraient de réduire sensiblement les tensions de programmation.



Les difficultés rencontrées par les mémoires Flash demandent de nouvelles architectures et l'utilisation de nouveaux matériaux. La suppression des extensions latérales de la grille de contrôle autour de la grille flottante est la plus importante rupture technologique. L'intégration de *High- κ* dans le diélectrique d'interpoly est une solution très fortement envisagée. Néanmoins, ces oxydes demandent encore un développement avant être intégré dans des produits. On notera aussi que la disparition des extensions latérales et l'intégration de *High- κ* a rendu plus attractif les mémoires à couche de piégeage nitrure et à nanocristaux. En effet, elles présentent une grande robustesse au couplage inter-cellule et au SILC. Couplées à l'intégration de *High- κ* , comme l' Al_2O_3 , elles sont annoncées comme une structure mémoire prometteuse pour le suivi du plan de route de l'ITRS.

³⁵On notera l'annonce de Toshiba qui utiliserait un matériau *High- κ* comme diélectrique d'interpoly (à base d'oxyde d'hafnium) et une grille métallique (métal siliciuré) pour sa technologie 43nm, prévue en 2010 [Tos08].

1.4 Les mémoires TANOS

1.4.1 Présentation de la technologie TANOS

Les difficultés pour maintenir un fort *coupling ratio* et pour réduire le couplage parasite entre les cellules adjacentes peuvent être contournées par l'intégration de mémoires à couche de piégeage discret. Ces mémoires sont donc envisagées par l'ITRS [ITR07] pour les nœuds technologiques inférieurs à $36nm$ car elles ne sont pas confrontées aux mêmes difficultés de miniaturisation que les mémoires à grille flottante continue, cf tableau 1.2

En particulier, la structure TANOS ($TaN/Al_2O_3/Si_3N_4/SiO_2/Si$) proposée par Samsung [Lee03] semble particulièrement adaptée aux mémoires NAND à forte densité pour les applications dédiées au stockage de données.

La cellule mémoire TANOS intègre du nitrure comme couche de piégeage, permettant la réduction de l'oxyde *tunnel* et d'être plus robuste à ces défauts dans ce dernier. De plus l'intégration d'un oxyde bloquant en *High- κ* permet un oxyde *tunnel* plus épais que pour les mémoires SONOS classiques³⁶. Enfin, la grille de contrôle en TaN , avec un travail de sortie de $4,8eV$, assure la réduction de la saturation de l'état effacé.

Année	2009	2010	2011	2012	2013	2014	2015
Nœud technologique F							
Flash NAND [nm]	40	36	32	28	25	22	20
Flash NOR (SONOS/NROM)[nm]	50	45	40	35	32	28	25
1 — Flash NAND – Couche de piégeage							
Endurance [cycles]	10 ⁵						
Nb max de bits par cellule	3	4					
Composition de l'oxyde <i>tunnel</i>	SiO_2 ou ONO						
EOT de l'oxyde <i>tunnel</i> [nm]	3 – 4						
Composition de l'oxyde de contrôle	SiO_2 ou Al_2O_3						
EOT de l'oxyde de contrôle [nm]	6 – 8						
Composition de la couche de piégeage	Si_3N_4						
Épaisseur de la couche de piégeage [nm]	5 – 7						4 – 6
Composition de la grille de contrôle	Poly-p/métal						Métal
2 — Flash NOR – SONOS/NROM							
EOT de l'oxyde <i>tunnel</i> [nm]	5	4,5			4		
Épaisseur de la couche de piégeage [nm]	5 – 7	4 – 6				4 – 5	
EOT de l'oxyde de contrôle [nm]	7 – 9	6 – 8				5 – 7	
Tension de programmation [V]	7 – 9	6 – 8				5 – 7	
Nb max de bits par cellule	4						6

TAB. 1.2: Résumé des objectifs technologiques pour les mémoires à couche de piégeage discret type NOR et NAND [ITR07].

³⁶Un oxyde *tunnel* de $\sim 2,5nm$ pour une SONOS classique contre $3 \sim 4nm$ pour une TANOS.

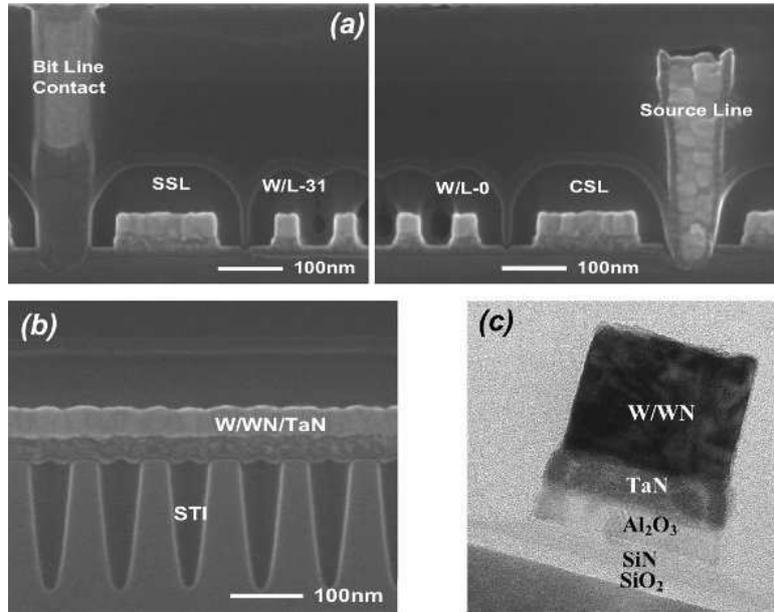


FIG. 1.25: Photographies MET a) d'une ligne de bit , b) d'une ligne de mot et c) une vue de coupe d'une mémoire TANOS pour le nœud technologique 63nm [Shi05].

La figure 1.25 regroupe des photographies MET d'une matrice mémoire TANOS pour le nœud technologique 63nm. On observe l'empilement TANOS, ainsi que les 32 cellules mémoires sur la ligne de bit et les transistors de sélection, caractéristiques d'une matrice NAND.

Ce dispositif a présenté des résultats intéressants. La figure 1.26 présente un ΔV_T de 6V pour une condition de programmation de 17V/100 μ s et d'effacement de -19V/10ms [Lee06]. De plus, pour la technologie 63nm, aucune interférence inter-cellule n'est observée. [Shi05] montre, qu'après 10^4 cycles d'écriture/effacement, aucun piégeage significatif dans l'alumine.

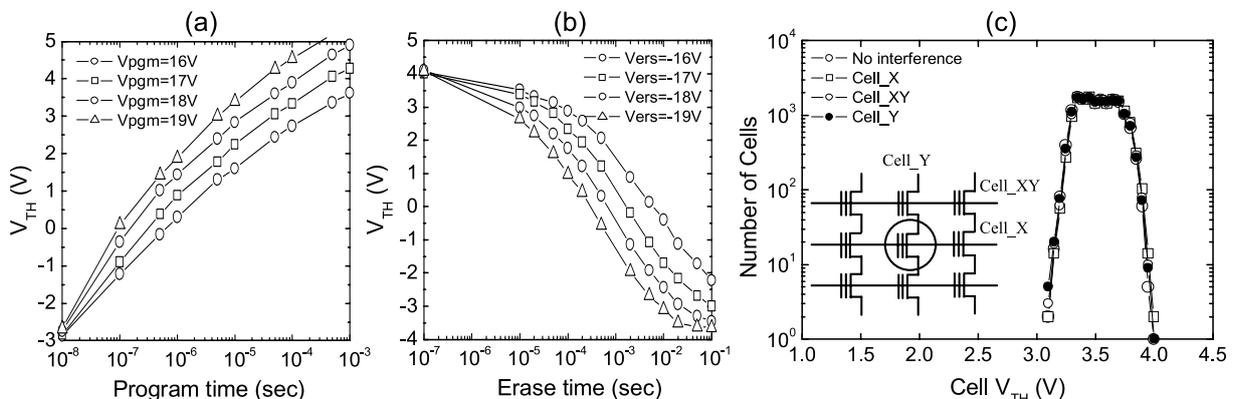
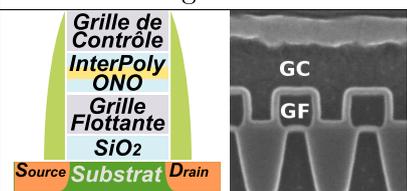
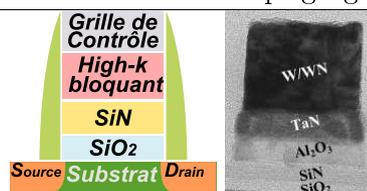


FIG. 1.26: Caractéristiques (a) d'écriture et (b) d'effacement typiques d'une cellule TANOS. (c) Variation de la tension de seuil en fonction de la programmation de cellules TANOS adjacentes pour le nœud technologique 63nm [Lee06].

	NAND à grille flottante	NAND à couche de piégeage
Structure		
Programmation :	injection d'électrons	injection d'électrons
Effacement :	injection d'électrons	injection de trous
Avantages	Meilleure rétention de l'information	Dispersion restreinte des V_T Plus prometteur pour la réduction des dimensions technologiques
Inconvénients	Dispersion anormale des V_T Moins prometteur pour la réduction des dimensions technologiques	Moins bonne rétention

TAB. 1.3: Résumé des propriétés des mémoires à grille flottante et à couche de piégeage [Cho09].

Le tableau 1.3 résume les différences entre les mémoires à grille flottante continue et les mémoires à couche de piégeage discret. On retiendra que les mémoires TANOS apparaissent comme plus propices à la poursuite de la miniaturisation, car elles sont plus résistantes au couplage inter-cellule et au SILC. De plus, la faible dispersion de leur V_T les rend plus attractives pour les applications à multi-niveaux (multi- V_T). Même si les mémoires TANOS offrent de grandes améliorations (oxyde tunnel plus épais, oxyde bloquant en *High- κ* et grille métallique) par rapport aux mémoires SONOS, elles souffrent d'une mauvaise rétention, défaut historique des SONOS³⁷. De plus, la technologie des mémoires à grille flottante continue est couramment intégrée dans les produits commerciaux et très bien contrôlée, alors qu'aujourd'hui aucun produit TANOS n'est commercialisé. Aussi, malgré que l'approche TANOS a initialement été prévue pour les générations technologiques inférieures à $50nm$ [Kim05], les approches classiques, à base de grille flottante polysilicium, subsistent toujours. Les mémoires TANOS sont maintenant envisagées pour les générations $20 \sim 30nm$.

Ainsi, il est encore nécessaire d'améliorer l'empilement mémoire pour améliorer l'effacement et la rétention. En effet, les fortes tensions de programmation dégradent fortement l'endurance et la rétention de ces dispositifs [Shi05]. Par exemple, [Par06] montre l'intérêt d'optimiser la couche d'alumine pour améliorer le comportement en endurance pour des mémoires de $40nm$. De même, des approches semi-cylindriques des TANOS ont montré des performances intéressantes pour limiter les effets canaux courts et améliorer les tensions de programmation [Kwa07].

³⁷Les mémoires SONOS doivent avoir une faible épaisseur d'oxyde *tunnel* ($\sim 2,5nm$) pour être effaçables ce qui les pénalisent grandement en rétention. De même, le nitrure s'active beaucoup, réduisant la rétention en température.

1.4.2 Difficultés des mémoires TANOS et leurs évolutions

Le problème historique des mémoires SONOS est leur mauvaise rétention. En effet, l'oxyde bloquant en oxyde de silicium et la couche de piégeage en nitrure nécessitent un oxyde *tunnel* fin $\sim 2,5nm$ pour pouvoir être effacés. Cette faible épaisseur est très pénalisante en rétention. Ainsi, de manière générale, les mémoires à couche de piégeage sont réputées pour avoir une mauvaise rétention ou être difficilement effaçables. Les mémoires TANOS, avec leur oxyde bloquant en alumine, permettent une nette augmentation de l'oxyde *tunnel* $3 \sim 4nm$ et leur grille métallique permet d'améliorer l'effacement. Mais le problème du couple effacement/rétention subsiste. Les principales solutions envisagées sont :

- L'ingénierie de bandes de l'empilement *tunnel*.
- L'ingénierie de la couche de piégeage.

1.4.2.1 L'ingénierie de bandes de l'empilement *tunnel*

L'ingénierie de bandes de l'empilement *tunnel* est envisagée pour améliorer l'effacement. Cette approche, appelée BE-SONOS³⁸, remplace usuellement l'oxyde *tunnel* par un ONO [Lue05]. De plus, l'intégration d'alumine en oxyde bloquant permet de réduire encore plus les temps d'effacement par rapport à une mémoire SONOS, cf figure 1.28 [Lai07].

L'objectif de l'ingénierie de bandes est d'ajuster les propriétés d'injection de l'oxyde *tunnel* en travaillant sur les matériaux qui le composent. Il s'agit de la même approche que les *Crested Barrier*, on recherche le plus fort courant à fort champ (pour l'écriture et l'effacement) et le plus faible à faible champ (pour la rétention).

L'approche la plus courante est de remplacer l'oxyde *tunnel* par un empilement tricouche Oxyde-Nitrure-Oxyde [Lai07]. Les deux couches de SiO_2 permettent la bonne isolation de l'empilement et le nitrure permet une meilleure sensibilité en champ. Mais il existe d'autres structures présentées dans la littérature à base :

- D'alumine, $SiO_2-2nm/Al_2O_3-8nm/HTO-2nm$ [Blo09a].
- De silicate d'hafnium, $SiO_2-1,9nm/HfSiON-3,6nm/HTO-4,1nm$ [Ver09].
- De bicouches, $SiO_2-2,9nm/Si_3N_4-1,6nm$ [Wan07, Lin08].
- De monocouches, $HfSiON$ [Van05, Van06b].

³⁸BE-SONOS : Bandgap Engineering SONOS.

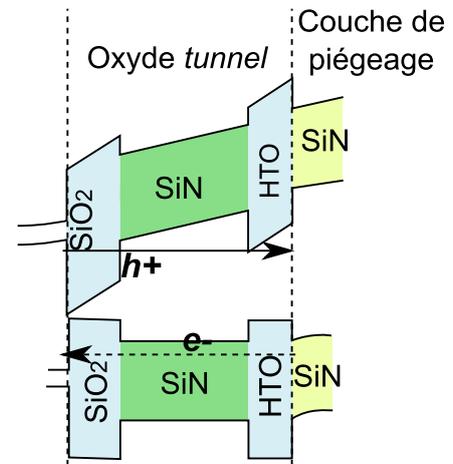


FIG. 1.27: Schémas de bandes d'une mémoire BE-SONOS en effacement et en rétention.

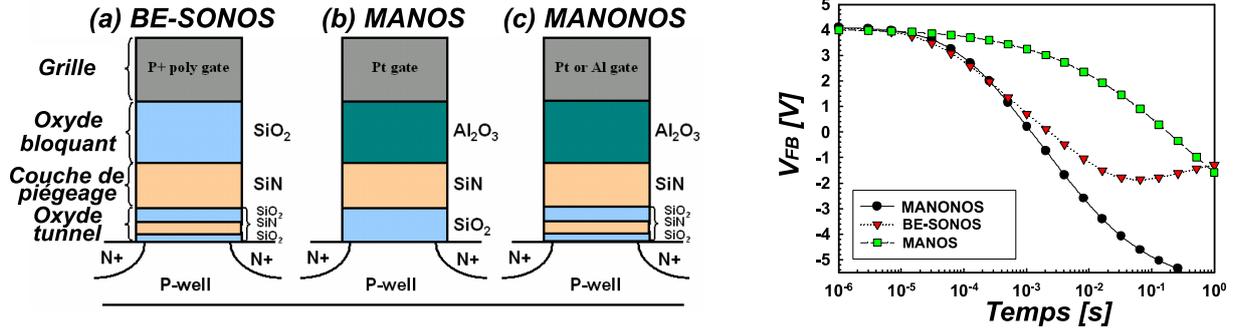


FIG. 1.28: Comparaison des caractéristiques d’effacement de mémoires SONOS (BE-SONOS), MANOS (Métal/Alumine/Nitrure/Oxyde) et MANONOS (Métal/Alumine/Nitrure/Oxyde/Nitrure/Oxyde) [Lai07].

1.4.2.2 L’ingénierie de la couche de l’empilement de grille

L’ingénierie de la couche de piégeage est étudiée par de nombreux groupes. Le tableau 1.4 regroupe quelques publications intégrant des *High- κ* comme couche de piégeage. On observe que l’ensemble de l’empilement de grille est optimisé, avec l’utilisation d’une grille métallique ou l’ajustement de l’oxyde bloquant et de l’oxyde *tunnel* en utilisant des matériaux *High- κ* . Le tableau 1.5 regroupe les caractéristiques mémoires de différents dispositifs présentés dans la littérature.

Les paramètres importants de la couche de piégeage sont :

- Une forte constante diélectrique pour limiter la chute de potentiel dans cette couche.
- Un important écart entre la bande de conduction de la couche de piégeage avec l’oxyde *tunnel* et l’oxyde bloquant pour réduire le courant de fuite en rétention.
- Des pièges profonds pour améliorer la rétention.

Oxyde <i>tunnel</i>	Couche de piégeage	Oxyde bloquant	Grille de contrôle	Institut	Publication
SiO_2	Si_3N_4	Al_2O_3	TaN	Samsung	[Shi05, Lee03]
SiO_2	$HfON$	$HfAlO$	TaN	Univ. Taiwan	[Lai06]
SiO_2	AlN	$HfAlO$	IrO_2	Univ. Taiwan	[Lai05]
SiO_2	$SiN, HfAlO, HfO_2, Al_2O_3$	$HfAlO, HfO_2, Al_2O_3, SiO_2$	HfN	Univ. Singapore	[Tan04] [Tan06]
SiO_2	$AlGaN$	$AlLaO_3$	TaN	Univ. Taiwan	[Chi05]
$HfAlO$	$HfSiO$	$HfAlO$	IrO_2	Univ. Singapore	[Wan05b]
HfO_2	Ta_2O_5	Al_2O_3, HfO_2	TaN	Univ. Texas	[Wan06, Wan04]
SiO_2	SiN	$HfO_2, HfSiON$	$Poly - Si$	NXP	[Van05]
SiO_2	SiN	Al_2O_3	$Poly - Si$	IMEC	[Cac07]
SiO_2	SiN	Al_2O_3	Pt	Macronix	EDL’07

TAB. 1.4: Description de publications intégrant des *High- κ* dans l’empilement de grille et comme couche de piégeage.

	Condition P/E rét. & end.	$\Delta V_{T_{init}}$ à 85 ° C	$\Delta V_T@10ans$ à 85 ° C	ΔV_T @cycles
<i>SiO₂-Si₃N₄-Al₂O₃-TaN</i> [Lee03] 2, 5 - 12 - 10 - 20	13, 5V 100μs -13V 10ms	4, 4V	2, 07V	4V@10 ⁵ à 85 ° C
FinFET <i>SiO₂-Si₃N₄-SiO₂</i> [Oh 04] 2, 3 - 9, 1 - 7, 1	13V 10μs -12V 1ms	5V	2, 9V	4, 2V@10 ⁴ à 85 ° C
<i>SiO₂-AlGaN-AlLaO₃-TaN</i> [Chi05] 2, 7 - 10 - 12 - 100	11V 100μs -11V 1ms	4, 8V	3, 3V	4V@10 ⁵ à 85 ° C
<i>SiO₂-AlN-HfAlO-IrO₂</i> [Lai05] 2, 8 - 12 - 13 - 50	13V 100μs -13V 1ms	5, 5V	3, 4V	4, 6V@10 ⁵ à 85 ° C
<i>HfAlO-HfSiO-HfAlO-IrO₂</i> [Wan05b] 6 - 6 - 11	12V 100μs -12V 100μs	2, 8V	2V	2, 3V@10 ⁵ à 25 ° C
<i>SiO₂-HfON-HfAlO-TaN</i> [Lai06] 2, 8 - 10 - 12 - 100	8V 100μs -8V 100μs	2, 5V	1, 45V	2, 1V
<i>SiO₂- NC Si-SiO₂</i> [Cho05] 2, 5 - 13 - 7, 5	14V 10ms -14V 100ms	4V	3, 5V	4, 5V@10 ⁵
Tri-gate <i>SiO₂-Si₃N₄-SiO₂</i> [Spe04]	11, 5V 3ms -11, 5V 100ms	1, 2V	1, 1V à 25 ° C	—
<i>SiO₂-LaAlO₃-HfON-Si₃N₄-LaAlO₃-SiO₂-TaN</i> [Lin08] 2, 5 - 2, 5 - 5 - 5 - 8 - 5 - 200	16V 100μs -16V 100μs	5, 6V	4, 1V	4, 9V@10 ⁵
<i>SiO₂-LaAlO₃-Si₃N₄-LaAlO₃-SiO₂-TaN</i> [Lin08] 2, 5 - 2, 5 - 5 - 8 - 5 - 200	16V 100μs -16V 100μs	3, 3V	2V	—

TAB. 1.5: Comparaisons des caractéristiques mémoires de différents dispositifs présents dans la littérature.

- Une forte densité de piège contrôlée pour avoir une grande fenêtre de programmation.

Ceux de l'oxyde bloquant sont :

- Une forte constante diélectrique pour assurer un bon couplage entre la grille et la zone de piégeage.
- Un grand gap pour réduire les courants de fuites en rétention et en programmation.
- Une faible densité de pièges pour limiter les courants assistés par pièges et assurer une bonne rétention.
- Une bonne compatibilité avec les autres matériaux.

Enfin, la grille de contrôle doit posséder une bonne compatibilité avec l'oxyde bloquant et un fort travail de sortie pour éviter l'injection d'électrons dans la couche de piégeage durant l'effacement.

Aussi des études sont menées sur l'optimisation du nitrure par le contrôle des concentrations en Si et en O. En effet, le *SiN* riche en silicium s'efface plus vite mais a une moins bonne rétention. Inversement, le *SiN* riche en oxygène réduit la fenêtre de programmation mais améliore la rétention, grâce à des pièges plus profonds [Van08, Cho08]. De même, l'épaisseur de la couche de piégeage peut impacter les propriétés de rétention [Mel08].

La figure 1.29 présente une évolution possible des empilements mémoires avec l'utilisation d'un bicouche *SiO₂/LaAlO₃* pour l'oxyde *tunnel* et l'oxyde bloquant. L'utilisation d'une couche de piégeage *Si₃N₄/HfON* augmenterait la profondeur des pièges.

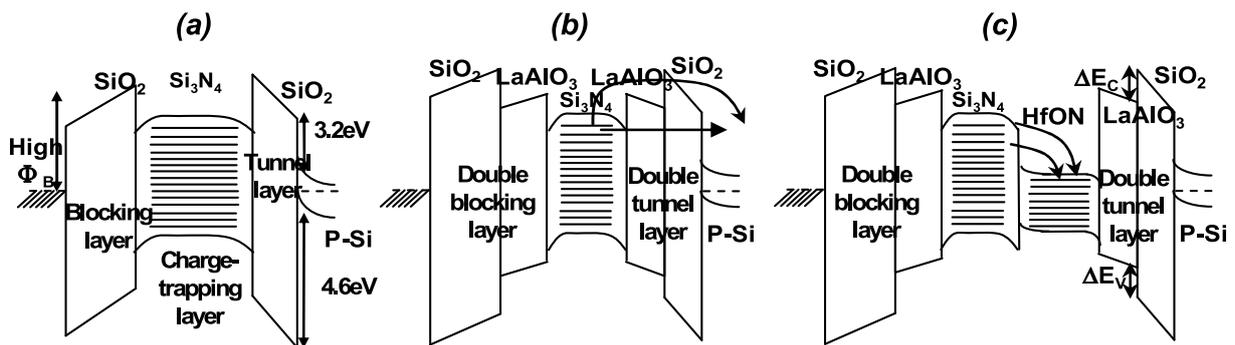


FIG. 1.29: Ingénierie de bandes proposée par [Lin08] pour améliorer l'effacement et la rétention grâce à l'optimisation de l'oxyde bloquant : $SiO_2/LaAlO_3$, de l'oxyde tunnel : $SiO_2/LaAlO_3$, et de la couche de piégeage $Si_3N_4/HfON$.



En conclusion les mémoires TANOS possèdent des caractéristiques intéressantes pour la réduction des dimensions et la poursuite de la *road map* des mémoires non-volatiles. La principale difficulté est de maintenir une bonne rétention en gardant un effacement rapide. L'ingénierie de bandes est une solution intéressante pour résoudre ce problème. Néanmoins, les mémoires TANOS représentent une rupture technologique importante que peu d'industriels sont prêts à franchir.

1.5 Conclusion

Comme nous venons de le voir les mémoires non-volatiles EEPROM-Flash sont confrontées à de nombreuses difficultés qui nécessitent une rupture technologique et l'utilisation de nouvelles approches. L'élément déclenchant reste la suppression des extensions latérales de la grille de contrôle autour de la grille flottante, imposant l'utilisation de *High- κ* pour le diélectrique d'interpoly. Cela a rendu plus attractif les mémoires à couche de piégeage discret pour les générations avancées, par exemple la TANOS. De plus, l'intégration de matériaux *High- κ* est un voie intéressante pour réduire les tensions de programmation pour les applications embarquées [Van06a, Pow08].

Les matériaux *High- κ* sont donc très étudiés, notamment l'alumine. Ainsi, de nombreuses publications présentent des études poussées sur son optimisation [Shu09], sur l'identification de ses défauts [Kit09, Ren02], sur la caractérisation de son courant de fuite [Beu09] ... D'autres travaux se sont intéressés à l'augmentation des propriétés d'injection soit par effet géométrique [Now08], soit par ingénierie de bandes [Lin08].

On notera aussi l'abondance des publications s'intéressant aux mémoires à couche de piégeage discret, notamment les empilements TANOS (*TaN / Al₂O₃ / Si₃N₄ / SiO₂ / Si*). En effet, la TANOS regroupe les intérêts des mémoires à couche de piégeage discret³⁹, des diélectriques *High- κ* ⁴⁰ et des grilles métalliques⁴¹. Ainsi, beaucoup de recherches sont effectuées sur la cellule TANOS pour améliorer ses performances [Kan07, Beu09] en optimisant, par exemple, la couche de piégeage en travaillant sur la stochiométrie [Van08, Cho08], l'épaisseur [Mel08] ou sur la structure de la couche [Huo07].

De nouvelles voies sont aussi fortement envisagées, comme par exemple la famille des mémoires résistives. Les PCRAM, mémoires à changement de phase, sont prometteuses. En effet, le confinement accélère leur programmation. De plus, elles peuvent être intégrées dans des structures *crossbar* permettant une grande densité d'intégration et des empilements 3D. Leurs performances électriques sont bonnes, notamment en écriture et en effacement. Néanmoins, cette technologie n'est pas encore mature (problèmes de fiabilité ...) pour des applications commerciales.

Dans ce contexte, nous allons nous intéresser aux différentes couches de l'empilement mémoire. Nous présenterons des analyses physico-chimiques et électriques des matériaux *High- κ* déposés au CEA-Léti (*HfO₂, HfAlO, Al₂O₃, HfSiO...*), en vue d'estimer leurs intérêts en tant que diélectrique d'interpoly. Puis nous étudierons ces matériaux *High- κ* une fois intégrés dans des dispositifs mémoires. Nous essaierons d'évaluer leurs impacts sur les caractéristiques mémoires des cellules à nanocristaux. Nous présenterons aussi une étude approfondie des mémoires à base de nitrure, pour comprendre les mécanismes physiques et permettre leurs améliorations.

³⁹Résistance au SILC, plus grande immunité au couplage inter-cellule.

⁴⁰Permettant le bon contrôle électrostatique de la couche de piégeage.

⁴¹Pour réduire la saturation de l'état effacé grâce à une grille à fort travail de sortie en limitant l'injection des électrons de la grille de contrôle vers la zone de piégeage durant l'effacement

Chapitre 2

Modélisation du courant et du piégeage dans les diélectriques

Dans ce chapitre, nous allons d'abord rappeler les différents modes de transport dans les diélectriques. Puis nous présenterons différentes approches pour modéliser le piégeage de charge dans ces couches. Enfin, en s'appuyant sur ces notions, nous décrirons deux modèles de piégeage pour les mémoires à couche de piégeage en nitrure que nous utiliserons, dans les chapitres suivants, pour des programmations Fowler-Nordheim, afin de mieux comprendre les mécanismes d'écriture, d'effacement et de rétention.

Sommaire

2.1	Conduction électronique dans les diélectriques	59
2.1.1	Conduction <i>tunnel</i> direct et Fowler-Nordheim	60
2.1.1.1	Expressions générales	60
2.1.1.2	Etats électroniques liés et continuum d'états	62
2.1.1.3	Expression des courants 2D	64
2.1.1.4	Expression des courants 3D	65
2.1.1.5	Courant d'émission d'une couche de piégeage	66
2.1.1.6	Calcul de la transparence	67
2.1.2	Modes de conduction assistée par des défauts	69
2.1.2.1	Courant Poole-Frenkel	70
2.1.2.2	Courant Poole	72
2.1.2.3	Identification des courants Poole et Poole-Frenkel	73
2.1.2.4	Courant TAT : Trap Assisted Tunneling	73
2.2	Mécanismes de piégeage dans les diélectriques	75
2.2.1	Les différentes approches pour la modélisation d'un piège	75
2.2.2	Modèle d'états indépendants	76
2.2.3	Aspect amphotérique des pièges	77
2.2.4	Modélisation des coefficients de piégeage	78

2.2.4.1	Échanges entre la BC et le piège du diélectrique	79
2.2.4.2	Échanges entre la BC du silicium et le piège	81
2.3	Modèles de piégeage pour les mémoires de type SONOS	83
2.3.1	État de l'art de quelques modèles SONOS	83
2.3.2	Modèle 1 : prise en compte de pièges amphotériques	86
2.3.3	Modèle 2 : prise en compte du transport dans la couche de piégeage	88
2.4	Conclusion	91

2.1 Conduction électronique dans les diélectriques

Dans cette partie nous allons présenter les différents modes de conduction présents dans les diélectriques. Nous commencerons par la conduction assistée par le champ électrique classiquement observée dans les oxydes fins, par exemple l'oxyde *tunnel* d'une mémoire Flash. Puis nous verrons les mécanismes assistés par des pièges présents dans les isolants.

La figure 2.1 résume les différents modes de conduction électroniques les plus courants dans les diélectriques. Ils peuvent être classés en deux catégories :

- Figure 2.1a, la conduction est limitée par l'injection. Pour ces modes de conduction le mécanisme limitant est l'injection des électrons dans le diélectrique. Ces courants sont donc contrôlés par la densité de porteurs pouvant traverser l'oxyde et leur probabilité de passage. Les mécanismes les plus courants sont :

- (1) **Émission thermoïonique** : les électrons dont l'énergie est très élevée et qui passent directement au-dessus de la barrière de potentiel du diélectrique, dans sa bande de conduction. Ce mécanisme apparaît à très haute température.
- (2) **Conduction Fowler-Nordheim** : les électrons traversent une barrière de potentiel triangulaire. L'épaisseur de l'oxyde n'intervient qu'à travers le champ électrique. C'est un mécanisme principalement contrôlé par le champ électrique dans le diélectrique. La température n'impacte que les densités d'états dans les électrodes.
- (3) **Conduction *tunnel* direct** : les électrons traversent une barrière de potentiel trapézoïdale. L'épaisseur de l'oxyde intervient donc fortement car elle détermine la distance que devrait parcourir les électrons dans le gap du diélectrique, contrairement à la conduction Fowler-Nordheim. Mais ce mécanisme possède les mêmes dépendances en champ et en température que celui-ci.
- (4) **Conduction *tunnel* des électrons de la bande de valence du substrat** : ce mode de conduction se produit lorsque la quantité d'états disponibles et la charge d'injection sont importantes.

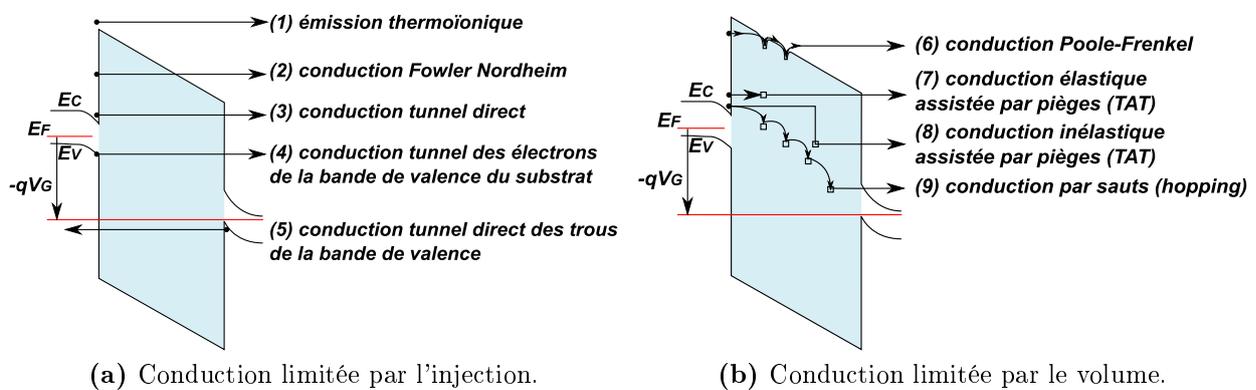


FIG. 2.1: Principaux modes de conduction à travers un diélectrique.

- (5) **Conduction *tunnel* direct des trous de la bande de valence** : les trous traversent la barrière de potentiel de l'oxyde. Il s'agit du même mécanisme que l'injection *tunnel* direct mais pour les trous.
- Figure 2.1b, la conduction est limitée par le volume. Pour ces modes de conduction [Geh04], l'injection d'électrons dans le diélectrique n'est pas le mécanisme limitant. Le transport est contrôlé par la conduction des porteurs dans le volume du diélectrique et est lié à la présence d'états localisés dans le gap du diélectrique, aussi appelés pièges :
- (6) **Conduction Poole et Poole-Frenkel** : les électrons passent de piège en piège par émission thermoïonique. C'est un mécanisme fortement activé en température et en champ électrique, le champ diminuant la barrière que doivent franchir les électrons.
- (7) **Conduction TAT¹ élastique** : les électrons traversent l'oxyde par effet *tunnel* par l'intermédiaire d'un piège sans perte d'énergie². Ce mécanisme possède les mêmes dépendances en température et en champ électrique que les conceptions *tunnel*.
- (8) **Conduction TAT inélastique** : les électrons traversent l'oxyde par effet *tunnel* par l'intermédiaire d'un piège avec perte d'énergie. Généralement, cette perte d'énergie se traduit par l'émission de phonons.
- (9) **Conduction par sauts (Hopping)** : les électrons passent de piège en piège par effet *tunnel*. Ce sont des mécanismes proches des conceptions TAT multisite mais qui nécessitent une grande densité de défauts.

2.1.1 Conduction *tunnel* direct et Fowler-Nordheim

Dans cette partie, nous présenterons les mécanismes de l'effet *tunnel*. Par souci de clarté, nous nous cantonnerons à l'étude du courant des électrons de la bande de conduction. Mais le formalisme proposé est le même pour les courants de trous ou les courants d'électrons de la bande de valence, seules les hauteurs de barrières, les masses effectives ou les polarisations changent.

2.1.1.1 Expressions générales

La figure 2.2 représente une structure MOS, Métal/Oxyde/Silicium. L'oxyde se comporte comme une barrière de potentiel pour les porteurs. En l'absence de champ électrique, la barrière est rectangulaire et la hauteur vue par les électrons du silicium est ϕ_B . Lorsque l'on applique une polarisation sur la grille, la barrière est modifiée³ et des électrons peuvent

¹TAT : Trap Assisted Tunneling.

²Le piège participant à la conduction a donc la même énergie que l'électron.

³En absence de charges dans le diélectrique, le champ est constant et le potentiel décroît linéairement.

Rappel : $\frac{\partial V_x^2}{\partial z^2} = -\rho$, où ρ densité volumique de charge [C/m^3] et V_x la différence de potentiel suivant x .

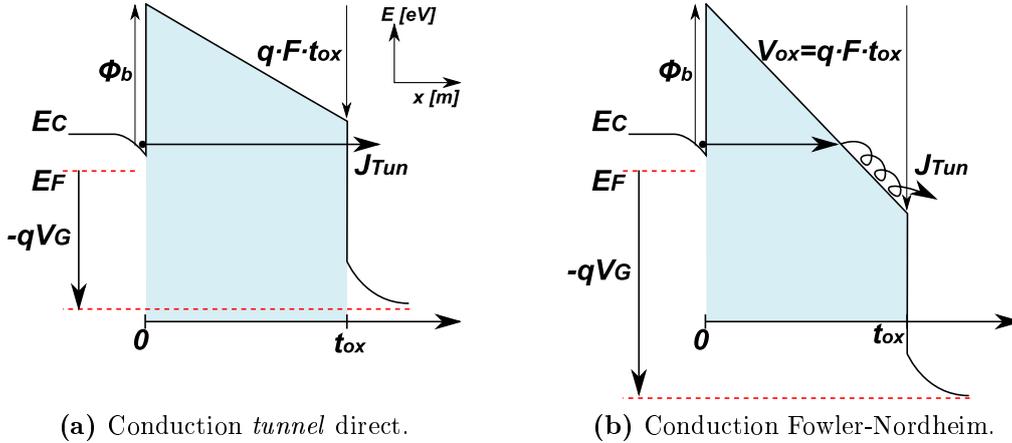


FIG. 2.2: Mécanismes de conduction (a) *tunnel* direct et (b) Fowler-Nordheim.

transiter par effet *tunnel* à travers l'oxyde.

On peut distinguer deux régimes d'injection de porteurs suivant la chute de potentiel aux bornes du diélectrique V_{ox} .

$V_{ox} < \phi_B$, J_{TD} , **barrière trapézoïdale, figure 2.2a** : les électrons pénètrent directement dans l'anode. La transition est considérée isoénergétique ou élastique⁴. C'est le *tunnel* direct.

$V_{ox} > \phi_B$, J_{FN} , **barrière triangulaire, figure 2.2b** : les électrons débouchent dans la bande de conduction du diélectrique. C'est le régime Fowler-Nordheim. Une fois sur la bande de conduction, les électrons peuvent subir des collisions avec les phonons de l'oxyde, et se relaxent.

On notera que le libre parcours moyen des électrons dans le SiO_2 est estimé entre $2nm$ à $4nm$ [Bri97, Cha84]. Pour les mémoires Flash, on peut considérer que l'ensemble des électrons injectés sera thermalisé sur la grille flottante, ce qui n'est pas le cas des mémoires de type SONOS [Fur07, Lee08].

Il est possible d'exprimer les courants *tunnel* sous la forme⁵ :

$$J_{TD/FN} = \int q \cdot T(E_{\perp}) \cdot N(E_{\perp}) dE_{\perp} \quad (2.1)$$

- $T(E_{\perp})$ [-] : Transparence du diélectrique à traverser à l'énergie E_{\perp} .
 $N(E_{\perp})$: Densité d'état d'électrons multipliée par leur vitesse normale à l'interface de l'oxyde à l'énergie E_{\perp} .
 $[m^{-2}s^{-1}J^{-1}]$
 E_{\perp} [J] : l'énergie des électrons normale au sens de propagation du courant.

Dans les applications mémoires, les empilements de grilles sont souvent constitués de plusieurs couches. Cette barrière de potentiel complexe entraîne de multiples configurations. Ainsi, une ingénierie de bandes est possible pour optimiser les caractéristiques re-

⁴Pas de perte d'énergie, pas d'interaction entre électron et oxyde.

⁵En supposant que le nombre d'états d'arrivés ne limite pas l'injection.

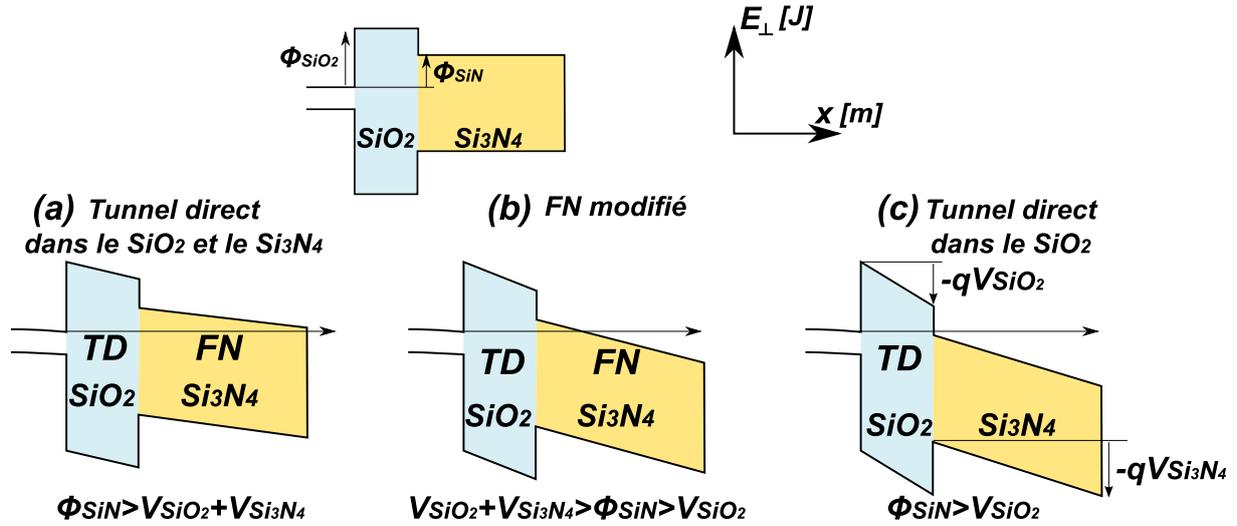


FIG. 2.3: Diagrammes de bandes pour trois polarisations. a) *Tunnel Direct* d'électrons dans le SiO_2 et le Si_3N_4 . b) *Tunnel Direct* d'électrons dans le SiO_2 et Fowler-Nordheim dans le Si_3N_4 . c) *Tunnel Direct* d'électrons dans le SiO_2 et bande de conduction du Si_3N_4 .

cherchées pour chaque couche : isolation, injection, confinement, ... Une configuration classique dans les mémoires SONOS comprend trois diélectriques, $SiO_2/Si_3N_4/SiO_2$. Ainsi, différentes conditions d'injection peuvent exister [Bac01] . Elles sont illustrées sur la figure 2.3.

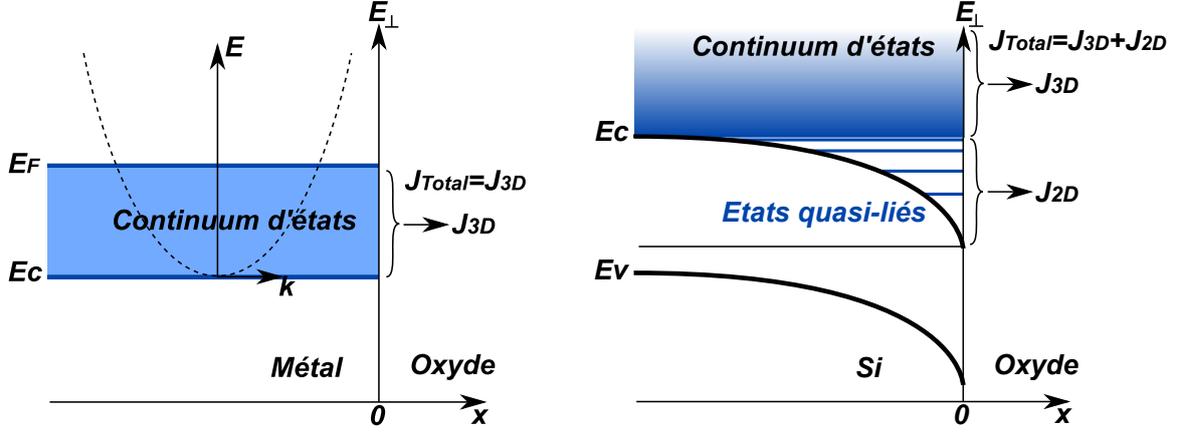
La plus classique est l'injection Fowler-Nordheim modifiée, où dans le cas de deux diélectriques (par exemple SiO_2/Si_3N_4), les électrons voient une barrière de potentiel trapézoïdale dans le SiO_2 (injection de type *tunnel direct*) et triangulaire dans le Si_3N_4 (injection de type Fowler-Nordheim).

2.1.1.2 Etats électroniques liés et continuum d'états

La forme de la densité d'états $N(E_{\perp})$ introduite dépend de la nature semi-conducteur ou métallique de l'électrode considérée. Dans le cas d'une cathode métallique, le courant *tunnel* est exclusivement issu d'un continuum d'états (cf figure 2.4a). Pour une cathode en silicium, le puits de potentiel à l'interface Si/oxyde, induit une quantification des états. Ces états liés introduisent une densité de courant dit 2D, J_{2D} (cf figure 2.4b). Pour les niveaux d'énergie électronique situés au-dessus du puits, il y a une contribution supplémentaire d'états continus. Ainsi, nous proposons de décomposer $N(E_{\perp})$ selon les deux contributions :

$$N(E_{\perp}) = N_{cont}(E_{\perp}) + N_{liés}(E_{\perp})$$

Où N_{cont} correspond au continuum de niveaux électroniques et $N_{liés}$ aux niveaux électroniques liés. Ainsi, pour une cathode métallique, $N_{liés} = 0$.



(a) Pour une grille métallique, le courant *tunnel* J_{Total} est dû à des électrons d'états faisant partie d'un continuum et transitant par effet *tunnel* à travers l'oxyde.

(b) Pour un substrat Si, le courant *tunnel* J_{Total} est dû aux électrons provenant d'états quasi-liés J_{2D} et du continuum d'états J_{3D} .

FIG. 2.4: États électroniques pour un métal et pour un substrat de silicium.

Pour le continuum d'états, dans le cadre d'une approximation de bande parabolique⁶, d'un gaz 3D d'électrons et d'une distribution de Fermi-Dirac [Cha84], on peut écrire :

$$N_{cont}(E_{\perp}) = \int_0^{\infty} \frac{4\pi m_{Si}}{h^3} f(E_{\perp} + E_{//}) dE_{//} = \frac{4\pi m_{Si}}{h^3} kT \cdot \ln \left[1 + \exp \left(-\frac{E_{\perp} - E_f}{kT} \right) \right] \quad (2.2)$$

- E_f [J] : Niveau de Fermi dans l'électrode considérée.
- m_{Si} [kg] : Masse effective des électrons dans l'électrode dans la direction normale à l'interface (dans le sens de propagation du courant *tunnel*).
- E_{\perp} [J] : Énergie des électrons de vitesse normale au sens de propagation du courant.
- $E_{//}$ [J] : Énergie des électrons de vitesse parallèle au sens de propagation du courant.
- $f(E)$ [-] : Fonction de fermie.

Pour les états liés, il est possible de leur assimiler une sous-bande distincte. Le courant total associé à ces états est donc la somme des courants d'électrons de chaque sous-bande [Shi98, Yan99]. Ainsi, la densité d'états liée est la somme des états sous chaque sous-bande :

$$N_{liés}(E_{\perp}) = \sum_i n(E_{\perp}^i) \cdot F_{imp}(E_{\perp}^i) \quad (2.3)$$

- $n(E_{\perp}^i)$ [m^{-2}] : Densité de porteurs sur la sous-bande d'énergie E_{\perp}^i .
- $F_{imp}(E_{\perp}^i)$ [Hz] : Fréquence d'impact des porteurs sur leur sous-bande.

Plusieurs approches existent pour obtenir les énergies de ces différents niveaux. Pour le calcul du niveau fondamental, correspondant à la sous-bande de plus basse énergie [Mat04,

⁶en fonction de k , le vecteur d'onde, cf figure 2.4a.

[Cle01, Gar04] utilisent l'approche dite variationnelle. Elle considère que la fonction d'onde associée aux électrons est approchée par une expression analytique unique, associée à toute la charge électronique d'inversion :

$$E_{fond} = \frac{3\hbar}{4\pi} \left[\frac{q^2 m_{Si}}{4\hbar \varepsilon_{Si}^2} (Q_{Cath} + 2Q_{Dep})^2 \right]^{1/3} \quad (2.4)$$

- q [C] : Charge de l'électron.
 m_{Si} [kg] : Masse de l'électron dans le Si dans le sens de propagation du courant *tunnel*.
 ε_{Si} [F/m] : Constante diélectrique du Si.
 Q_{Cath} [C/m²] : Charge dans la cathode.
 Q_{Dep} [C/m²] : Charge de désertion dans le silicium, $Q_{Dep} \sim -\sqrt{2 \cdot \varepsilon_{Si} \cdot q \cdot N_A \cdot \varphi_F}$
 avec N_A nombre de dopants [m⁻³] et $\varphi_F \sim \frac{kT}{q} \ln \left(\frac{N_A}{n_i} \right)$

Q_{Cath} peut être calculée de manière pratique en intégrant la caractéristique C_V de la structure.

La fréquence d'impact est directement proportionnelle à l'énergie du fondamental [Cle01] :

$$F_{imp}^{fond} = \frac{4E_{fond}}{3\hbar} \quad (2.5)$$

On notera que pour un silicium normalement dopé ($10^{17} \sim 10^{18} \text{cm}^{-3}$), la fréquence d'impact est le l'ordre de $F_{imp}^{fond} \sim 10^{13} \text{Hz}$.

2.1.1.3 Expression des courants 2D

Le courant 2D correspond à celui généré par les électrons des états liés (cf figure 2.4b). Dans un régime d'inversion ou d'accumulation forte⁷, la charge totale dans la cathode est principalement portée par le niveau fondamental⁸. Ainsi, d'après [Shi98, Yan99], la contribution de ce niveau au courant vaut :

$$J_{Tun}^{2D} = Q_{Cath} \cdot F_{imp}^{fond} \cdot T(E_{fond}) \quad (2.6)$$

⁷Cas où le courant 2D devient prédominant sur le courant 3D

⁸En fait, Il apparaît que les deux premiers niveaux contribuent au courant *tunnel*. Le niveau fondamental a plus de charges, une fréquence d'impact plus forte mais une transparence plus faible, alors que le deuxième niveau a moins de charges, une fréquence d'impact plus faible mais une transparence plus forte. Ainsi, en considérant que toute la charge est sur le niveau fondamental, on obtient une bonne approximation du courant *tunnel*.

2.1.1.4 Expression des courants 3D

Pour les courants 3D, il n'y a plus de quantification dans le substrat de silicium ou pour une grille métallique. Ainsi, le courant peut s'écrire :

$$J_{Tun}^{3D} = \int_{E_{cl}}^{\infty} q \cdot T(E_{\perp}) \cdot N_{cont}(E_{\perp}) dE_{\perp} \quad (2.7)$$

Où E_{cl} est le niveau énergétique à partir duquel on peut considérer un continuum d'états. En très faible inversion/accumulation ou pour un métal, il peut être pris égal à E_c . En pratique, [Wei77] prend comme E_{cl} le deuxième niveau quantifié. Ainsi, l'équation 2.8 donne une expression simplifiée du courant 3D à partir des équations 2.2 et 2.7.

$$J_{Tun}^{3D} = \frac{4\pi \cdot q \cdot m_{Si} \cdot kT}{h^3} \int_{E_{cl}}^{\infty} T(E_{\perp}) \ln \left[1 + \exp \left(\frac{Ef - E_{\perp}}{kT} \right) \right] dE_{\perp} \quad (2.8)$$

Pour être plus rigoureux il faudrait prendre en compte le nombre d'états disponibles côté anode. Dans ce cas, on obtient :

$$J_{Tun}^{3D} = \frac{4\pi \cdot q \cdot m_{Si} \cdot kT}{h^3} \int_{E_{cl}}^{\infty} T(E_{\perp}) \ln \frac{\left[1 + \exp \left(\frac{Ef - E_{\perp}}{kT} \right) \right]}{\left[1 + \exp \left(\frac{Ef - E_{\perp} - qV_G}{kT} \right) \right]} dE_{\perp} \quad (2.9)$$

Il existe d'autres expressions du courant *tunnel*. On notera l'approche analytique de [Pan95] donnant l'expression du courant 3D à $T=0$ °K dans l'équation 2.10 ainsi qu'un développement pour la dépendance en température (cf [Pan95] – équation 13).

$$J_{Tun}^{3D}(T = 0K) \sim A \cdot F^2 \exp \frac{-B}{F} \quad (2.10)$$

- A : $\frac{m_e^{Si} \cdot q^3}{8\pi \cdot h \cdot m_e^{ox} \cdot \phi_b}$.
 B : $\frac{8\pi \sqrt{2m_e^{ox}}}{3 \cdot h \cdot q} \left[\phi_b^{3/2} - (\phi_b - qt_{ox}F)^{3/2} \cdot H(\phi_b - qt_{ox}F) \right]$.
 F [V/m] : Champ dans l'oxyde.
 t_{ox} [m] : Épaisseur de l'oxyde .
 ϕ_b [J] : Hauteur de barrière vue par les électrons de la cathode.
 m_e^{ox} [kg] : Masse effective des électrons dans l'oxyde.
 m_e^{Si} [kg] : Masse effective des électrons dans la cathode (ici, en silicium) dans le sens de propagation du courant *tunnel*.
 $H(x)$: Fonction de Heaviside, définie comme : $H(x < 0) = 0$ $H(x \geq 0) = 1$.

2.1.1.5 Cas particulier d'un courant d'émission d'une couche de piégeage

Dans le cas de mémoires de type SONOS, où les électrons sont stockés dans un diélectrique, usuellement le Si_3N_4 , nous allons être amenés à modéliser le courant des électrons partant de la zone de stockage de charge. Cette expression du courant devra être fonction de la densité de charge sur la bande de conduction du nitrure⁹ : $n [m^{-3}]$. Nous pouvons écrire ce courant sortant sous la forme de l'équation 2.11. Ainsi, on définit v_T (cf équation 2.12) comme étant la vitesse moyenne pondérée par la densité de porteurs ($\propto f_{\perp}(E_{\perp})$) et la probabilité de passage ($T(E_{\perp})$).

$$J_{out} = q \cdot n \cdot v_T \quad (2.11)$$

$$v_T = \frac{\int_0^{\infty} f_{\perp}(E_{\perp}) v_{\perp}(E_{\perp}) T(E_{\perp}) dE_{\perp}}{\int_0^{\infty} f_{\perp}(E_{\perp}) dE_{\perp}} \quad (2.12)$$

$v_{\perp} : \sqrt{\frac{2E_{\perp}}{m_{\perp}}}$ la vitesse [m/s], pour les électrons d'un gaz 3D, normale au plan d'injection à l'énergie E_{\perp} [Via08, Via09b].

On notera que $E_{\perp} = 0$ correspond au bas de la bande de conduction du nitrure, soit E_C .

De plus, dans le cas d'un diélectrique, nous pouvons supposer que le niveau de Fermi est loin des bandes de conduction et de valence, l'approximation de Boltzmann est donc justifiée ($f_{\perp} = N_C \cdot \exp\left(-\frac{E_{\perp} - E_{f_{Si_3N_4}}}{kT}\right)$). Ainsi, nous obtenons les expressions :

$$v_T = \frac{\int_0^{\infty} e^{-\frac{E_{\perp}}{kT}} \sqrt{\frac{2E_{\perp}}{m_{\perp}}} \cdot T(E_{\perp}) dE_{\perp}}{\int_0^{\infty} e^{-\frac{E_{\perp}}{kT}} dE_{\perp}} \quad (2.13)$$

$$v_T \approx \sqrt{\frac{2}{m_{\perp}}} \cdot \frac{\int_0^{\infty} e^{-\frac{E_{\perp}}{kT}} \cdot T(E_{\perp}) dE_{\perp}}{\int_0^{\infty} \frac{e^{-\frac{E_{\perp}}{kT}}}{\sqrt{E_{\perp}}} dE_{\perp}}$$

$$v_T \approx v_{T1} = \sqrt{\frac{2}{m_{\perp} \cdot \pi \cdot kT}} \cdot \int_0^{\infty} e^{-\frac{E_{\perp}}{kT}} \cdot T(E_{\perp}) dE_{\perp} \quad (2.14)$$

Il est possible d'aller plus loin dans la simplification en utilisant la même approximation que dans [De 01]¹⁰ et obtenir cette expression :

$$v_T \approx v_{T2} = \sqrt{\frac{2kT}{m_{\perp} \cdot \pi}} \cdot T(E_C) \quad (2.15)$$

⁹En effet, comme nous allons le montrer dans la partie 2.3, pour modéliser une mémoire de type SONOS, nous calculons la densité d'électrons sur la bande de conduction du nitrure. Ainsi pour exprimer le courant sortant de notre structure, il nous faut un courant dépendant de n , densité d'électrons.

¹⁰Cette approximation considère que l'énergie des électrons est comprise entre E_C et $E_C + kT$, soit $\int_{E_C}^{\infty} e^{-\frac{E_{\perp}}{kT}} \cdot T(E_{\perp}) dE_{\perp} \sim T(E_C) \cdot kT$.

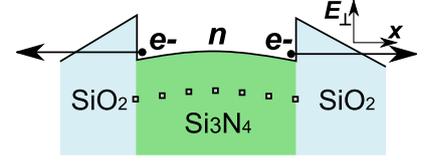


FIG. 2.5: Diagramme de bandes d'une structure ONO pouvant émettre des électrons.

- m_{\perp} [kg] : Masse effective des électrons dans le nitrure (diélectrique de départ) dans la direction de l'injection.
- E_c [J] : Niveau de la bande de conduction du nitrure.

La figure 2.6a compare les différentes approximations de la vitesse moyenne pondérée. Le diélectrique traversé est un *HTO* de $10nm$ et la cathode est en nitrure. On observe le même comportement pour les différentes approximations. La figure 2.6b représente l'erreur relative. On remarque qu'elle n'est pas négligeable, mais à fort champ elle est constante, elle pourra donc être compensée par les paramètres d'injection (hauteurs de barrières, masses effectives)

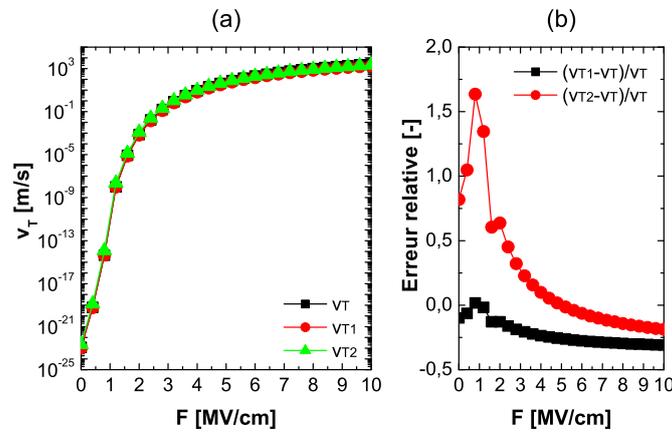


FIG. 2.6: Comparaison entre les différentes approximations de la vitesse moyenne pondérée : (a) valeur de la vitesse et (b) erreur relative. Diélectrique de cathode : Si_3N_4 , $\Delta E_C = 2eV$, $m_{\perp} = 0,5m_0$. Oxyde bloquant à traverser : *HTO*, $\Delta E_C = 2,8eV$, $m_{\perp} = 0,48m_0$, $t_{ox} = 10nm$.

2.1.1.6 Calcul de la transparence

La transparence d'une barrière de potentiel est définie comme la probabilité de passage d'un électron à travers celle-ci. Sa valeur est forcément dans l'intervalle $[0:1]$:

- $T=1$ correspond à une probabilité de passage de 100%, soit l'absence de barrière.
- $T=0$ correspond à une probabilité de passage de 0%, soit une barrière infranchissable.

Le passage *tunnel* d'un électron à travers une barrière, correspond à la pénétration de sa fonction d'onde dans celle-ci. Le calcul de cette probabilité se réalise à partir de la fonction d'onde, solution de l'équation de Schrödinger pour le profil de potentiel associé à la barrière en question. Pour une forme de barrière de potentiel quelconque, la solution de l'équation de Schrödinger est rarement analytique, elle est donc difficilement utilisable.

Nous présenterons ici la méthode de calcul par approximation WKB (Wentzel-Kramers-Brillouin). C'est une approche simple et efficace, mais elle ne prend pas en compte les réflexions électroniques aux interfaces des diélectriques. D'un point de vue pratique, il est

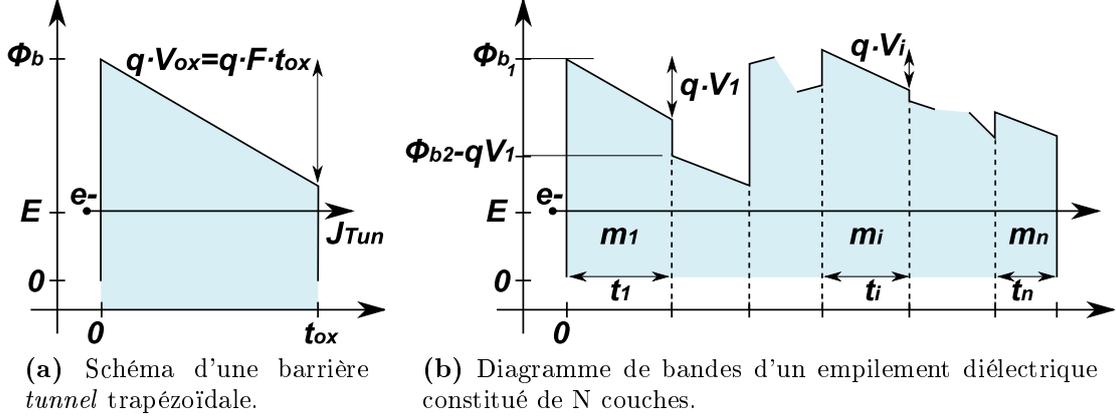


FIG. 2.7:

possible de prendre en compte ces réflexions par un facteur correctif [Coi09]. De plus, la non prise en compte de ce phénomène sera compensée par les paramètres utilisés (hauteurs de barrières et masses effectives).

Dans son principe, l'approximation WKB s'applique lorsque l'on considère que les longueurs d'onde associées aux fonctions d'onde des porteurs sont petites devant la dimension caractéristique des variations de la barrière de potentiel à traverser. Cela revient à dire que ce potentiel varie lentement dans l'espace. Cette approche correspond au cas dit quasi-classique.

Nous considérons d'abord une seule barrière *tunnel* (cf figure 2.7a). Dans ce cas, nous pouvons exprimer la transparence comme :

$$T_{WKB} = \exp\left(\frac{-2\sqrt{2}}{\hbar}U(E, V_{ox})\right) \quad (2.16)$$

Avec $U(E, V_{ox})$ la fonction énergie potentielle dont la définition est :

$$U(E, V_{ox}) = \int_0^{t_{ox}} \sqrt{m_{ox}(\phi_b - q \cdot V_{ox} - E) \cdot H(\phi_b - q \cdot V_{ox} - E)} dx \quad (2.17)$$

$$U(E, V_{ox}) = \frac{2\sqrt{m_{ox}t_{ox}}}{3V_{ox}} \left[(\phi_b - E)^{\frac{3}{2}} H(\phi_b - E) - (\phi_b - qV_{ox} - E)^{\frac{3}{2}} H(\phi_b - qV_{ox} - E) \right] \quad (2.18)$$

E [J] : Énergie électronique.

V_{ox} [V] : Chute de potentiel dans l'oxyde.

m_{ox} [kg] : Masse effective des électrons dans l'oxyde.

t_{ox} [m] : Épaisseur de l'oxyde.

ϕ_b [J] : Hauteur de barrière de l'oxyde.

H : Fonction de Heaviside. Elle nous permet de différencier les configurations *tunnel* direct et Fowler-Nordheim.

Appliquons maintenant l'équation 2.17 au cas d'une barrière multicouche constituée d'un empilement de diélectriques, cf figure 2.7b. On obtient :

$$T_{WKB}(E, V_{ox}) = \exp \left(\frac{-4\sqrt{2}}{3\hbar} \sum_{i=1}^n \frac{\sqrt{m_i} t_i}{V_i} \left[(\phi_{b_i} - R_{i-1})^{\frac{3}{2}} \cdot H(\phi_{b_i} - R_{i-1}) - (\phi_{b_i} - R_i)^{\frac{3}{2}} \cdot H(\phi_{b_i} - R_i) \right] \right) \quad (2.19)$$

$$\text{avec } R_i = q \sum_{j=1}^i V_j + E$$

- m_i [kg] : Masse effective des électrons dans la couche i .
- t_i [m] : Épaisseur de la couche i .
- ϕ_{b_i} [J] : Hauteur de barrière de la couche i .
- V_i [V] : Chute de potentiel dans la couche i .
- V_{ox} [V] : $\sum_{i=1}^n V_i$, chute de potentiel dans l'empilement complet.
Avec l'indexation choisie, il est bon de préciser que $R_0 = E$.

Il est important de rappeler que les champs dans la couche sont reliés entre eux par le théorème de Gauss, cf équation 2.20. Cela permet de calculer la chute de potentiel dans chacune des couches.

$$\varepsilon_i \cdot F_i - \varepsilon_{i+1} \cdot F_{i+1} = \sigma_i \quad (2.20)$$

$$F_{i=1} = - \left[V_{ox} + \sum_{j=2}^n \left(\frac{t_j}{\varepsilon_j} \cdot \sum_{p=1}^{j-1} \sigma_p \right) \right] \cdot \left[t_1 + \varepsilon_1 \cdot \sum_{j=2}^n \frac{t_j}{\varepsilon_j} \right]^{-1}$$

- ε_i [F/m] : Permittivité électrique de la couche i .
- σ_i [C/m²] : Charge à l'interface des couches i et $i + 1$.

2.1.2 Modes de conduction assistée par des défauts

Pour les modes de conduction assistée par des défauts [Geh04], l'injection d'électrons dans le diélectrique n'est pas le mécanisme limitant. Le transport est limité par la conduction des porteurs dans le volume du diélectrique et est lié à la présence de pièges. Ces défauts, pièges, états localisés dans le gap, d'origines diverses (lacunes, impuretés...), assistent la conduction en piégeant puis libérant les électrons qui traversent le diélectrique. Les mécanismes abordés ici sont la conduction Poole-Frenkel, la conduction Poole et le TAT (*Trap Assisted Tunneling* ou mécanisme *tunnel* assisté par piège).

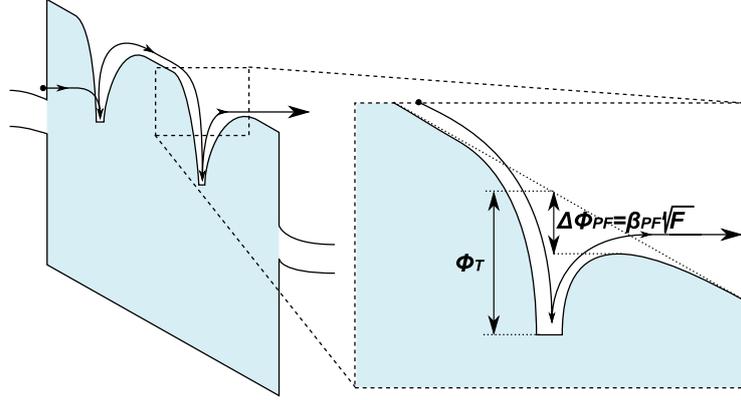


FIG. 2.8: Diagramme de bandes schématique dans le cas d'une conduction Poole-Frenkel. Le puits de potentiel associé au piège génère un abaissement de la barrière de $\Delta\phi_{PF}$.

2.1.2.1 Courant Poole-Frenkel

Le transport Poole-Frenkel correspond à l'émission thermoïonique d'électrons situés dans les pièges d'un diélectrique vers la bande de conduction de celui-ci. Ces électrons libérés, sont ensuite re-capturés par des pièges plus proches de l'anode. Cela correspond à des échanges entre pièges isolés activés en champ et en température. Ce type de mécanisme est prépondérant à haute température, vis-à-vis des mécanismes d'échange *tunnel* de piège à piège, comme le TAT ou le Hopping.

Comme le montre la figure 2.8, l'émission thermoïonique est facilitée lorsque l'on abaisse la barrière de potentiel. En effet, on suppose un piège localement chargé générant un puits coulombien dans la structure, soit :

$$V(x) = \phi_T - \left| \frac{q}{4\pi\epsilon_{diel}x} \right| - F \cdot x \quad (2.21)$$

- F [V/m] : Champ électrique dans le diélectrique.
- ϵ_{diel} [F/m] : Permittivité du diélectrique.
- ϕ_T [J] : Profondeur des pièges.

Cela génère une baisse de la barrière de potentiel de :

$$\Delta\phi_{PF} = \beta_{PF} \cdot \sqrt{F}$$

avec β_{PF} la constante de Poole-Frenkel [Fre38] :

$$\beta_{PF} = \sqrt{\frac{q^3}{\pi \cdot \epsilon_{diel}}} \quad (2.22)$$

Ainsi, [Hil71] exprime le courant Poole-Frenkel selon les deux expressions 2.23 et 2.24, pour un diélectrique amorphe ou cristallin :

– Courant Poole-Frenkel semi-cristallin :

$$J_{PF}^{Cristallin} = A_{Cristallin} \cdot (kT)^2 \cdot F \cdot \exp\left(-\frac{\phi_T - \beta_{PF} \cdot \sqrt{F}}{kT}\right) \quad (2.23)$$

– Courant Poole-Frenkel amorphe :

$$J_{PF}^{Amorphe} = A_{Amorphe} \cdot (kT)^2 \cdot \sqrt{F} \cdot \exp\left(-\frac{\phi_T - \beta_{PF} \cdot \sqrt{F}}{kT}\right) \quad (2.24)$$

avec $A_{Cristallin}$ et $A_{Amorphe}$ sont deux constantes du système¹¹.

Le produit $\exp\left(-\frac{\phi_T - \beta_{PF} \cdot \sqrt{F}}{kT}\right)$ représente la probabilité d'émission d'un site Poole-Frenkel isolé. De plus, la différence entre diélectrique cristallin et amorphe réside dans le choix des vitesses des porteurs effectué par [Hil71]¹².

Cependant, les courants donnés par les équations 2.23 et 2.24 ne considèrent que le flux d'électrons dans le sens direct¹³. La prise en compte des courants inverses modifie globalement ces expressions. [Hil71] montre qu'elles deviennent :

$$J_{PF}^{Cristallin} = 2A_{Cristallin} \frac{(kT)^4}{\beta_{PF}} \cdot \exp\left(-\frac{\phi_T}{kT}\right) \cdot \alpha_{PF}^2 \sinh(\alpha_{PF}) \quad (2.25)$$

$$J_{PF}^{Amorphe} = 2A_{Amorphe} \frac{(kT)^3}{\beta_{PF}} \cdot \exp\left(-\frac{\phi_T}{kT}\right) \cdot \frac{\alpha_{PF} \cosh(\alpha_{PF}) - \sinh(\alpha_{PF})}{\alpha_{PF}} \quad (2.26)$$

$$\text{Avec } \alpha_{PF} = \frac{q\Delta\phi_{PF}}{kT} = \frac{q \cdot \beta_{PF} \cdot \sqrt{F}}{kT}.$$

¹¹Dans [Hil71] $A = q\mu N_i$ avec μ mobilité dans le diélectrique [$m^2V^{-1}s^{-1}$] et $N_i = \frac{\partial N_a}{\partial E} \propto [m^{-3}J^{-2}]$ où N_a est la densité d'états dans le piège

¹²[Hil71] choisie :

– Une dérive en champ classique pour le cristallin : $\frac{\partial v_x}{\partial E} \cdot dE = \mu_{Cristallin} \cdot F$

– Une conduction par procédé de diffusion aléatoire pour l'amorphe : $\frac{\partial v_x}{\partial E} \cdot dE = \mu_{Amorphe} \cdot \sqrt{F}$

Il est à noter que la dépendance en $J \propto F$ ou en $J \propto \sqrt{F}$ en fonction de la cristallinité peut être discutée. Dans tous les cas, elle n'est valable que dans une certaine gamme de V_G (champs forts). En effet, quelque soit le type de matériau, il existe toujours une zone de conduction ohmique à faibles (voir très faibles) champs, soit une dépendance en $J \propto F$. Aussi, on ne retiendra des conductions Poole et Poole-Frenkel que leur dépendance en $J \propto \exp\left(-\frac{\phi}{kT}\right)$

¹³du piège à la bande de conduction du diélectrique

2.1.2.2 Courant Poole

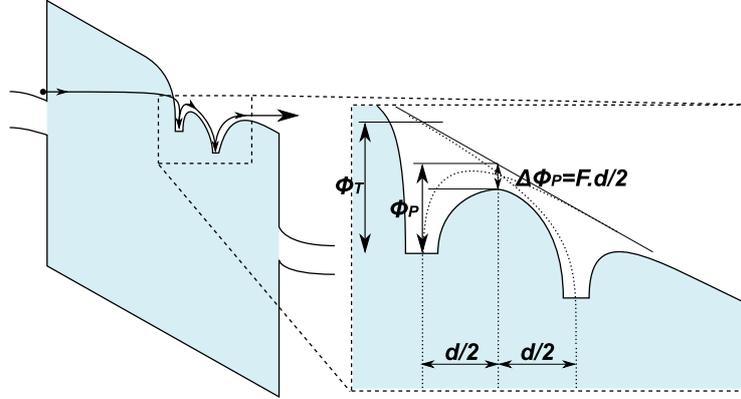


FIG. 2.9: Diagramme de bandes schématisé dans le cas d'une conduction Poole. L'interaction électrostatique entre deux pièges induit un abaissement supplémentaire de ϕ_P .

Lorsque la densité des pièges dans le diélectrique devient très importante, il peut y avoir recouvrement des potentiels coulombiens entre deux pièges, séparés d'une distance d (cf figure 2.9). Aussi, pour un site multiple de pièges, la hauteur de barrière initiale vue par le piège, ϕ_T , est réduite à ϕ_P , même en l'absence de champ électrique, due à l'interaction entre les pièges.

$$\phi_P = \phi_T - \frac{\beta_{PF}^2}{d}$$

L'application d'un champ induit un abaissement de barrière supplémentaire de :

$$\Delta\phi_p = \frac{F \cdot d}{2}$$

Aussi, pour des conditions de faible champ, et pour des émissions thermoïoniques au-dessus d'une barrière $\phi_P - \Delta\phi_p$, [Hil71] montre que le courant peut se mettre sous une forme générique simple :

$$J_{Poole} = A_P \cdot T^2 \cdot \exp\left(-\frac{\phi_P}{kT}\right) \cdot f_P(\alpha_P) \quad (2.27)$$

$$\alpha_P : = \frac{q\Delta\phi_P}{kT} = \frac{q \cdot d \cdot F}{2kT}$$

A_P : Constante du système

$$f_P : \text{Fonction définie par : } f_{Poole}^{Cristallin} = \sinh(\alpha) \text{ ou } f_{Poole}^{Amorphe} = \frac{\alpha \cdot \cosh(\alpha) - \sinh(\alpha)}{\alpha^2}$$

2.1.2.3 Identification expérimentale des courants Poole et Poole-Frenkel

De manière générale les courants Poole et Poole-Frenkel peuvent se mettre sous la forme :

$$J_P = A_P \cdot T^n \cdot \exp\left(-\frac{\phi}{kT}\right) \cdot f(\alpha_P) \quad (2.28)$$

Ainsi, expérimentalement, pour identifier un courant Poole ou Poole-Frenkel à travers un diélectrique, on utilise un diagramme appelé *tracé de Hill*. La variable réduite Γ , donnée par l'équation 2.29 et dépendant du courant expérimental, est tracée en fonction de $T^{-1}\sqrt{F}$ (exemple figure 3.21). Si les tracés de Γ à différentes températures se superposent, le courant est bien un courant Poole ou Poole-Frenkel

$$\Gamma = \frac{J_G}{T^n \cdot \exp\left(-\frac{\phi}{kT}\right)} \quad (2.29)$$

avec :

- $\phi = \phi_P$ et $n=2$ et pour le courant Poole.
- $\phi = \phi_T$ et $n=3$ pour le courant Poole-Frenkel dans un diélectrique amorphe et $n=4$ pour un diélectrique cristallin.

2.1.2.4 Courant TAT : Trap Assisted Tunneling

La conduction TAT, assistée par pièges, est un courant *tunnel*. Il existe deux familles de TAT :

- TAT élastique : dans ce cas, les électrons de la cathode transitent par des pièges situés dans la bande interdite du diélectrique, sans perdre d'énergie avant d'arriver dans l'anode.
- TAT inélastique : pour ce mode de conduction, les électrons de la cathode se relaxent dans les pièges qu'ils rencontrent en émettant l'énergie $\hbar\omega$ (ω : pulsation associée à la fonction d'onde de l'électron [s^{-1}]) sous forme de phonons avant d'arriver dans l'anode.

Ainsi, le piège met un certain temps pour capturer ($\tau_c(x)$ [s]) et émettre ($\tau_e(x)$ [s]) un électron. Ces échanges conduisent à l'établissement de courants entrant, $J_{in}(x)$, et sortant, $J_{out}(x)$, du piège. Une expression simple des courants peut être [Via06] :

$$J_{in}(x) \propto qN_T(x) \cdot T_{in}(x) (1 - f_T(x)) \quad J_{out}(x) \propto qN_T(x) \cdot T_{out}(x) f_T(x)$$

- $N_T(x)$ [m^{-3}] : Densité de pièges en x.
- $f_T(x)$ [-] : Remplissage du piège en x.
- σ [m^2] : Section de capture des pièges.
- T_{in}, T_{out} [-] : Transparences d'entrée et de sortie des électrons du piège.

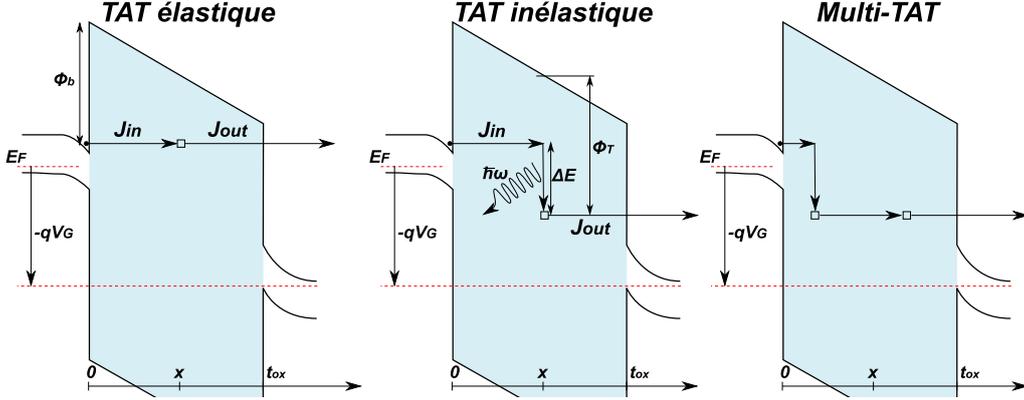


FIG. 2.10: Schémas des conductions *tunnel* assistées par pièges : TAT.

La conduction TAT exprime un régime de conduction établi en dehors de tout effet transitoire [Geh03, Iel00], ainsi :

$$J_{in}(x) = J_{out}(x) = \sigma N_T(x) \frac{T_{in}(x) \cdot T_{out}(x)}{T_{in}(x) + T_{out}(x)}$$

Dans ce cas, l'expression générale du TAT est [Gov06] :

$$J_{TAT} = q \int_0^{t_{ox}} \frac{N_T(x)}{\tau_c(x) + \tau_e(x)} dx \quad \text{ou} \quad \int_0^{t_{ox}} \sigma N_T(x) \frac{T_{in}(x) \cdot T_{out}(x)}{T_{in}(x) + T_{out}(x)} dx \quad (2.30)$$

Plusieurs publications [Her95, Suz86, Geh03] donnent différentes expressions de $\tau_c(x, E)$ et de $\tau_e(x, E)$ en fonction des mécanismes considérés (élastiques, inélastiques, émission multiphonon, ...) et l'on retrouve plusieurs types de répartition de pièges [Cho97, Lar01] (uniforme, Gaussienne, ...) impactant le courant simulé.

Il est intéressant de noter l'approche analytique simple de [Tak99] pour un TAT inélastique, où il exprime :

$$J_{in}(x) = \sigma N_T(x)(1 - f_T(x))J(\phi_B, x, F)$$

$$J_{out}(x) = \sigma N_T(x)f_T(x)J(\phi_T, t_{ox} - x, F)$$

$f_T(x)$ [-] : Remplissage du piège.

$J(\phi_B, x, F)$: Courant *tunnel* du substrat (cf équation 2.8) jusqu'à x , la position du piège.

Hypothèse :

Il obtient ainsi une forme compacte du TAT (cf équation 2.31).

$$J_{TAT} = \frac{t_{eff} \cdot \sigma \cdot N_T(x_T)}{2} J(\phi_B, x_T, F) \quad (2.31)$$

t_{eff} [m] : Épaisseur effective de *tunnel*. Soit $t_{eff} = \frac{\int_0^{t_{ox}} J(\phi_B, x, F) dx}{J(\phi_B, x_T, F)}$

x_T [m] : Position pour laquelle le courant $\frac{J(\phi_B, x, F) \cdot J(\phi_T, t_{ox} - x, F)}{J(\phi_B, x, F) + J(\phi_T, t_{ox} - x, F)}$ est maximal.

F [V/m] : Champ dans l'oxyde.

On notera que, pour les mémoires non-volatiles, les courants TAT sont souvent utilisés pour modéliser les courants de fuite à travers les diélectriques (notamment l'alumine). On soulignera que l'aspect multi-pièges est utilisé pour expliquer les dispersions statistiques des propriétés mémoires, soit par des approches analytiques [Via06], soit par des simulations [Gov06].

2.2 Mécanismes de piégeage dans les diélectriques

Dans cette partie nous allons nous intéresser à la modélisation des pièges dans un diélectrique. Comme nous allons le voir, les pièges peuvent être définis par :

- Le niveau énergétique : ϕ_T .
- La densité volumique (ou surfacique) : N_T .
- Le coefficient de piégeage : $c_{n/p}$.
- La capacité à émettre un trou ou un électron : $e_{n/p}$.
- etc.

2.2.1 Les différentes approches pour la modélisation d'un piège

Un défaut est une anomalie cristalline pouvant modifier la densité d'états du matériau. Pour qu'un défaut induise un piège, il faut qu'il crée un état dans le gap, ce qui n'est pas toujours le cas [Xio07]. Un piège, quant à lui, est un état énergétique dans le gap du diélectrique capable soit de donner, soit d'accepter, des électrons. Un piège est donc électriquement actif. Les échanges de charge avec cet état, ou les charges qu'il peut contenir, peuvent perturber son environnement.

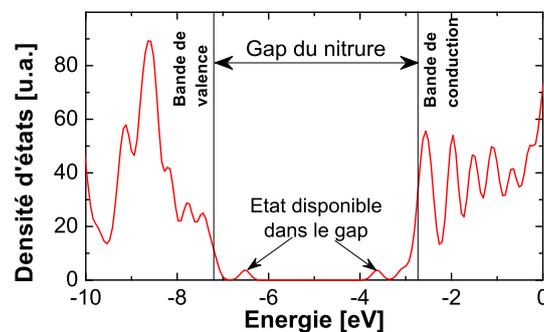


FIG. 2.11: Densité d'états du Si_3N_4 sous phase β calculée en DFT par SIESTA code. Le défaut introduit correspond à une lacune de nitrure neutralisée par un hydrogène (une liaison $Si - H$ et une liaison $Si - Si$) [Via09c].

La figure 2.11 représente la densité d'états dans le nitrure [Via09c]. Elle a été obtenue par simulation atomistique calculée en DTF (pour *Density Functional Theory*). Il s'agit de la Théorie de la Fonctionnelle de la Densité qui permet, à partir de la structure électronique de la matière (atome, structure, liaison, ...), d'obtenir la densité d'états du système. Dans le cas présent, un défaut a été introduit. Il correspond à une lacune de nitrure neutralisée par un atome d'hydrogène (une liaison $Si - H$ et une liaison $Si - Si$). Il est intéressant d'observer que ce défaut crée deux états dans le gap : un proche de la bande de conduction et l'autre proche de la bande de valence. [Via09c] montre que l'état du "bas" peut contenir deux électrons pouvant être émis. De même, l'état du haut peut capturer des électrons. Nous observons ainsi la profondeur du piège (ϕ_T en [J]) et sa densité d'états disponibles (N_T en [m^{-3}]).

En s'appuyant sur cette répartition des états, il est possible de modéliser les pièges suivant différentes approches :

- En considérant chaque état de manière indépendante.
- En considérant un seul piège capable d'accepter les trous ou les électrons : vision amphotérique.
- En considérant un seul piège capable d'accepter des électrons (interdisant de se charger positivement).

2.2.2 Modèle d'états indépendants

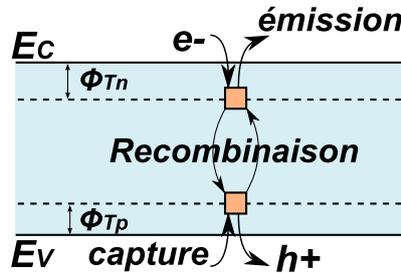


FIG. 2.12: Schéma de bandes des pièges.

L'approche utilisée ici est celle de deux états distincts, l'un ne pouvant piéger que des trous et l'autre ne pouvant piéger que des électrons [Pau06, Son08b, Son08a]. L'un et l'autre peuvent se recombiner. Il est alors possible d'exprimer le remplissage de chaque état [Sho52] :

$$\frac{\partial f_{Tn}}{\partial t} = c_n \cdot (1 - f_{Tn}) - e_n \cdot f_{Tn} - R(f_{Tn}, f_{Tp}) \quad (2.32)$$

$$\frac{\partial f_{Tp}}{\partial t} = c_p \cdot (1 - f_{Tp}) - e_p \cdot f_{Tp} - R(f_{Tn}, f_{Tp}) \quad (2.33)$$

[Son08a] met en évidence, par les faibles valeurs des sections de capture utilisées, que la recombinaison est presque immédiate entre les deux états. De plus, si nous considérons la même densité de pièges pour les électrons et les trous ($N_T = N_{Tn} = N_{Tp}$), la capture

des trous est compensée par la capture des électrons. Ainsi, il serait possible d'écrire le système sous la forme d'une équation :

$$\mathbf{1} \geq \mathbf{f}_T \geq \mathbf{0} : \frac{\partial f_T}{\partial t} = c_n \cdot (1 - f_T) - e_n \cdot f_T - c_p$$

$$-1 \leq \mathbf{f}_T \leq \mathbf{0} : \frac{\partial f_T}{\partial t} = c_n - c_p \cdot (1 + f_T) + e_p \cdot f$$

Utilisant f_T , la densité des charges piégées, où :

$\mathbf{f}_T \geq \mathbf{0}$, $\mathbf{f}_T = \mathbf{f}_{Tn}$ lorsque le piège proche de la bande de conduction se remplit d'électrons.

$\mathbf{f}_T \leq \mathbf{0}$, $\mathbf{f}_T = -\mathbf{f}_{Tp}$ lorsque le piège proche de la bande de valence se remplit de trous.

La densité volumique des charges piégées serait donc $q \cdot f_T \cdot N_T$.¹⁴

2.2.3 Aspect amphotérique des pièges

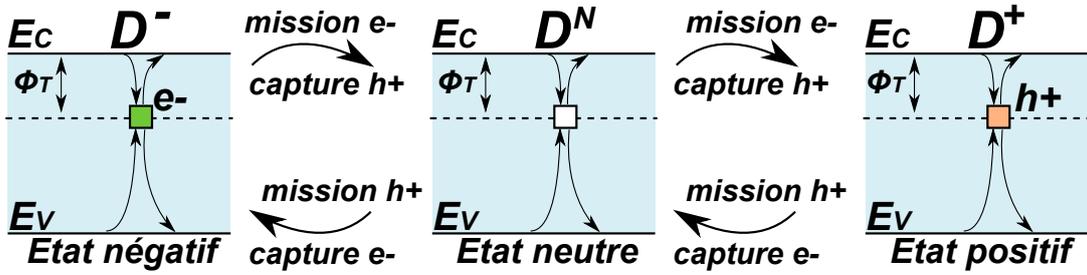


FIG. 2.13: Schémas des différents passages entre les trois états des pièges amphotériques.

La vision amphotérique des pièges est une approche très répandue [Lib90, Yan00, Arr07]. Elle suppose qu'un piège peut prendre trois états différents (cf figure 2.13) :

- L'état D^N : le piège est neutre, non-chargé.
- L'état D^+ : le piège est chargé positivement. Cet état traduit la capacité du défaut à piéger des trous.
- L'état D^- : le piège est chargé négativement. Cet état traduit la capacité du défaut à piéger des électrons.

On associe à chaque état une fonction de remplissage : f_T^N , f_T^+ et f_T^- .

$$\frac{\partial f_T^N}{\partial t} = c_n^+ \cdot f_T^+ - c_n^N \cdot f_T^N - c_p^N \cdot f_T^N + c_p^- \cdot f_T^- \quad (2.34)$$

$$\frac{\partial f_T^+}{\partial t} = c_p^N \cdot f_T^N - c_n^+ \cdot f_T^+ \quad (2.35)$$

$$\frac{\partial f_T^-}{\partial t} = c_n^N \cdot f_T^N - c_p^- \cdot f_T^- \quad (2.36)$$

¹⁴À noter que les coefficients d'émission des électrons et des trous doivent se référer au bon piège, soit : $e_n = e_n^0 \exp\left(\frac{\phi_{Tn}}{kT}\right)$ et $e_p = e_p^0 \exp\left(\frac{\phi_{Tp}}{kT}\right)$ avec e_n^0 et e_p^0 les coefficients d'émission intrinsèque, cf partie 2.2.4.

Les expressions des équations 2.34, 2.35 et 2.36 [Lib90] expriment les remplissages des états seulement en fonction des coefficients de piégeage, $c_{n/p}^{-/N/+}$ [s^{-1}]. Pour tenir compte de l'émission, ces équations deviennent [Yan00, Arr07] :

$$\frac{\partial f_T^+}{\partial t} = c_p^N \cdot f_T^N - e_p^+ \cdot f_T^+ - c_n^+ \cdot f_T^+ + e_n^N \cdot f_T^N \quad (2.37)$$

$$\frac{\partial f_T^-}{\partial t} = c_n^N \cdot f_T^N - e_n^- \cdot f_T^- - c_p^- \cdot f_T^- + e_p^N \cdot f_T^N \quad (2.38)$$

$$f_T^- + f_T^N + f_T^+ = 1 \quad (2.39)$$

Un piège ne peut contenir qu'un nombre limité (cf équation 2.39) d'électrons (pour $f_T^- = 1$) ou de trous (pour $f_T^+ = 1$). Lorsqu'il y a ni trou ni électron, il est entièrement vide (neutre) (pour $f_T^N = 1$).

A partir des équations d'un piège amphotérique, il est possible de retrouver l'expression d'un piège capable de piéger seulement des électrons, cf l'équation 2.40, très couramment utilisée pour les mécanismes de génération/recombinaison SRH dans les semi-conducteurs ou pour des approches mémoires qui ne nécessitent pas le piégeage de trous.

$$\frac{\partial f_T}{\partial t} = c_n \cdot (1 - f_T) - e_n \cdot f_T - c_p \cdot f_T + e_p \cdot (1 - f_T) \quad (2.40)$$

2.2.4 Modélisation des coefficients de piégeage

Dans cette partie nous allons essayer d'explicitier les coefficients de capture, c_n , et d'émission, e_n . Nous pouvons distinguer deux principaux types d'échange :

- Entre la bande de conduction du silicium et le piège, cf figure 2.14a
- Entre la bande de conduction et le piège du diélectrique, cf figure 2.14b

Pour des raisons de simplicité et de symétrie, nous ne traiterons que le cas des électrons, les trous pouvant être obtenus de manière similaire.

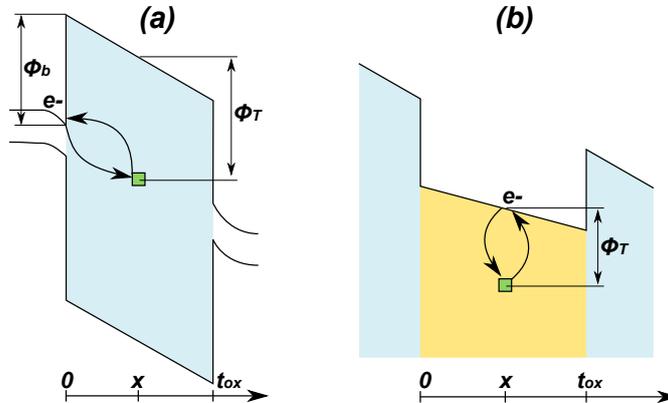


FIG. 2.14: Schémas des deux modes d'échange : a) entre la bande de conduction du silicium et le piège et b) entre la bande de conduction et le piège du diélectrique.

2.2.4.1 Échanges entre la bande de conduction et le piège du diélectrique

Mécanisme de génération-recombinaison de type SRH :

Ce type d'échange correspond aux mécanismes de génération-recombinaison assistés par des centres de recombinaisons : les pièges [Sho52, Mat04]. Ces échanges très utilisés dans les semi-conducteurs sont souvent appelés : mécanismes SRH (Shockley-Read-Hall). Dans cette approche le coefficient de capture des électrons est défini par l'équation 2.41, comme la moyenne de la $v_n \cdot \sigma_n^{15}$ pour les états d'énergie E . Le coefficient d'émission est défini d'après l'équation 2.42.

$$c_n = 2 \cdot \int \sigma_n \cdot \frac{f_n(k)}{(2\pi)^3} \cdot \frac{1}{\hbar} \cdot \sqrt{\left(\frac{\partial E_n^x}{\partial k_n^x}\right)^2 + \left(\frac{\partial E_n^y}{\partial k_n^y}\right)^2 + \left(\frac{\partial E_n^z}{\partial k_n^z}\right)^2} \cdot d^3k \quad (2.41)$$

$$e_n = c_n \cdot \exp\left(\frac{E_T - E_F}{kT}\right) \quad (2.42)$$

- $\sigma_n [m^2]$: Section de capture du piège.
- $\hbar [J \cdot s]$: Constante de Planck réduite.
- $E_n [J]$: $(E_n^x E_n^y E_n^z)$ énergie électronique des électrons selon les trois directions de l'espace.
- $k_n [2\pi/m]$: $(k_n^x k_n^y k_n^z)$ vecteur d'onde associé aux électrons d'énergie $(E_n^x E_n^y E_n^z)$.
- $f_n [-]$: Fonction de distribution des électrons.
- $E_T [J]$: Énergie du piège.
- $E_F [J]$: Énergie du niveau de Fermi du réservoir de porteurs qui échange avec le piège.

En supposant un régime de bandes plates, où les électrons n'ont pas de direction privilégiée de déplacement, la vitesse moyenne des électrons peut être supposée identique dans toutes les directions de l'espace. La vitesse moyenne peut être considérée comme la vitesse thermique des électrons :

$$\frac{1}{\hbar} \cdot \sqrt{\left(\frac{\partial E_n^x}{\partial k_n^x}\right)^2 + \left(\frac{\partial E_n^y}{\partial k_n^y}\right)^2 + \left(\frac{\partial E_n^z}{\partial k_n^z}\right)^2} = \sqrt{\frac{3 \cdot kT}{m_{Si}}} = v_n^{th} \quad (2.43)$$

Le coefficient de capture peut donc s'écrire :¹⁶

$$c_n = \int \sigma_n \cdot v_n^{th} \cdot f_n(k) \cdot \frac{2}{(2\pi)^3} \cdot d^3k = \int \sigma_n \cdot v_n^{th} \cdot f_n(E) \cdot N_C(E) \cdot dE \quad (2.44)$$

En utilisant l'approximation de Boltzmann¹⁷, l'expression devient :

$$c_n = \sigma \cdot v_{th} \cdot N_C \cdot \exp\left(\frac{E_F - E_C}{kT}\right) \quad (2.45)$$

$$N_C(E) [m^{-3}] = 2 \cdot \left(\frac{2\pi \cdot m \cdot kT}{h^2}\right)^{3/2}, \text{ le nombre de places disponibles sur la bande de conduction.}$$

¹⁵ v_n : vitesse des électrons et σ_n la section de capture du piège

¹⁶ $N_C(E)dE = \frac{2}{(2\pi)^3} \cdot d^3k$ avec $N_C(E)dE$ le nombre de places disponibles entre l'énergie E et $E + dE$.

¹⁷ $f_n(E) = \exp\left(\frac{E_f - E}{kT}\right)$

Application aux pièges dans un diélectrique :

Il s'agit d'une reprise de la génération-recombinaison présentée au-dessus. Ainsi, nous utilisons le même formalisme utilisé pour le silicium mais appliquée pour une couche de piégeage, usuellement du nitrure [Sho52, Nas04, Arn75, Gri03] :¹⁸

$$c_n = n \cdot v_n \cdot \sigma_n \quad (2.46)$$

- n [m^{-3}] : Densité d'électrons sur la bande conduction du diélectrique.
 v_n [m/s] : Vitesse des électrons. Dans la littérature, on retrouve, soit la vitesse des porteurs suivant un modèle de Dérive-Diffusion [Lib90], soit une vitesse thermique : $v_{th} = \sqrt{\frac{3kT}{m_{oxn}}}$ [Gu 07].
 σ_n [m^2] : Section de capture du piège.

Comme le résume [Nas04], il existe plusieurs approches pour expliciter l'émission. Mais toutes ont pour base commune les mécanismes de génération-recombinaison et la statistique SRH. Ainsi, on retrouve le coefficient d'émission exprimé plus haut dans beaucoup de publications [McW90, Arr07, Wan05a, Yan00] correspondant à une émission assistée en température (cf équation 2.47).

$$e_n = \sigma_n \cdot v_{th} \cdot N_C \exp\left(\frac{\phi_T}{kT}\right) = e_n^0 \exp\left(\frac{\phi_T}{kT}\right) \quad (2.47)$$

Certains auteurs [Arn75, Lui97, Nas04, Fur06, Gu 07] rajoutent une activation en champ de type Poole-Frenkel explicitée dans l'équation 2.48.

$$e_n = e_n^0 \cdot \exp\left(\frac{\phi_T - \beta_{PF} \cdot \sqrt{F}}{kT}\right) \quad (2.48)$$

- $e_n^0 = \sigma_n \cdot \sqrt{\frac{3 \cdot kT}{m_N}} \cdot 2 \cdot \left(\frac{2\pi \cdot m_N \cdot kT}{h^2}\right)^{3/2}$
 ϕ_T [J] : Profondeur du piège.
 m_N [kg] : Masse effective dans le nitrure.
 T [$^{\circ}K$] : Température.
 β_{PF} [$J\sqrt{m/V}$] : Constante de Poole-Frenkel (cf équation 2.22).

Néanmoins, les approches Poole-Frenkel ne prennent pas en compte l'ionisation à basse température où l'émission est fortement contrôlée par des mécanismes *tunnel*. Ainsi [Lui97] propose un modèle d'émission thermique assistée par *tunnel*, cf équation 2.49. Cette approche considère un continuum d'états entre le piège et la bande de conduction. Ce modèle permet de considérer l'émission Poole-Frenkel à haute température et l'émission assistée par *tunnel* à basse température.

Il est intéressant de constater que ces différentes approches, bien que différentes, comportent le terme SRH. Les autres membres traduisent les nouveaux phénomènes pris en compte.

¹⁸On peut noter que certains auteurs utilisent $c_n^0 = v_n \cdot \sigma_n$ comme paramètre du système.

$$e_n = \frac{e_n^0}{kT} \int_0^{w-\beta\sqrt{F}} \exp\left(-\frac{E}{kT} - \frac{2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m_N(qV(x) - E)} dx\right) dE + e_n^0 \cdot \exp\left(\frac{\phi_T - \beta_{PF} \cdot \sqrt{F}}{kT}\right) \quad (2.49)$$

$$x_{1,2} = \frac{1}{2} \frac{\phi_T - E}{qF} \left[1 \mp \left(\frac{qF}{\pi\epsilon\epsilon_\infty \cdot (\phi_T - E)^2} \right)^{1/2} \right]$$

$V(x)$ [V] : Potentiel d'un puits coulombien, cf équation 2.21.

[Nas04] propose aussi un modèle d'émission par ionisation des pièges par multi-phonons. Cette approche est donnée par l'expression :

$$e_n = \frac{qF}{2\sqrt{2m^*W_{opt}}} \exp\left(-\frac{4}{3} \frac{\sqrt{2m^*}}{\hbar qF} W_{opt}^{3/2} + 4 \frac{m^*W_{ph}}{\hbar^2} \cdot \frac{W_{opt}(W_{opt} - W_t)}{q^2F^2} \cosh \frac{W_{ph}}{2T}\right) \quad (2.50)$$

W_{opt}, W_t : Énergies optique et thermique de ionisation des pièges.

W_{ph} : Énergie des phonons.



Pour conclure, dans l'ensemble de nos simulations, nous avons choisi une émission activée en température et en champ de type Poole-Frenkel décrite dans l'équation 2.48

2.2.4.2 Échanges entre la bande de conduction du silicium et le piège

Dans cette partie, nous partons des équations décrivant les échanges entre les pièges neutres d'un diélectrique et des électrodes en Si [Iel00, Gar07]. Ces équations se basent sur la statistique de recombinaison des électrons et des trous SRH présentée dans [Sho52]. Dans notre cas les échanges ne se font pas entre des pièges situés dans le gap du silicium mais dans un diélectrique. Ainsi, il s'agit du même type d'échange déporté à travers une barrière de potentiel, d'où l'utilisation de la transparence vue entre le réservoir de porteurs (ici le silicium) et les pièges.

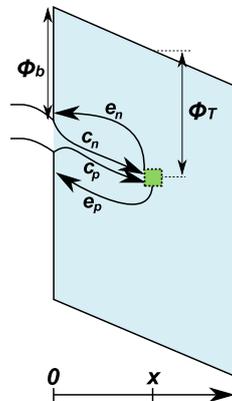


FIG. 2.15: Schéma des échanges entre la bande de conduction du silicium et les pièges.

Ainsi, le coefficient de capture s'écrit d'après l'équation 2.51 et le coefficient d'émission d'après l'équation 2.52. En développant l'équation 2.51, nous obtenons l'équation 2.53

$$c_n = 2 \int \frac{\sigma_n \cdot f_n(k_n) \cdot T_n(k_n)}{\hbar(2\pi)^3} \sqrt{\left(\frac{\partial E_n^x}{\partial k_n^x}\right)^2 + \left(\frac{\partial E_n^y}{\partial k_n^y}\right)^2 + \left(\frac{\partial E_n^z}{\partial k_n^z}\right)^2} \cdot d^3 k_n \quad (2.51)$$

$$e_n = c_n \cdot \exp\left(\frac{E_T - E_F}{kT}\right) \quad (2.52)$$

$$c_n = \frac{4\pi m_{Si} kT}{h} \sigma_n \int_{E_{inf}}^{\infty} \ln\left(1 + \exp\left(\frac{E_F - E}{kT}\right)\right) T(E) dE \quad (2.53)$$

- σ_n [m^2] : Section de capture du piège.
 \hbar [$J \cdot s$] : Constante de Planck réduite.
 E_n [J] : $(E_n^x E_n^y E_n^z)$ énergie électronique des électrons selon les trois directions de l'espace.
 k_n [$2\pi/m$] : $(k_n^x k_n^y k_n^z)$ vecteur d'onde associé aux électrons d'énergie $(E_n^x E_n^y E_n^z)$.
 f_n [-] : Fonction de distribution des électrons.
 T_n [-] : Transparence associée à toute barrière de potentiel que les électrons doivent franchir pour transiter entre le piège et le réservoir de porteurs.
 E_T [J] : Énergie du piège.
 E_F [J] : Énergie du niveau de Fermi du réservoir de porteurs qui échange avec le piège.
 E_{inf} [J] : Maximum de l'énergie des pièges et des électrons du réservoir de porteurs.



Il est intéressant de noter qu'il existe une analogie très forte entre ce mécanisme et les courants de type *tunnel*. En effet, les expressions des coefficients de capture peuvent aussi s'écrire $c_n = \frac{\sigma_n}{q} \cdot J_{Tun}$, où J_{Tun} est le courant *tunnel* entre le réservoir d'électrons (ici le silicium) et le piège.

2.3 Modèles de piégeage pour les mémoires de type SONOS

Dans cette partie nous allons présenter des modèles d'écriture, d'effacement et de rétention des mémoires SONOS¹⁹ dans le cas d'une programmation Fowler-Nordheim²⁰. Pour ce faire nous allons nous appuyer sur l'ensemble des concepts présentés précédemment.

Il existe un grand nombre de publications traitant de ce sujet [Arr08a, Gri03, McW90, Pau06, Son08b, Wan05a, Yan00]. Aussi, dans un premier temps, nous présenterons très succinctement quelques modèles de mémoires SONOS de la littérature. Puis nous détaillerons deux modèles de piégeage sur lesquels nous nous sommes appuyés dans les chapitres suivants. Notre objectif est d'avoir un modèle physique permettant de comprendre et d'analyser le comportement expérimental de nos dispositifs. Les détails des équations et des mécanismes considérés pour chacun d'entre eux sont donnés dans les parties 2.3.2 et 2.3.3.

2.3.1 État de l'art de quelques modèles SONOS

[Buc06, Gar07] présentent un modèle de piégeage et de déchargement. Ils ne considèrent que les échanges directs entre le canal, la grille et les pièges. Ils se basent sur les équations 2.53 et 2.52. On notera dans cette approche, l'utilisation d'états liés, une résolution semi-analytique et la prise en compte des effets quantiques dans le canal.

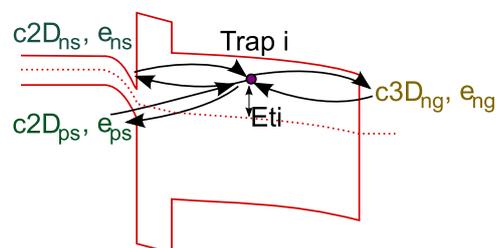


Fig. 2.a : Échanges pris en compte par [Gar07].

[Fur08] présente un modèle de rétention. Il considère l'émission de pièges vers la bande de conduction du nitrure. Il s'agit d'une émission assistée par phonons. Les électrons émis sur la bande de conduction du nitrure sont séparés en deux groupes, ceux qui sont émis par la grille et ceux qui sont émis côté canal et qui sont repiégés. De plus, il considère une transparence égale à un pour l'alumine. On notera que cette approche ne prend pas en compte l'émission d'électrons à travers le SiO_2 , mais considère l'injection de trous du canal vers le nitrure.

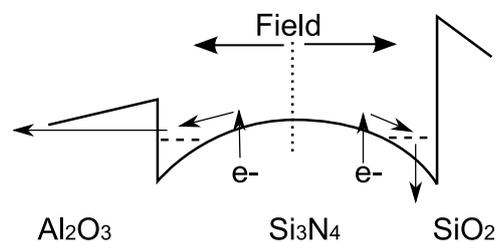


Fig. 2.b : Diagramme de bandes et séparation de mécanismes d'émission considérés dans [Fur08].

¹⁹Mémoire à couche de piégeage à base de nitrure.

²⁰En opposition à l'injection par porteurs chauds des mémoires NOR.

[Fur07] présente un modèle d'écriture. Ses principales hypothèses sont :

- La triangulation de l'énergie des électrons dans le canal,
- L'approximation WKB pour les transparences,
- Une répartition gaussienne de la profondeur des pièges (de moyenne $1,8eV$),
- Une émission de type Poole-Frenkel.

De plus, il considère un libre parcours moyen des électrons dans le nitrure. Ainsi, toute la charge injectée dans le nitrure n'est pas piégée et une partie continue vers la grille de contrôle. Le libre parcours moyen permettrait d'expliquer la cinétique de piégeage différente des mémoires SANOS par rapport aux mémoires à grille flottante continue. L'émission Poole-Frenkel permettrait, quant à elle, d'ajuster la saturation des caractéristiques de programmation.

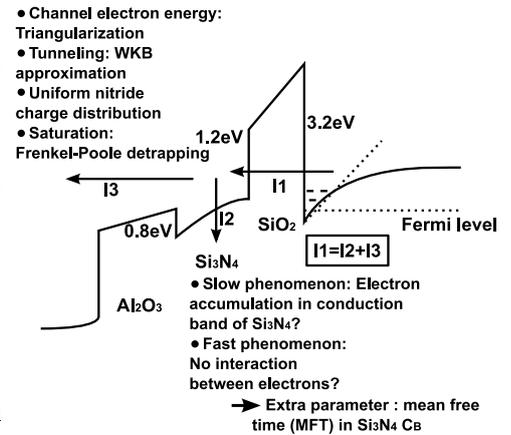


Fig. 2.c : Diagramme de bandes et mécanismes pris en compte dans [Fur07].

[Gri03] présente un modèle d'écriture/effacement qui inclut le transport des électrons et des trous sur les bandes de conduction et de valence du nitrure. Il considère un mécanisme de dérive/diffusion avec une vitesse fixe. L'émission entre les pièges et la bande de conduction du nitrure est une ionisation des pièges assistée par multi-phonons (cf l'équation 2.50). Les conditions aux limites du courant de dérive/diffusion sont prises aux interfaces SiO_2/Si_3N_4 par des courants Fowler-Nordheim.

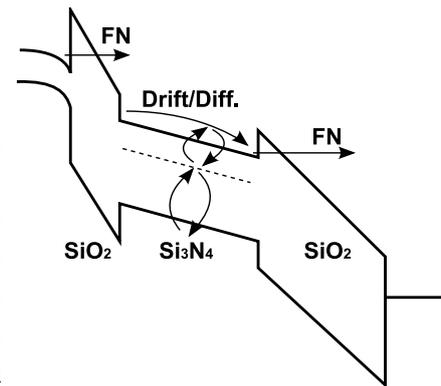


Fig. 2.d : Diagramme de bandes et mécanismes pris en compte dans [Gri03].

[Wan05a] présente un modèle analytique de rétention des SONOS. Il considère les mécanismes d'émission des électrons de type thermoïomique (cf équation 2.47) ainsi que les échanges du piège vers le canal. Mais pour permettre la résolution analytique du système (cf équation 2.40), il ne prend pas en compte la modification des coefficients d'émission due à la perte de charge (modifiant les champs et les transparences dans l'empilement).

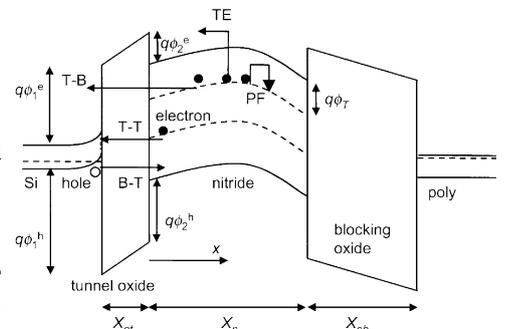


Fig. 2.e : Diagramme de bandes et émissions considérées dans [Wan05a].

[Arr08a] présente un modèle de rétention qui considère les mêmes mécanismes que [Wan05a]. La grande différence réside dans la résolution numérique du système, la répartition en énergie des pièges et à la prise en compte de la modification des coefficients d'émission due à la perte de charge.

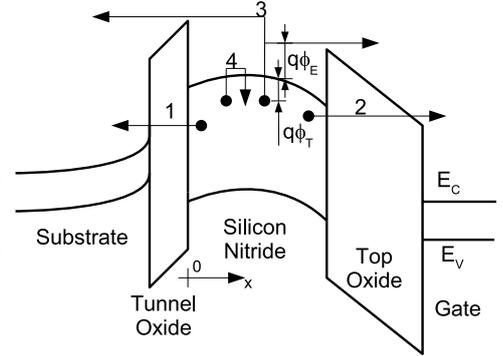


Fig. 2.f : Diagramme de bandes et émissions considérées dans [Arr08a].

[Lee08] présente des simulations numériques de programmation. Ce modèle prend en compte l'injection d'électrons dans le nitrure (J_{T-in}). Une partie de ces électrons peut traverser l'oxyde bloquant (J_{T-out}). Ceux qui restent dans le nitrure peuvent être capturés (J_C). Les électrons piégés peuvent être émis sur la bande de conduction du nitrure par émission Poole-Frenkel puis évacués par l'oxyde bloquant (J_{e-out}). Les champs électriques dans l'empilement sont modifiés par la charge présente dans le nitrure. Il suppose que la charge est uniformément piégée à une seule profondeur énergétique et que la localisation spatiale est définie par un centroïde de charge.

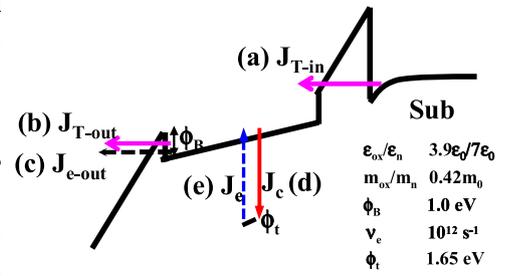


Fig. 2.g : Diagramme de bandes et courants considérés dans [Lee08].

[Son08b] présente des simulations basées sur un programme prenant en compte l'écriture, l'effacement et la rétention. Il s'agit d'un modèle *self-consistent* (c'est-à-dire qu'il prend en compte l'ensemble des modifications du système à chaque instant de la simulation) considérant l'évolution spatiale et temporelle du potentiel, des charges piégées, des charges sur les bandes de conduction/valence et des courants d'injection et d'émission. Ses principales hypothèses sont :

- Un mécanisme TAT pour les courants d'injection,
- Une activation Poole-Frenkel pour l'émission des électrons piégés,
- Deux niveaux de piège (un pour les électrons et un pour les trous) pouvant recombiner entre eux.

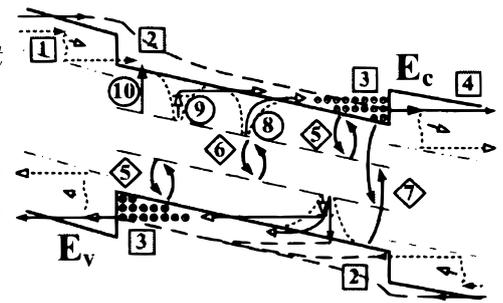


Fig. 2.h : Diagramme de bandes et ensemble de mécanismes pris en compte dans [Lee08].

2.3.2 Modèle 1 : prise en compte de pièges amphotériques

Dans cette approche nous avons voulu prendre en compte les mécanismes liés aux électrons et aux trous. Ainsi, un modèle amphotérique du piège a été adopté. Néanmoins, pour simplifier la résolution du système d'équations différentielles, nous avons choisi de ne pas considérer le transport des porteurs (e^- et h^+) dans la couche de piégeage. Nous avons donc supposé une densité uniforme de charge disponible sur les bandes de conduction et de valence. Ainsi, nous ne considérons qu'une densité surfacique de charge positive, Q_p , et négative, Q_n ²¹.

On soulignera que les différents types d'injection, *tunnel* direct, Fowler-Nordheim et Fowler-Nordheim modifié, ont été pris en compte. De plus, les pièges peuvent échanger avec les bandes de conduction/valence de la couche de piégeage mais aussi avec le substrat, d'après les équations 2.46, 2.48, 2.52 et 2.53. De plus, les effets quantiques dans le substrat sont pris en compte par l'intermédiaire :

- des niveaux d'énergie utilisés pour l'injection.
- de la tension de substrat.
- de la charge dans le canal.

D'un point de vue pratique, nous utilisons des tables obtenues par des simulations Poisson-Schrödinger, nous donnant la charge et la tension dans le canal, et le niveau fondamental en fonction du dopage et de la tension de grille.

Ainsi, l'ensemble du système peut être modélisé par le système des équations ci-dessous :

- Charge négative disponible sur la bande de conduction du diélectrique :

$$\frac{\partial Q_n}{\partial t} = J_{in_n} - J_{out_n} - (c_n^N \cdot f_T^N - e_n^- \cdot f_T^- + c_n^+ \cdot f_T^+ - e_n^N \cdot f_T^N) \cdot N_T \cdot q \cdot t_{diel} \quad (2.54)$$

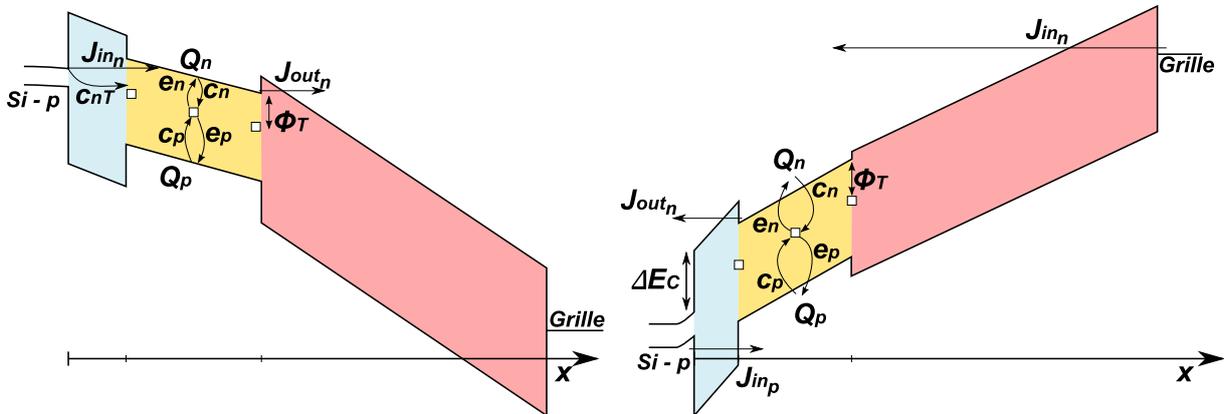


FIG. 2.16: Schémas des mécanismes considérés dans le modèle 1.

²¹ $n \cdot t_{diel} = Q_n$ $p \cdot t_{diel} = Q_p$ avec n et p les densités volumiques de charges sur les bandes de conduction et de valence du nitrure. t_{diel} est l'épaisseur du nitrure.

- Charge positive disponible sur la bande de valence du diélectrique :

$$\frac{\partial Q_p}{\partial t} = J_{in_p} - J_{out_p} - (c_p^N \cdot f_T^N - e_p^+ \cdot f_T^+ + c_p^- \cdot f_T^- - e_p^N \cdot f_T^N) \cdot N_T \cdot q \cdot t_{diel} \quad (2.55)$$

La variation de la charge correspond à la différence des courants entrant et sortant dans le diélectrique, moins la charge piégée.

- Remplissage des pièges amphotériques avec prise en compte des échanges entre pièges \longleftrightarrow substrat et pièges \longleftrightarrow diélectrique :

$$\frac{\partial f_T^+}{\partial t} = (c_p^N + c_{pT}^N) \cdot f_T^N - (e_p^+ + e_{pT}^+) \cdot f_T^+ - (c_n^+ + c_{nT}^+) \cdot f_T^+ + (e_n^N + e_{nT}^N) \cdot f_T^N \quad (2.56)$$

$$\frac{\partial f_T^-}{\partial t} = (c_n^N + c_{nT}^N) \cdot f_T^N - (e_n^- + e_{nT}^-) \cdot f_T^- - (c_p^- + c_{pT}^-) \cdot f_T^- + (e_p^N + e_{pT}^N) \cdot f_T^N \quad (2.57)$$

$$f_T^- + f_T^N + f_T^+ = 1 \quad (2.58)$$

- Répartition du potentiel dans la structure :

$$\varepsilon_{diel} \frac{\partial^2 V}{\partial x^2} = q \cdot (n - p + f_T^- \cdot N_T - f_T^+ \cdot N_T) \quad (2.59)$$

- Coefficients de capture entre les pièges et les bandes de conduction ou de valence du diélectrique de piégeage [s^{-1}] :

$$c_{n/p}^{+/N/-} = n \cdot v_{n/p}^{+/N/-} \cdot \sigma_{n/p}^{+/N/-}$$

- Coefficient d'émission entre les pièges et les bandes de conduction ou de valence du diélectrique de piégeage en fonction des différents états du piège [s^{-1}] :

$$e_{n/p}^{+/N/-} = v_{n/p}^{+/N/-} \cdot \sigma_{n/p}^{+/N/-} \cdot N_C \cdot \exp\left(\frac{\phi_{T_{n/p}} - \beta_{PF} \cdot \sqrt{F}}{kT}\right)$$

- Coefficient de capture entre les pièges et le substrat en fonction des différents [s^{-1}] :

$$c_{n/pT}^{+/N/-} = \frac{4\pi m_{Si} kT}{h} \sigma_{n/pT}^{+/N/-} \int_{E_{inf}}^{\infty} \ln\left(1 + \exp\left(\frac{E_F - E}{kT}\right)\right) T(E) dE$$

- Coefficient d'émission entre les pièges et le substrat en fonction des différents états du piège [s^{-1}] :

$$e_{n/pT}^{+/N/-} = c_{n/pT}^{+/N/-} \cdot \exp\left(\frac{E_T - E_F}{kT}\right)$$

- Nombre de places disponibles sur la bande de conduction ou de valence [m^{-3}] :

$$N_{C_{n/p}} = 2 \left(\frac{2\pi m_{ox_{n/p}} kT}{h^2} \right)^{3/2}$$

V [V]	: Potentiel dans l'empilement.
$f_T^{+/N/-}$ [-]	: Remplissage des pièges.
$Q_{n/p}$ [C/m ²]	: Charge présente sur la bande de conduction/valence de la couche de piégeage.
N_T [m ⁻³]	: Densité volumique de pièges dans le diélectrique de piégeage.
ε_{diel} [F/m]	: Permittivité du diélectrique.
$\phi_{T_{n/p}}$ [J]	: Profondeur des pièges par rapport à la bande de conduction ou de valence.
E_T [J]	: Énergie des pièges.

Ainsi les principales hypothèses sont :

- Des pièges amphotériques dans la couche de piégeage.
- L'absence de transport et une répartition uniforme de charge sur les bandes de conduction et de valence de la couche de piégeage.
- Une statistique SRH pour le calcul des coefficients d'émission.
- Une activation en température de type Poole-Frenkel pour les émissions vers les bandes de conduction et de valence de la couche de piégeage.
- L'absence d'échanges directs de piège à piège.
- Des conductions de type *tunnel* (direct, Fowler-Nordheim et Fowler-Nordheim modifié) dans l'oxyde *tunnel* et l'oxyde bloquant (pas de mécanismes assistés par pièges dans l'oxyde bloquant).
- Une statistique de Fermi dans le canal.
- Une statistique de Boltzmann dans la couche de piégeage.
- L'absence de courant de trous entre la couche de piégeage et la grille de contrôle.

L'hypothèse la plus lourde reste l'absence de transport dans la couche de piégeage (sur les bandes de conduction et de valence ainsi qu'entre les pièges). L'amélioration la plus pertinente de ce modèle serait donc la prise en compte de ce phénomène. Dans un deuxième temps, les mécanismes de conduction assistée par pièges pourraient être implémentés pour être en accord avec les mesures des courants de fuite à travers les matériaux *High- κ* actuellement envisagés comme oxyde bloquant (comme l'alumine par exemple).

2.3.3 Modèle 2 : prise en compte du transport dans la couche de piégeage

Ce modèle a été développé par l'université d'Udine et est détaillé dans [Via09b]. Nous avons utilisé cette approche pour plusieurs analyses de résultats d'écriture (cf partie 4.2.3.1) et de rétention (cf partie 4.2.3.3).

La particularité de cette modélisation est de prendre en compte le transport des électrons dans la couche de piégeage. Nous considérons que les électrons injectés au-dessus de la bande de conduction du nitrure sont thermalisés sur la bande de conduction de la couche de piégeage. Le transport de ces électrons est modélisé par un mécanisme de dérive-diffusion (ce courant est défini par l'équation 2.60) et de piégeage (le taux net de piégeage est défini par l'équation 2.61). Les coefficients de capture et d'émission sont définis par les équations 2.46 et 2.48. Les conditions aux limites pour ce courant sont imposées par les

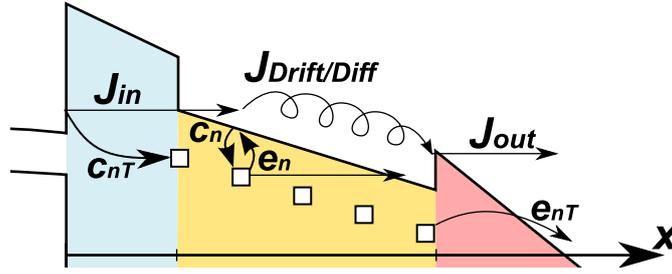


FIG. 2.17: Schéma des mécanismes considérés dans le modèle 2.

courants d'injections *tunnel*, J_{in} et J_{out} , définis par les équations 2.6, 2.8 et 2.11. Comme pour le modèle de la partie 2.3.2, nous avons considéré les différents modes d'injection à travers l'oxyde *tunnel* et l'oxyde bloquant.

$$\vec{i}_n(x, t) = q \left(n(x, t) \cdot \mu_n \cdot \overrightarrow{\text{grad}}(V(x, t)) + \frac{kT}{q} \cdot \mu_n \cdot \overrightarrow{\text{grad}}(n(x, t)) \right) \quad (2.60)$$

$$r_n(x, t) = c_n \cdot (1 - f_T(x, t)) - e_n \cdot f_T(x, t) \quad (2.61)$$

Les échanges directs entre canal et piège ont été considérés. Il s'agit d'un piègeage inélastique modélisé par les termes c_{nT} et e_{nT} de l'équation 2.63. Ainsi, le système peut être synthétisé suivant les équations différentielles 2.62, 2.63 et 2.64.

- Transport dans la bande de conduction de la couche de piégeage :

$$\frac{\partial n(x, t)}{\partial t} = \text{div} \left(\frac{\vec{i}_n(x, t)}{q} \right) - r_n(x, t) \cdot N_T \quad (2.62)$$

- Remplissage des pièges :

$$\frac{\partial f_T(x, t)}{\partial t} = r_n(x, t) + c_{nT} \cdot (1 - f_T(x, t)) - e_{nT} \cdot f_T(x, t) \quad (2.63)$$

- Répartition du potentiel dans la structure :

$$\varepsilon_{\text{diel}} \frac{\partial^2 V(x, t)}{\partial x^2} = q \cdot (n(x, t) + f_T(x, t) \cdot N_T) \quad (2.64)$$

$V(x, t)$ [V]	: Potentiel dans l'empilement.
$n(x, t)$ [m^{-3}]	: Densité volumique d'électrons sur la bande de conduction du diélectrique de piégeage.
$f_T(x, t)$ [-]	: Remplissage des pièges.
N_T [m^{-3}]	: Densité volumique de pièges dans le diélectrique de piégeage.
$\vec{i}_n(x, t)$ [A/m^2]	: Courant de dérive-diffusion des électrons sur la bande de conduction de la couche de piégeage.
$r_n(x, t)$ [s^{-1}]	: Taux net de piégeage entre les pièges et la bande de conduction du diélectrique de piégeage.
c_n [s^{-1}]	: Coefficient de capture entre les pièges et la bande de conduction du diélectrique de piégeage, cf l'équation 2.46.
e_n [s^{-1}]	: Coefficient d'émission entre les pièges et la bande de conduction du diélectrique de piégeage, cf l'équation 2.48.
c_{nT} [s^{-1}]	: Coefficient de capture entre les pièges et le substrat, cf l'équation 2.53.
e_{nT} [s^{-1}]	: Coefficient d'émission entre les pièges et le substrat, cf l'équation 2.52.
ε_{diel} [F/m]	: Permittivité du diélectrique.
μ_n [$m^2 s^{-1} V^{-1}$]	: Mobilité des électrons sur la bande de conduction de la couche de piégeage.

On notera, dans cette approche, que l'on ne prend pas en compte les effets quantiques, ni la discrétisation des états dans le canal. Ainsi pour compenser cet effet, les hauteurs des barrières ont été abaissées $\sim 0,2eV^{22}$. On soulignera aussi la prise en compte de l'émission des électrons vers la bande de conduction du nitrure par effet *tunnel* (cf figure 2.17). Enfin, cette approche nécessite des précautions dans la résolution du système²³ dues à l'équation de dérive diffusion. Ce modèle nous permet une très bonne simulation de l'écriture et de la rétention, mais nous interdit une modélisation correcte de l'effacement du fait de la non prise en compte des trous.

	Modèle 1 cf paragraphe 2.3.2	Modèle 2 cf paragraphe 2.3.3
Dérive-Diffusion	non	oui
Trous	Piège amphotérique	non
Permet la modélisation de :		
L'écriture	mauvaise dépendance en V_G pas de répartition de la charge	oui
L'effacement	oui	non
La rétention	pas de redistribution de charges	oui

TAB. 2.1: Résumé des propriétés des deux modèles.

²²Pour le SiO_2 , le ΔE_C devient $2,9eV$ au lieu des $3,15eV$ habituel.

²³L'approche utilisée ici est du type "Backward Euler Methods"

2.4 Conclusion

L'objectif de ce chapitre était d'introduire l'ensemble des éléments théoriques permettant l'analyse des résultats expérimentaux qui seront exposés dans les deux prochains chapitres.

Aussi, nous avons tout d'abord présenté des rappels sur les modes de conduction électronique dans les diélectriques telle que la conduction *tunnel*, Fowler-Nordheim ou les conductions assistées par pièges (conduction Poole-Frenkel, TAT, ..). Nous avons aussi explicité le courant partant d'un diélectrique vers une anode, courant de sortie que l'on rencontre souvent dans les mémoires de type SONOS.

Nous avons donné plusieurs représentations des pièges dans un diélectrique. Cela a permis de modéliser les différents échanges que les pièges peuvent avoir avec leur environnement : échange piège \Leftrightarrow bande de conduction/valence du diélectrique ou échange piège \Leftrightarrow bande de conduction/valence d'une électrode (substrat ou grille).

Enfin, nous avons présenté deux modèles de piégeage pour les mémoires de type SONOS, à injection Fowler-Nordheim. Ces modèles étant complémentaires, nous avons regroupé leurs avantages et leurs inconvénients dans le tableau 2.1. Ils nous permettront de couvrir un grand spectre de phénomènes physiques. Ainsi, dans les chapitres suivants, nous nous appuierons souvent sur ces modèles pour mieux comprendre les phénomènes de piégeage, de rétention ou d'effacement observés expérimentalement.

Chapitre 3

Etude des matériaux *High- κ*

Ce chapitre consiste en une présentation et une analyse des matériaux *High- κ* . En s'appuyant sur les analyses physico-chimique réalisées au CEA-Léti, nous allons évaluer leurs intérêts par des caractérisations électriques. Nous allons étudier :

- Les alliages à base d'alumine et d'oxyde d'hafnium : les aluminates d'hafnium.
- L'effet de la nitruration sur ces matériaux.
- L'impact des traitements thermiques sur l'alumine.
- Les silicates d'hafnium pour des intégrations en tant qu'oxyde *tunnel*.

Enfin, en se basant sur ces mesures, nous évaluerons les potentiels de ces matériaux une fois intégrés en tant que diélectrique d'interpoly dans une cellule mémoire Flash.

Sommaire

3.1	Présentation des matériaux <i>High-κ</i>	95
3.1.1	Introduction générale des matériaux <i>High-κ</i>	95
3.1.2	Procédés de fabrication	97
3.2	Alliages à base de HfO_2 et Al_2O_3 : Les aluminates d'hafnium	97
3.2.1	Généralités sur les <i>HfAlO</i>	97
3.2.1.1	Procédés de fabrication	97
3.2.1.2	Stœchiométrie, gap et cristallinité	99
3.2.2	Étude des charges fixes dans les aluminates d'hafnium	102
3.2.2.1	Effet de la concentration en <i>Hf</i>	102
3.2.2.2	Localisation des charges fixes dans les aluminates	103
3.2.3	Courant de fuite dans les <i>HfAlO</i> : épaisseurs et natures	105
3.2.3.1	Influence de l'épaisseur des <i>HTO</i> dans un tricouche	106
3.2.3.2	Problématique du piégeage et mesure de courant	108
3.2.3.3	Influence de la couche <i>High-κ</i> à faible et à fort champ	109
3.2.3.4	Modes de conduction et d'activation en température	111
3.2.4	Propriétés de piégeage des <i>HfAlO</i>	115
3.3	Impact de la nitruration sur les <i>HfAlO</i>	118
3.3.1	Impact sur les propriétés physico-chimiques	119

3.3.1.1	Incorporation de l'azote dans le film	119
3.3.1.2	Influence sur le gap optique	120
3.3.1.3	Cristallinité	121
3.3.2	Effets du recuit NH_3 sur les propriétés de piégeage	123
3.4	Impact du traitement thermique sur l'alumine	126
3.4.1	Les propriétés physico-chimiques	127
3.4.1.1	Étude du stress, de la densité et de l'épaisseur	127
3.4.1.2	Mesures FTIR-ATR	128
3.4.1.3	Concentration d'hydrogène dans l'empilement mémoire	130
3.4.1.4	Simulation atomistique	131
3.4.2	Étude des caractéristiques électriques	134
3.4.2.1	Étude du courant de fuite	134
3.4.2.2	Étude des propriétés de piégeage	135
3.5	Silicate d'hafnium : $HfSiON$	137
3.5.1	Analyses matériaux	137
3.5.2	Étude du courant de fuite	140
3.5.2.1	Structure bicouche $SiO_2/HfSiON$	140
3.5.2.2	Structure tricouche $SiO_2/HfSiON/HTO$	142
3.6	Évaluation des performances mémoires	144
3.6.1	Simulation de la fenêtre de programmation	144
3.6.2	Simulation de la rétention	148
3.7	Conclusion	151

3.1 Présentation des matériaux *High- κ*

La rapide croissance de l'industrie microélectronique depuis les années 50 est liée à l'amélioration continue des performances et de la densité d'intégration des composants. La réduction rapide des dimensions caractéristiques des transistors a conduit à des longueurs de dispositifs toujours plus faibles et à des épaisseurs d'isolation de plus en plus fines. De nouveaux problèmes apparaissent, comme l'augmentation du courant de fuite à travers l'oxyde de grille où l'épaisseur est difficilement réductible. Pour traiter ce problème, l'industrie microélectronique étudie différentes approches, dont l'une consiste à remplacer l'isolant de grille en SiO_2 , par un matériau à plus forte constante diélectrique, permettant d'avoir un contrôle électrostatique du canal amélioré et moins de fuites de grille.

Le fort engouement porté sur ces matériaux, pour les applications logiques, les a conduit à être étudiés pour les mémoires non-volatiles. Comme nous l'avons vu dans le chapitre 1, les mémoires à grille flottante fonctionnent à des tensions élevées et la poursuite de la réduction de leurs dimensions est délicate. C'est donc pour cela que de nombreux efforts de recherche se portent actuellement sur les *High- κ* , pour les intégrer aussi bien en tant qu'interpoly que couche de piégeage ou bien encore comme diélectrique *tunnel*.

3.1.1 Introduction générale des matériaux *High- κ*

Le terme *High- κ* peut se traduire littéralement comme "haute permittivité relative"¹. Dans sa définition la plus stricte, un *High- κ* est un oxyde métallique ayant une constante diélectrique plusieurs fois supérieure à celle du SiO_2 . En pratique, il arrive couramment que soient assimilés aux *High- κ* des diélectriques dont la permittivité ne vaut qu'un facteur deux de celle de la silice², le plus célèbre dans les applications mémoires étant l'alumine. Les modélisations physiques employées pour décrire l'isolation électrique du SiO_2 , lui associent un gap et une masse effective de façon à garder un formalisme compatible avec celui employé pour le silicium.³

De nombreux diélectriques alternatifs au SiO_2 sont aujourd'hui à l'étude (voir le tableau 3.1). De façon empirique, on constate que le gap est d'autant plus faible que sa permittivité est élevée (voir la figure 3.1, [Rob02]) et il en est de même pour l'écart entre bande de conduction du Si et du *High- κ* (ΔE_C). Partant de cette constatation et en vue des différentes contraintes des applications logiques, mémoires flash ou DRAMs, il

¹le κ , kappa en grec, représente la permittivité relative chez les anglo-saxons : $\varepsilon = \kappa \cdot \varepsilon_0 = \varepsilon_r \cdot \varepsilon_0$.

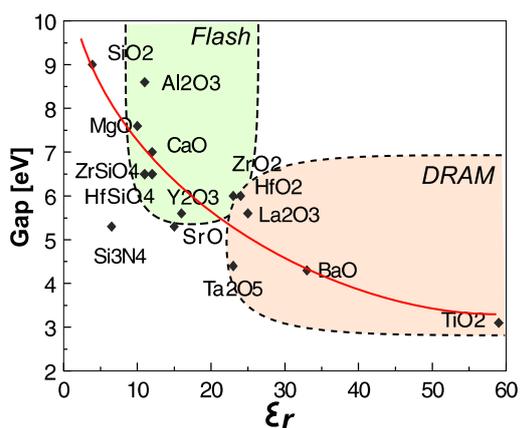
²dont la constante diélectrique relative vaut 3,9.

³D'un point de vue théorique, il serait légitime de remettre en cause une telle représentation de la silice, qui contrairement au Si, est amorphe. Néanmoins, des calculs ab initio de [Tan98] et une étude de [Nea00] ont montré que le SiO_2 conserve sa structure de bande et ses caractéristiques d'isolation de volume pour une épaisseur aussi fine que 0,7nm.

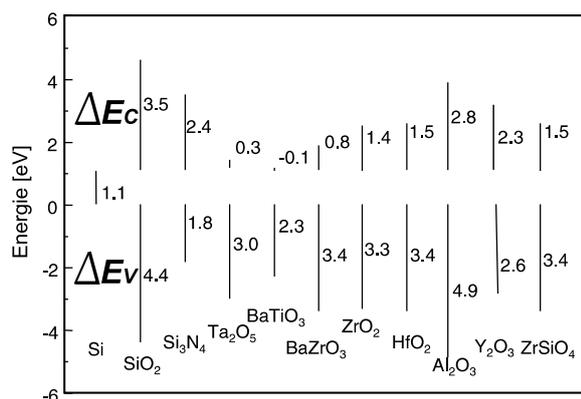
Oxyde	ϵ_r ou κ	E_G [eV]	ΔE_C [eV]	ΔE_V [eV]	
<i>SiO</i> ₂	3,9	8,9	3,2	4,6	[Wil01]
<i>Si</i> ₃ <i>N</i> ₄	7	5,1	2	2	[Wil01]
<i>Al</i> ₂ <i>O</i> ₃	9	8,7	2,8	4,8	[Wil01]
<i>HfO</i> ₂	25	5,7	1,5	3,1	[Wil01, Rob00]
<i>HfAlO</i>	9 → 20	5,6~6,4	1,9~2,3	2,5~3,1	[Yu 02]
<i>HfSiO</i>	3,9~25	8,9→5,5	3→1,4	4,9~3,1	[Ito04, Pun04]
<i>HfSiON</i>	4~20	8,9→3	3,2→1		[Kam05, Pun04]
<i>TiO</i> ₂	80~86	3,5~3,2	1,2	1,2~1	[Wil01]
<i>ZrO</i> ₂	25	5,8	1,4	3,3	[Wil01, Rob00]
<i>AlN</i>	~9	5,8	1	3,7	[Lai05]
<i>GaN</i>	9	3,4	-0,6	2,9	[Chi05]
<i>AlGaN</i>	10	3,4~5,8	0,3		[Chi05]
<i>La</i> ₂ <i>O</i> ₃	~30	~6	~2,3	~2,6	[Rob04]
<i>Y</i> ₂ <i>O</i> ₃	15	6	2,3	2,6	[Rob04]
<i>Ta</i> ₂ <i>O</i> ₅	22	4,4	0,35	2,85	[Rob04]

TABLE 3.1: Propriétés des oxydes métalliques : *High-κ*.

est possible de cibler les *High-κ* les plus intéressants en fonction des applications visées [Kit09, Eng07]. Parmi l'ensemble des diélectriques envisageables pour les mémoires non-volatiles, nous avons eu l'opportunité d'étudier des dispositifs à base d'*Al*₂*O*₃, d'*HfO*₂, d'*HfAlO*, d'*HfSiO* et d'*HfSiON*.



(a) Gap optique en fonction de la constante diélectrique.



(b) Décalage des bandes de conduction et de valence des matériaux *High-κ* sur silicium [Rob02].

FIG. 3.1: Rappels sur les *High-κ* : Constante diélectrique, E_G et ΔE_C .

3.1.2 Procédés de fabrication

Parmi les différentes techniques de fabrication des *High- κ* présentées dans la littérature nous retrouvons :

- La PVD (Physical Vapor Deposition) [Ha 02, Lee99].
- La MOCVD (Metal-Organic Chemical Vapor Deposition) [Lee02, Kim03].
- L'ALD (Atomic Layer Deposition) [Gro02, Kim02, Mor02].

La technique PVD nécessite un précurseur en phase solide qui est placé dans une chambre sous vide, en même temps que le substrat de Si sur lequel on souhaite faire le dépôt (réalisé entre 400°C et 500°C).

Le principe de base de cette méthode est de vaporiser un matériau en phase solide de façon à le déposer sur le substrat. La PVD a été utilisée dans [Gou03] pour fabriquer du Y_2O_3 et du La_2O_3 , dans [Man01] pour du Al_2O_3 , et dans [Yam05] pour du $HfSiON$.

La technique MOCVD, est basée sur des précurseurs organométalliques (comme par exemple $Ta(OC_2H_5)_5$ pour fabriquer l'oxyde de tantale Ta_2O_5). Pour cette méthode de dépôt, des précurseurs en phase vapeur sont amenés à se décomposer par pyrolyse à la surface du substrat en Si (dont la température est plus élevée que le reste de la chambre de réaction) avant éventuellement de réagir avec d'autres éléments à proximité.

La méthode ALD est celle qui a été retenue au LETI pour la fabrication des diélectriques Al_2O_3 et HfO_2 qui sont étudiés ici. Elle est détaillée dans le paragraphe 3.2.1.1.

3.2 Alliages à base de HfO_2 et Al_2O_3 : Les aluminates d'hafnium

Dans cette section nous allons étudier les aluminates d'hafnium, alliages d'oxyde d'hafnium (HfO_2) et d'alumine (Al_2O_3). Leur intérêt est qu'ils sont modulables. En effet, le contrôle de la concentration en Hf permet l'ajustement de leurs caractéristiques physico-chimiques et électriques, par exemple la température de cristallisation, le gap ou la permittivité. Avant l'étude des propriétés électriques des $HfAlO$, nous rappellerons leurs procédés de dépôt et nous présenterons une synthèse des résultats morphologiques obtenus au LÉTI sur ces matériaux.

3.2.1 Généralités sur les $HfAlO$

3.2.1.1 Procédés de fabrication

La méthode utilisée au LÉTI pour déposer des monocouches d'oxydes métalliques tels que le HfO_2 , Al_2O_3 et les $HfAlO$ sur un substrat de silicium est l'ALD ou ALCVD (Atomic Layer Chemical Vapor Deposition) [Hau00, Les02]. Cette méthode date du début

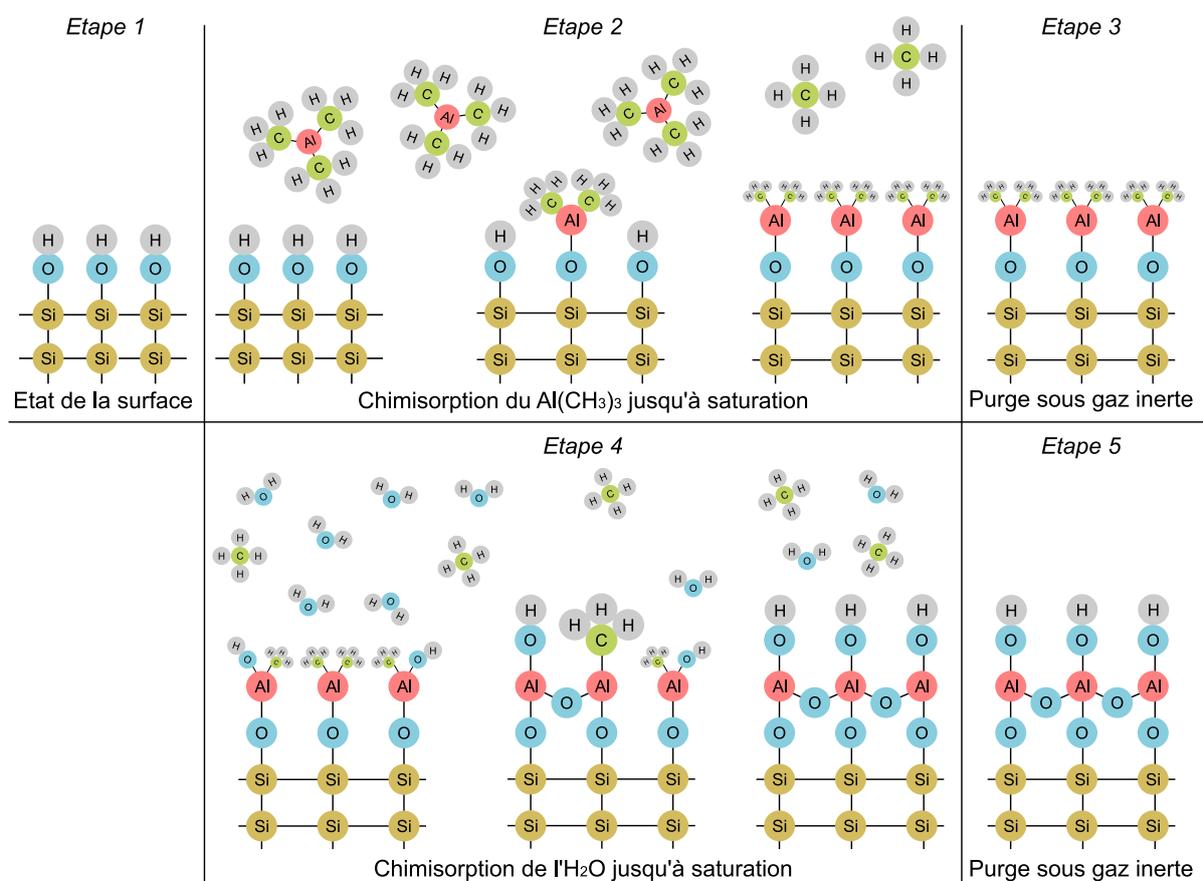


FIG. 3.2: Résumé des étapes de fabrication d'un diélectrique lors d'un cycle d'ALCVD, pour un dépôt d' Al_2O_3 [Hau00].

des années 70 et constitue une modification de la CVD. La CVD est une croissance de matière résultant de la réaction chimique entre deux précurseurs. Pour l'ALD⁴ les précurseurs sont introduits de façon alternative. Après saturation d'un des réactants avec la surface du substrat, la chambre est purgée avant l'introduction du précurseur suivant. C'est ainsi que l'on obtient une croissance contrôlée, à la couche atomique près. La vitesse de croissance est de l'ordre de 0,05 à 0,15nm par cycle dont la durée varie entre 0,5 et 5 secondes. Les épaisseurs déposées peuvent aller de 1nm à quelques dizaines de nanomètres. Les températures de dépôt sont comprises entre 150 ° C et 500 ° C et les pressions sont de quelques centaines de mTorr. Pour l' Al_2O_3 les précurseurs sont l' $Al(CH_3)_3$ (TMA : triméthyle d'aluminium) et H_2O (l'eau) ; pour le HfO_2 ce sont l' $HfCl_4$ et H_2O . Le principal défaut de cette méthode est qu'il reste toujours un peu d'eau dans la structure, ce qui induit des défauts et rend le matériau peu stable. Aussi les oxydes métalliques subissent généralement un recuit pour les stabiliser et désorber l'eau.

La figure 3.2 est un exemple des différentes étapes de dépôts ALD pour l'alumine. Cette croissance se fait à partir de $Al(CH_3)_3$ et H_2O entre 150 °C et 350 °C.

⁴contrairement à la CVD classique.

Étape 1 La surface de départ est constituée de groupements OH .

Étape 2 Le TMA est ensuite introduit à l'aide d'un flux de gaz inerte, et commence à réagir avec les groupements OH . | Il se crée ainsi des liaisons chimiques avec le substrat. Il y a alors formation de méthane. La réaction s'arrête spontanément lorsque tous les groupements $Si-OH$ en regard du flux de TMA ont réagi.

Étape 3 Le méthane et le TMA résiduels sont ensuite purgés à l'aide d'un gaz inerte.

Étape 4 De l'eau est introduite dans l'enceinte afin de créer de nouveaux sites OH , disponibles pour la prochaine couche à déposer et il y a à nouveau formation de méthane.

Étape 5 Puis une nouvelle purge des gaz est effectuée.

À la fin du cycle, la surface du Si est semblable à celle du schéma de la figure 3.2-étape 5⁵. Mais un seul cycle d'ALD ne permet pas de déposer une monocouche complète d'oxyde. En fait, simplement 14% à 17% de la surface sont recouverts en raison de l'encombrement des molécules fabriquées. Une optimisation du procédé permet d'atteindre un taux de recouvrement allant jusqu'à 40%. Ces étapes sont directement transposables au cas du HfO_2 . En changeant de précurseurs après plusieurs cycles de croissance : TMA puis $HfCl_4$, il est possible d'obtenir du $HfAlO$. Ainsi, pour indiquer le rapport des cycles ($Hf:Al$) de dépôt, on note $HfAlO-Hf:Al$ (exemple : $HfAlO-1:4$).

Pour désorber l'eau et stabiliser la matériau déposé, des recuits N_2 ou O_2 sont utilisés. L'impact de ces recuits sera discuté dans ce chapitre. Les recuits standards sont pour le HfO_2 et les $HfAlO$: 750°C sous N_2 ; et pour l' Al_2O_3 : 700°C sous O_2



Ainsi, pour résumer, on retiendra que les $HfAlO$, Al_2O_3 et HfO_2 sont déposés entre 300 et 350°C par ALCVD. Les précurseurs utilisés sont, le $HfCl_4/H_2O$ pour le HfO_2 , le $Al(CH_3)_3/H_2O$ pour l' Al_2O_3 et l'on indique le rapport de cycles de dépôt des $HfAlO$ par ($Hf:Al$). Enfin les recuits standards utilisés au Létis sont effectués à 750°C sous N_2 pour le HfO_2 et les $HfAlO$; et sous O_2 pour l' Al_2O_3 . Nous verrons par la suite comment améliorer leurs caractéristiques en ajustant ce recuit.

3.2.1.2 Stœchiométrie, gap et cristallinité

Nous venons de voir (3.2.1.1) qu'il était possible d'obtenir différents types de $HfAlO$ en modifiant le rapport de cycles de dépôt de Hf et de Al . La figure 3.3-a donne la concentration en Hf et en Al de l'alliage en fonction des cycles de dépôt. La composition a été déterminée par VPD-ICPMS (Vapor Phase Chemical Decomposition Inductively Coupled Plasma Mass Spectroscopy) et corrélée à du WDXRF (Wavelength Dispersive X-Ray Fluorescence). On remarque que les rapports des concentrations ne correspondent pas aux rapports des cycles de dépôt, le Hf ayant une concentration plus forte. Ceci est expliqué par le fait que le $HfCl_4$ consomme en partie le Al .

⁵On note qu'en fonction de sa température de dépôt et des recuits subis, l'alumine peut être soit cristalline, soit amorphe.

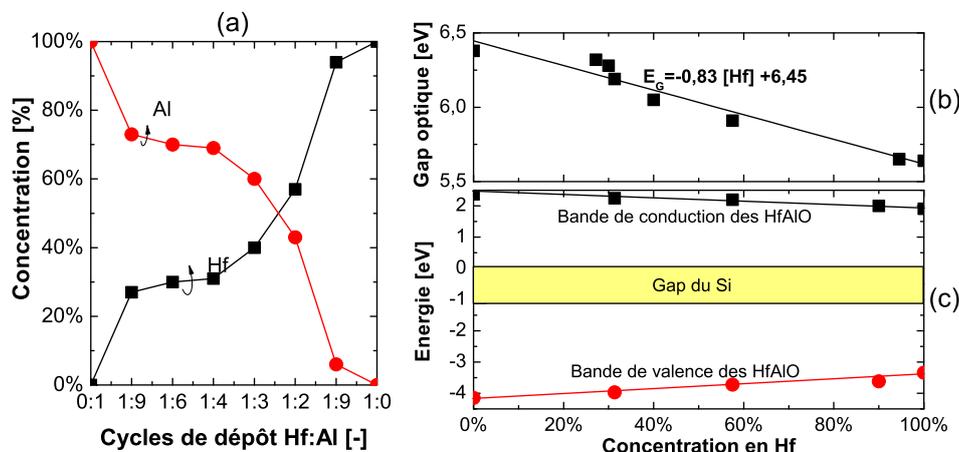


FIG. 3.3: (a) Concentration en *Hf* et en *Al* en fonction du rapport de cycles de dépôt et (b) gap optique extrait sur les *HfAlO* déposés au Léti. (c) Niveaux de la bande de conduction et de valence des *HfAlO* donnés par [Yu 02] en fonction de la concentration en *Hf*.

Le gap optique est extrait par ellipsométrie spectroscopique. L'épaisseur et le gap du film sont déterminés grâce au modèle de Tauc Lorentz [Boh04, Ngu05]. À noter que ce modèle est valide pour les matériaux amorphes, c'est pourquoi une légère erreur est faite pour le HfO_2 et les *HfAlO* riches en *Hf* qui sont cristallins (figure 3.4).

La figure 3.3-b nous donne le gap optique en fonction de la concentration en *Hf*. On observe une variation linéaire du gap, comprise entre 6,4eV pour le Al_2O_3 et 5,6eV pour le HfO_2 pur confirmant le mélange du HfO_2 et du Al_2O_3 durant le dépôt ALD [Ngu05]. Ces valeurs sont en très bon accord avec les résultats de la littérature [Ngu05, Yu 02] obtenus sur des films d'*HfAlO* déposés par ALCVD.

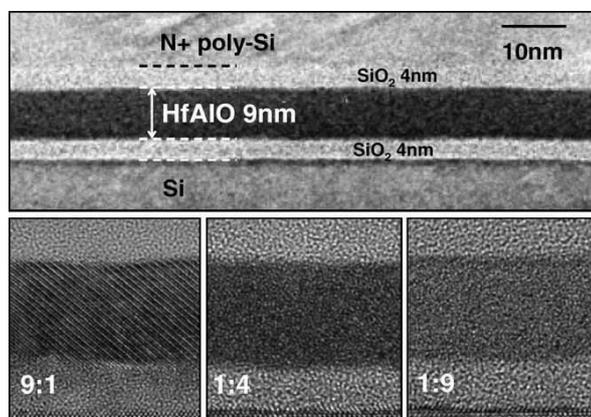


FIG. 3.4: Photographies MET d'aluminat d'hafnium, pour différentes compositions [Mol06].

$HfCl_4$: $Al(CH_3)_3$ cycle de déposition	Concentration en		Gap	ΔE_C	ΔE_V	ϵ_r	Cristallinité
	Hf [%]	Al [%]	[eV]	[eV]	[eV]	(ou κ)	après recuit
0 :1 (Al_2O_3)	0	100	6,4	2,37	3,03	8	
1 :9	27	73	6,35			11,5	Amorphe
1 :6	30	70	6,3				
1 :4	31	69	6,2	2,25	2,85	15	↓
1 :2	57	43	5,9	2,2	2,6		
9 :1	94	6	5,65			17	Polycristallin
1 :0 (HfO_2)	100	0	5,65	1,91	2,22	20	

TAB. 3.2: Concentration en Hf , en Al , gap, permittivité relative et écarts avec la bande de conduction (valence) du Si en fonction du nombre de cycles de dépôt des $HfAlO$. Les valeurs de ΔE_C et ΔE_V ont été extraites de [Yu 02].

La figure 3.3-c représente l'écart entre les bandes de conduction (resp. valence) du Si et du $High-\kappa$: ΔE_C (resp. ΔE_V). Elles sont extraites de [Yu 02] ce qui explique la légère différence entre les gaps de la figure 3.3-b et c. On constate qu'avec l'augmentation de la concentration en hafnium, la diminution du gap se répercute plus sur la bande de valence que sur la bande de conduction.

La figure 3.4 est une photographie HR-TEM⁶ (High Resolution Transmission Electron Microscopy) d'une section d'empilements tricouche HTO-4nm/ $HfAlO$ -9nm/HTO-4nm. Les deux oxydes encapsulant le $HfAlO$ sont des oxydes de silicium déposés à haute température (ici 730°C) en utilisant du silane comme précurseur ; on les appelle couramment HTO (High Thermal Oxides). Il apparaît que le $HfAlO$ -9 : 1 est cristallin dû à sa forte concentration en Hf alors que le $HfAlO$ -1 : 9 est amorphe. Ceci est en accord avec la littérature : la température de cristallisation augmente de façon monotone avec la concentration en Al , l'action du Al joue comme un stabilisateur de la phase amorphe [Zhu02].

Des mesures d'EOT ont été effectuées sur des capacités tricouches (cf. photographie MET 3.4) pour plusieurs épaisseurs et natures de $High-\kappa$. Les permittivités relatives (ou constantes diélectriques) ont été extraites et reportées dans le tableau 3.2, ainsi que les valeurs présentées dans cette section. On peut remarquer que la permittivité électrique augmente avec la concentration en Hf , et reste comprise entre celle du HfO_2 et celle de Al_2O_3 [Zhu02]. On peut donc contrôler un certain nombre de propriétés physiques des $HfAlO$ en contrôlant la concentration en Hf et en Al .

⁶En français : TEM=MET Microscope électronique à transmission.

3.2.2 Étude des charges fixes dans les aluminates d'hafnium

3.2.2.1 Effet de la concentration en *Hf*

L'une des difficultés de l'utilisation de *High-κ* en microélectronique est la présence de charges fixes pouvant causer :

- Une dispersion des tensions de seuil, difficile à maîtriser.
- Une dégradation de la mobilité à faible champ dans le cas de transistors.

Mais elles peuvent aussi améliorer les performances pour les applications mémoires en faisant office de barrière dans le *High-κ*, limitant ainsi le courant de fuite pour des applications interpoly ou oxyde bloquant [Jeo06].

Dans tous les cas, il ne faut pas confondre charges fixes et pièges. Les pièges sont des défauts qui peuvent accepter ou donner des électrons, ils peuvent échanger avec la bande de conduction et/ou de valence ; les charges fixes sont des défauts qui modifient uniquement les propriétés électriques de la structure.

Pour évaluer les charges fixes dans les empilements *High-κ*, nous avons étudié une structure capacitive tricouche typiquement intégrée en tant qu'interpoly ou oxyde bloquant (cf figure ci-contre). Le *High-κ* est encapsulé entre deux *HTO* (High Thermal Oxide) d'épaisseur visée 4nm et déposés à 730°C. Les différents *High-κ* ont été déposés par ALCVD (cf 3.2.1.1) :

- HfO_2 à 350°C, recuit à 600°C pendant 15min sous N_2 .
- Al_2O_3 à 350°C, recuit à 700°C pendant 15min sous O_2 .
- les $HfAlO$ à 300°C, recuits N_2 1min 750°C pour le 9:1 et 900°C pour le 1:4 et le 1:9).

Une couche de *HTO* de 10nm a servi de référence.

La figure 3.7 montre le décalage des caractéristiques $C_G - V_G$ vers les tensions positives comparé à la référence *HTO* de 10nm. Cela est dû à la présence de charges fixes négatives dans la couche de *High-κ*. Le décalage augmente avec la concentration en *Al*, partant du HfO_2 en allant jusqu'au Al_2O_3 . La présence de charges fixes dans l'alumine est souvent référencée dans la littérature [Lee00], ainsi dans l' $HfAlO$, elles sont attribuées à la composante alumine. Néanmoins, leur origine reste incertaine. Pour l' Al_2O_3 , plusieurs hypothèses existent :

- Une, selon laquelle l' Al_2O_3 amorphe est dissociable en éléments négatifs $(AlO_{4/2})^-$, compensés par des éléments positifs Al^{3+} . Cependant,

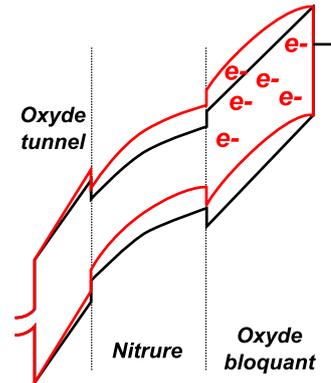


FIG. 3.5: Schéma de bandes d'une mémoire SONOS : impact des charges fixes négatives dans l'oxyde bloquant.



FIG. 3.6: Structure capacitive tricouche étudiée.

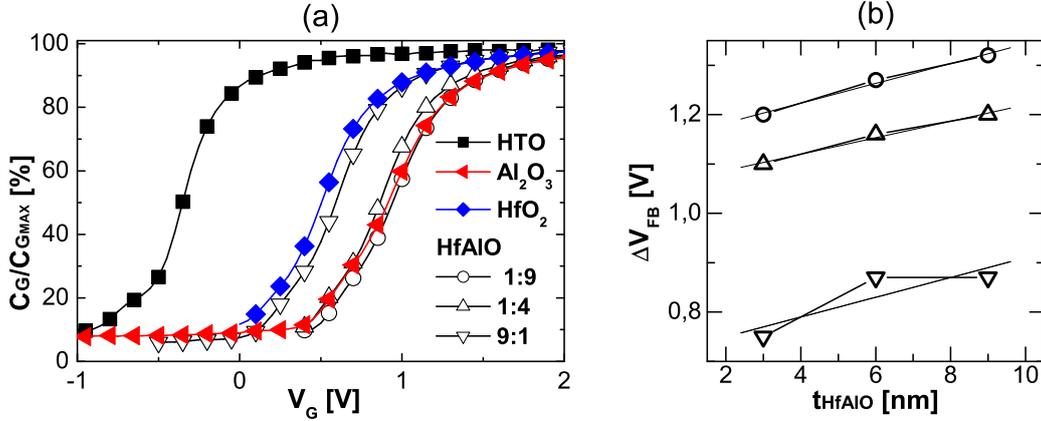


FIG. 3.7: (a) Caractéristiques $C_G/C_{G_{MAX}}-V_G$ d'empilements HTO-4nm / *High- κ* -6nm / HTO-4nm intégrant différents types de *High- κ* comparés à une référence de HTO de 10nm. (b) Tension de bande plate en fonction de l'épaisseur des *HfAlO*.

près de l'interface SiO_2/Al_2O_3 , cette compensation ne peut avoir lieu, il apparaît alors des charges fixes avec une distribution surfacique [Luc00].

- D'autres auteurs sont plutôt favorables à une distribution volumique, dont l'origine serait des groupements OH [Eri97].

Aussi il est important de localiser ces charges fixes pour connaître leur nature.

3.2.2.2 Localisation des charges fixes dans les aluminates

La figure 3.7b représente le décalage de la tension de bande plate par rapport à un dispositif sans charge (ΔV_{FB}) en fonction de l'épaisseur *HfAlO*. On observe une augmentation linéaire du ΔV_{FB} . Nous allons montrer que la dépendance du ΔV_{FB} en fonction de l'épaisseur du *High- κ* dépend de la localisation de la charge. En considérant, la répartition de charges de la figure ci-contre, où :

- t_{Haut} : Épaisseur de l'Oxyde-Haut [m].
- $t_{High-\kappa}$: Épaisseur du *High- κ* [m].
- σ_1 : Charges à l'interface Oxyde-Haut/*High- κ* en [Coulomb/m²].
- σ_2 : Charges à l'interface Oxyde-Bas/*High- κ* en [Coulomb/m²].
- ρ : Charge volumique dans le *High- κ* en [Coulomb/m³].
- ε_{SiO_2} : Permittivité du SiO_2 Haut, soit $3,9\varepsilon_0$ [F/m].
- $\varepsilon_{High-\kappa}$: Permittivité du *High- κ* [F/m].

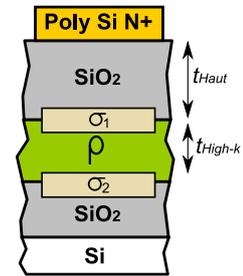


FIG. 3.8: Localisation possible des charges dans une structure capacitive tricouche.

<i>HfCl₄</i> : <i>Al(CH₃)₃</i> cycle de déposition	t_T (nm)		$\varepsilon_{High-\kappa}$ ou κ
	ellipso	MET	
9:1	3,82	3,6→3,9	17
1:4	4,99	4,5	15
1:9	5,02	4,6→5	11,5

TAB. 3.3: Ensemble des paramètres physiques des empilements : t_T et t_B mesurés par ellipsométrie ; $\varepsilon_{High-\kappa}$ extrait à partir des $C - V$.

On obtient :

$$\Delta V_{FB} = \frac{\sigma_1 \cdot S}{C_{Haut}} + \int_{t_{Haut}}^{t_{Haut}+t_{High-\kappa}} \frac{\rho \cdot S \cdot dt}{C(t)} + \frac{\sigma_2 \cdot S}{C_{Bas}} \quad (3.1)$$

$$\Delta V_{FB} = \sigma_1 \cdot \frac{t_{Haut}}{\varepsilon_{SiO_2}} + \sigma_2 \cdot \left(\frac{t_{Haut}}{\varepsilon_{SiO_2}} + \frac{t_{High-\kappa}}{\varepsilon_{High-\kappa}} \right) + \rho \cdot \left(\frac{t_{Haut} \cdot t_{High-\kappa}}{\varepsilon_{SiO_2}} + \frac{t_{High-\kappa}^2}{2 \cdot \varepsilon_{High-\kappa}} \right) \quad (3.2)$$

Les équations 3.1 et 3.2 expriment le décalage de la tension de bande plate pour les trois localisations de la charge considérées⁷.

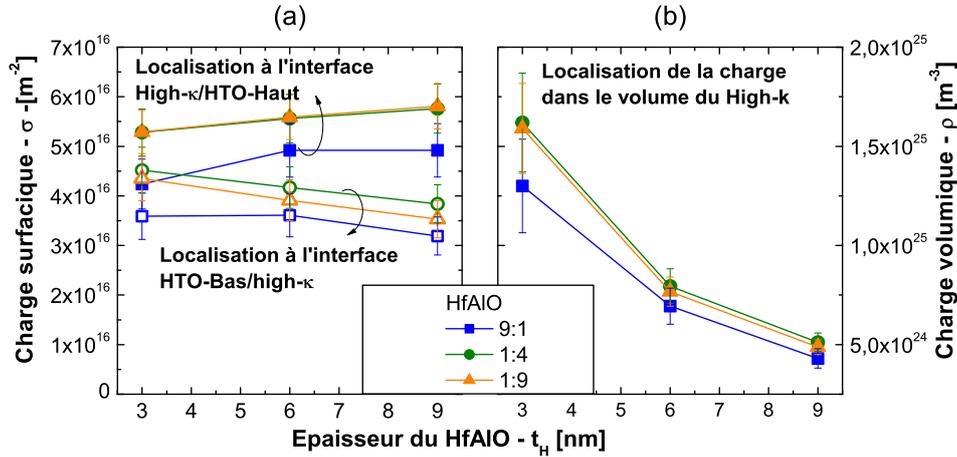


FIG. 3.9: Densité de charges fixes pour trois modes de normalisation : a) charge localisée aux interfaces *High-κ*- *HTO* Haut ou Bas ; b) charge localisée dans le volume du *High-κ*. Empilement considéré *HTO*-4nm / *High-κ* / *HTO*-4nm.

⁷Il est important de noter que ces expressions ne sont pas seulement valables pour des charges fixes mais pour tous types de charges, ainsi elles pourront être utilisées pour localiser la charge piégée. De plus, seule l'équation 3.1 est valable pour une densité volumique de charges non constante.

D'après l'équation (3.2), les charges situées à l'interface *High- κ* /Oxyde-Haut n'induisent aucune dépendance sur le ΔV_{FB} en fonction de l'épaisseur du *High- κ* alors que les charges situées à l'interface Oxyde-Bas/*High- κ* induisent une dépendance linéaire du ΔV_{FB} en fonction de l'épaisseur du *High- κ* . Quant à une densité volumique de charges constante, elle donne une dépendance parabolique.

En s'appuyant sur cette équation et sur les paramètres physiques de nos empilements extraits par MET et par mesures électriques (cf. tableau 3.3), nous avons déduit des densités de charges pour les différentes localisations (figure 3.9). En prenant en compte les erreurs et incertitudes de nos paramètres, il apparaît qu'une normalisation volumique nous donne une densité de charges fixes qui varie beaucoup avec l'épaisseur du *High- κ* , réfutant cette hypothèse. En revanche, les normalisations surfaciques donnent des densités plutôt uniformes.

Ainsi, le caractère volumique des charges fixes n'est pas dominant pour des épaisseurs d'aluminates variant entre 3 et 9nm. Malgré tout, nous ne pouvons rien affirmer sur la domination d'une interface plutôt qu'une autre pour les charges surfaciques.



En conclusion, ce comportement surfacique correspondrait à l'hypothèse faite sur l'alumine (dissociable en $(AlO_{4/2})^-$ et en Al^{3+}). Les charges fixes dans *HfAlO* seraient donc bien dues à l'*Al*, augmentant avec sa concentration, et localisées aux interfaces.

3.2.3 Courant de fuite dans les *HfAlO* : Impact des épaisseurs et de la concentration en *Hf*

Dans les mémoires EEPROM-Flash ou de type SONOS, le courant de fuite à travers les diélectriques est l'un des points les plus critiques. Des fuites dans la couche entre la grille de contrôle et la grille flottante peuvent induire une saturation de l'écriture et de l'effacement ainsi qu'une mauvaise rétention. Cette couche, en fonction des applications⁸, possède différents noms : diélectrique d'interpoly ou oxyde bloquant. Dans les mémoires Flash, ce rôle est classiquement assuré par des empilements de type Oxyde-Nitride-Oxyde (ONO) d'épaisseurs typiques 5nm-8nm-5nm [ITR07]. Ce type d'empilement est encore couramment utilisé en production.

Dans cette partie nous nous intéresserons au courant de fuite dans des empilements tricouches similaires, où le nitride a été remplacé par un *High- κ* et encapsulé entre deux HTO. Les procédés de dépôt sont similaires à ceux présentés dans le paragraphe 3.2.2.1.

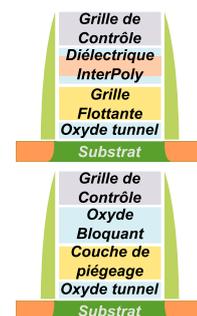


FIG. 3.10: Schémas de cellules mémoires à grille flottante et à couche de piégeage.

⁸Mémoires Flash à grille flottante continue ou mémoires de type SONOS

Pour l'ensemble des courbes présentées, nous avons exprimé le courant en fonction de V_G/EOT ⁹ de façon à décorrélérer le courant de fuite de la permittivité des matériaux étudiés. Les caractéristiques $I_G - V_G$ ont été mesurées sur des substrats de type N et P respectivement pour les tensions positives et négatives. Aussi, l'interface Si/oxydes est accumulée¹⁰ pour tous les points présentés.

Nous verrons tout d'abord l'impact de l'épaisseur des *HTO* Haut et Bas sur le courant de fuite. Puis nous mettrons en évidence l'importance de la permittivité électrique du matériau *High-κ* ainsi que la difficulté de caractériser le courant dans des empilements qui piègent des électrons. Enfin, nous étudierons l'activation du courant en température.



FIG. 3.11: Structure capacitive tricouche étudiée.

3.2.3.1 Influence de l'épaisseur des *HTO* dans un tricouche

Le premier objectif de l'intégration des *High-κ* dans les mémoires non-volatiles est d'augmenter le couplage entre la grille de contrôle et la zone de piégeage, sans augmenter les courants de fuite. La structure tricouche, couramment utilisée (type ONO), présente les meilleures performances en terme d'isolation mais pour un gain en couplage réduit. La suppression de l'un d'eux (bicouche) ou des deux (monocouche), sans augmenter considérablement le courant de fuite, permettrait une augmentation notable du couplage.

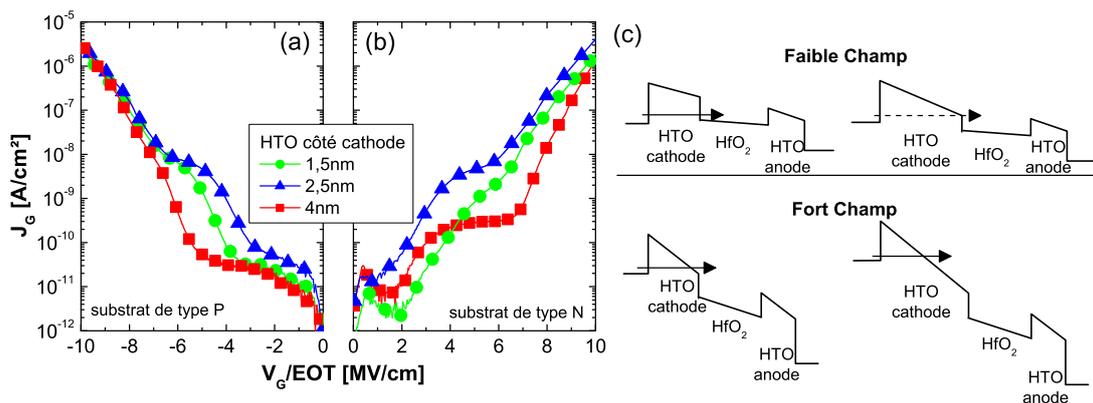


FIG. 3.12: Caractéristiques $J_G - V_G/EOT$ d'un empilement $HTO/HfO_2-8nm/HTO-4nm$, pour différentes épaisseurs de *HTO* côté cathode : 1,5nm, 2,5nm et 4nm.

⁹Obtenu par mesure CV.

¹⁰Porteurs majoritaires.

Ici nous nous intéressons à un tricouche à base d'oxyde d'hafnium intégrant plusieurs épaisseurs de HTO. L'objectif est d'évaluer l'impact de l'épaisseur des HTO sur le courant de fuite..

La figure 3.12 (resp. figure 3.13) regroupe les caractéristiques $J_G - V_G/EOT$ d'empilement où l'on a fait varier l'épaisseur du HTO côté cathode (resp. anode), c'est à dire le HTO du côté de l'injection, soit :

- le HTO Bas pour les tensions positives (resp. negatives)
- le HTO Haut pour les tensions négatives (resp. positives)

On remarque que le HTO côté cathode contrôle fortement la conduction à faible champ. En effet, comme le montre le schéma de la figure 3.12-c, à bas champ, l'injection des électrons depuis la cathode s'effectue par *tunnel* direct. Aussi l'épaisseur de l'oxyde côté cathode a une importance prépondérante pour la conduction. Tandis qu'à fort champ, l'injection est de type Folwer-Nordheim et l'épaisseur d'oxyde n'est plus entièrement vue par les électrons.

L'impact du HTO côté anode semble moins important. À fort champ, le schéma de bande de la figure 3.13-c montre qu'un électron injecté dans l'empilement ne devrait pas voir le HTO côté anode. En effet, les électrons ont une énergie supérieure à la bande de conduction du HTO côté anode. Or, on observe une légère influence. Elle souligne que les électrons peuvent perdre une partie de leur énergie et voir l'impact de l'épaisseur de ce HTO. Cela coïncide avec un libre parcours moyen dans le SiO_2 d'environ 1nm [Bri97].



On peut donc conclure qu'une réduction du HTO côté anode n'est pas préjudiciable à l'isolation. Aussi la suppression du HTO Haut n'impacterait pas l'écriture et la rétention. Il est à noter que le même comportement a été observé pour le nitrure dans les empilements ONO [De 99].

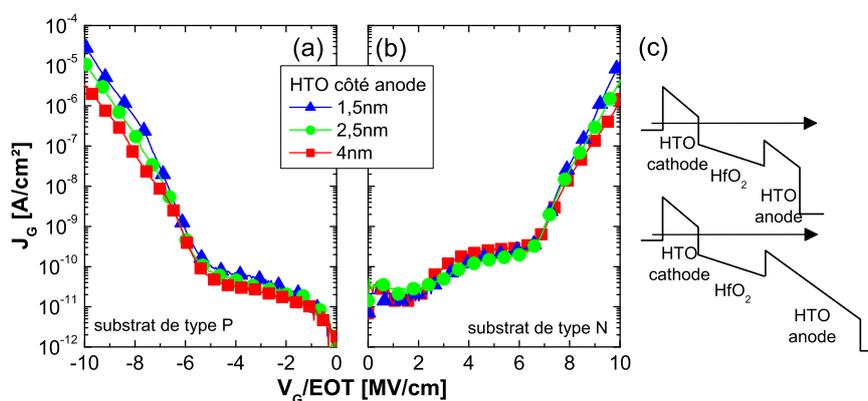


FIG. 3.13: Caractéristiques $J_G - V_G/EOT$ d'un empilement HTO-4nm/ HfO_2 -8nm/HTO, où l'épaisseur du HTO côté anode varie : 1,5nm, 2,5nm et 4nm.

3.2.3.2 Piégeage dans le *High-κ* épais : perturbations des mesures de courant

La présence de pièges dans les *High-κ* est un point critique pour leur intégration dans les mémoires. Ils perturbent lourdement la caractérisation des empilements, peuvent augmenter le courant de fuite à faible champ par des mécanismes assistés par pièges (cf partie 2.1.2), réduire le courant à fort champ par modification de la structure de bande.

La figure 3.14a représente la caractéristique $J_G - V_G/EOT$ effectuée lors d'un aller-retour en tension jusqu'à 11MV/cm. La figure 3.14b regroupe les caractéristiques $J_G - Temps$ à champs fixes. Lors du retour de la caractéristique $J_G - V_G/EOT$, le courant chute. De même sur $J_G - Temps$, le courant décroît en fonction du temps. Le piégeage important dans des tricouches contenant des *High-κ* peut expliquer ce phénomène. En effet, les électrons piégés créent une charge négative qui modifie les barrières de potentiel, rendant l'injection plus difficile et ainsi diminuant le courant à travers l'empilement, comme le montre la figure 3.14c.

Ainsi, on observe une variation de près de deux décades entre la caractéristique aller et celle retour. On notera que la valeur du courant mesurée lors du retour coïncide avec les valeurs de saturation des caractéristiques $J_G - Temps$. De plus, on peut expliquer la variation plus faible du courant à fort champ, par :

- Un piégeage à des temps plus rapides, inférieurs à 1s.
- Une modification relative du champ plus faible, causée par les charges piégées (même quantité de charge mais champ plus fort).

Cela nous informe sur la difficulté de caractériser les empilements à base de *High-κ*. En effet, en fonction du temps d'intégration de la mesure, les valeurs du courant varient.

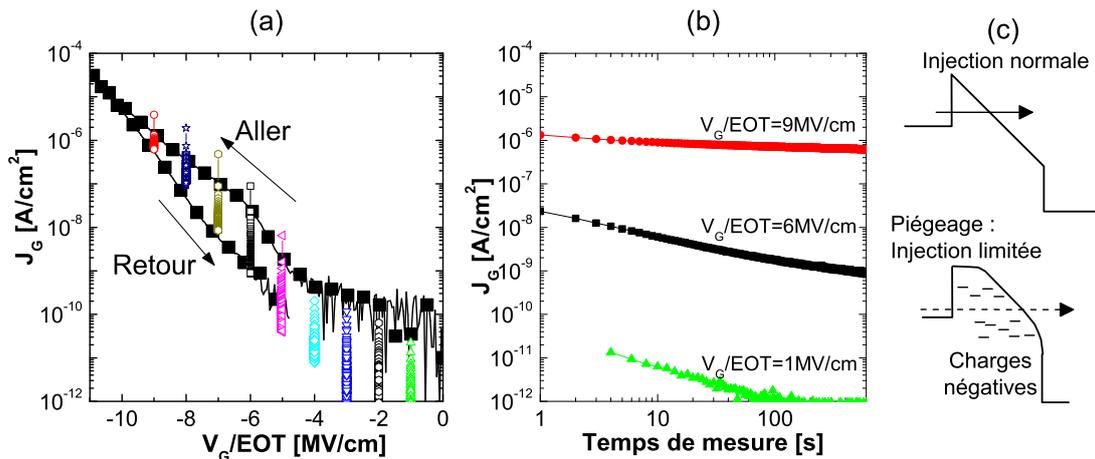


FIG. 3.14: a) Caractéristiques $J_G - V_G/EOT$ d'un empilement HTO-4nm/ *HfAlO*-9nm/HTO-4nm en aller-retour. b) Évolution de la densité de courant avec le temps de mesure. c) Modification du diagramme de bandes dues au piégeage dans un oxyde.

Il faut donc se placer dans des conditions similaires au régime de fonctionnement où l'on utilise ces champs :

- A fort champ, une mesure rapide, comme lors de l'écriture et l'effacement des dispositifs mémoires.
- A faible champ, une mesure à densité de courant établi pour simuler la rétention. Dans ce cas, nous avons donc réalisé des mesures en courant à V_G/EOT constant, et nous avons reporté la valeur du courant établi au bout de 10 minutes.

3.2.3.3 Influence de la couche $High-\kappa$ à faible et à fort champ

Les figures 3.16 représentent la densité de courant d'empilements OHO (Oxyde/ $High-\kappa$ /Oxyde) avec différentes compositions de $HfAlO$. Différents types de mesures sont présentés. La figure 3.16a représente la densité de courant pour une mesure rapide, avec un temps d'intégration de $640\mu s$. La figure 3.16b représente la densité de courant établi pour des temps de 10min. À fort champ, dans les deux cas, on peut observer que malgré une structure plus cristalline (cf figure 3.4) et une hauteur de barrière plus faible (cf tableau 3.2), la capacité d'isolation augmente avec la concentration en Hf . Cette tendance est conservée pour des épaisseurs de 3nm à 9nm de $High-\kappa$ (cf. figure 3.18).

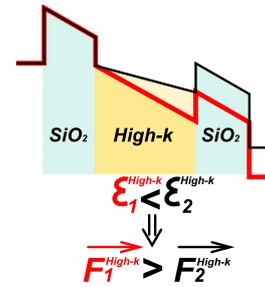


FIG. 3.15: Schéma de bande d'un tricouche pour différentes permittivités du $High-\kappa$.

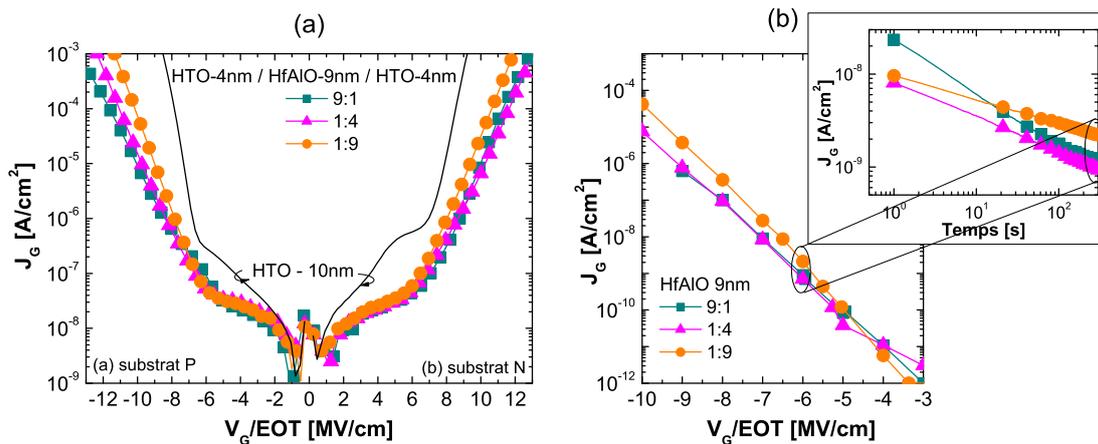


FIG. 3.16: Caractéristiques $J_G - V_G/EOT$ d'empilements HTO-4nm / $HfAlO$ -9nm / HTO-4nm pour différentes concentrations en Hf pour a) des mesures standards et b) des mesures à densité de courant établi.

Cette meilleure isolation s'explique par la réduction du champ dans le *High-κ* avec l'augmentation de sa permittivité. En effet, l'équation 3.3 exprime le champ dans le *High-κ* ($\overrightarrow{F_{High-\kappa}}$) en fonction du champ dans le *SiO₂* ($\overrightarrow{F_{SiO_2}}$) en l'absence de charges dans l'empilement. Il apparaît que plus la permittivité du *High-κ* ($\varepsilon_{High-\kappa}$) est grande, plus le champ est faible. De plus, avec la concentration en *Hf*, l'abaissement du ΔE_C aurait pu augmenter le courant de fuite, ce qui n'est pas le cas. Aussi l'abaissement du ΔE_C n'est donc pas critique pour un tricouche à fort champ. On peut estimer qu'à fort champ, à V_G/EOT fixé, le courant total est dominé par le champ dans le *High-κ*.

$$\overrightarrow{F_{High-\kappa}} = \frac{\varepsilon_{SiO_2}}{\varepsilon_{High-\kappa}} \cdot \overrightarrow{F_{SiO_2}} \quad (3.3)$$

Pour les champs inférieurs à 6MV/cm, on observe une inversion de tendance pour les densités de courant établi (cf figure 3.16-b). Les *HfAlO* riches en *Al* ont un courant de fuite plus faible. Qualitativement, pour des champs inférieurs à $V_G/EOT=6MV/cm$, les électrons injectés depuis le substrat à travers le *HTO* Bas en *tunnel* direct arrivent dans le gap du *High-κ*. La valeur du gap du *High-κ* devient alors critique et la meilleure isolation du *HfAlO* riche en *Al* peut être due à son gap plus important (cf figure 3.3). De plus, une densité de pièges plus importante pour les *HfAlO* riches en *Hf* peut favoriser les courants assistés par pièges prépondérants à faible champ.

La figure 3.18 montre des densités de courant très similaires quelle que soit l'épaisseur du *High-κ*. La normalisation des champs induit des champs identiques dans tout l'empilement (cf. équation 3.3), il est donc cohérent que le courant de fuite soit, à fort champ, identique pour différentes épaisseurs de *High-κ*.

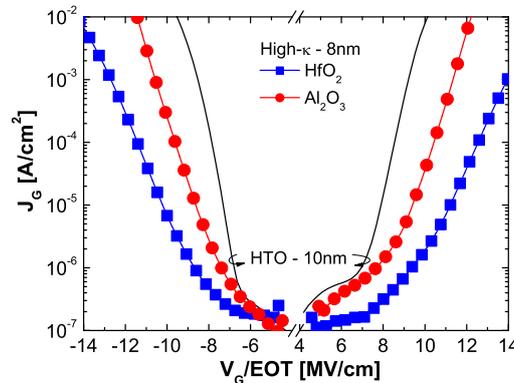


FIG. 3.17: Caractéristiques $J_G - V_G/EOT$ d'empilements *HTO*-4,7nm / *High-κ*-8nm / *HTO*-5,6nm pour différents types de *High-κ* : *HfO₂* et *Al₂O₃* pour des mesures standards.

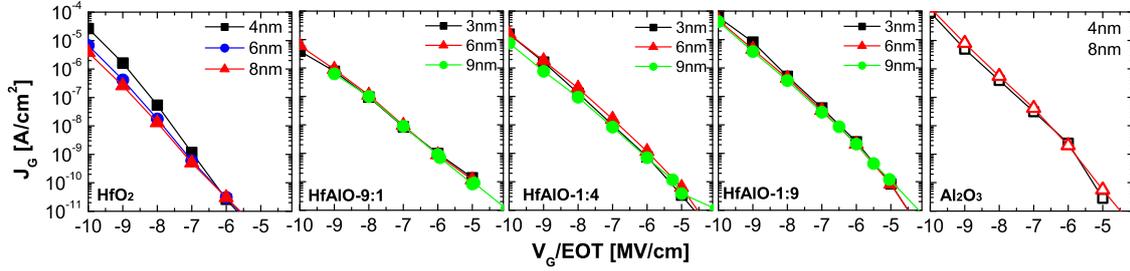


FIG. 3.18: Ensemble des caractéristiques $J_G - V_G/EOT$ d'empilements $HTO/High-\kappa/HTO$ pour plusieurs concentrations en Hf et pour différentes épaisseurs de $High-\kappa$.

3.2.3.4 Modes de conduction et d'activation en température

Dans ce paragraphe nous étudierons les mécanismes de conduction à travers les empilements tricouches $HTO/HfAlO/HTO$, pour les différentes compositions de $HfAlO$. Nous avons effectué des mesures $J_G - V_G/EOT$ à différentes températures, de $25^\circ C$ à $200^\circ C$ (figure 3.19). On observe un courant de fuite fortement activé en température. Aussi, à partir de cette mesure, nous avons tracé la courbe d'Arrhenius à $10MV/cm$ pour les trois compositions d'aluminates (figure 3.19-insert). Le caractère linéaire de ces courbes indique une activation du niveau de courant en :

$$J_G = \exp\left(-\frac{Ea}{kT} + cst\right) \quad (3.4)$$

$Ea [J]$: Énergie d'activation.

Le tableau 3.20 regroupe les différentes énergies d'activation (Ea) relatives à nos différents $High-\kappa$. Il apparaît clairement que l'énergie d'activation augmente avec la concentration en Hf , ce qui signifie que le niveau de courant de fuite de l'empilement est plus sensible à la température. De plus les énergies d'activation du $HfAlO$ sont comprises entre celles de l' Al_2O_3 et du HfO_2 , ce dernier présentant la plus forte valeur. Néanmoins, à fort champ, le $HfAlO-9:1$ présente encore la meilleure isolation par rapport au $HfAlO$ moins riche en Hf , même à $200^\circ C$.

Pour identifier clairement les différents modes de conduction dans nos empilements tricouches, nous avons tracé pour chaque composition de $HfAlO$, les diagrammes de Hill [Hil71] (figure 3.21). Ils représentent la fonction réduite Γ en fonction de $\sqrt{F} \cdot T^{-1}$, où Γ est donnée par l'équation 3.5.

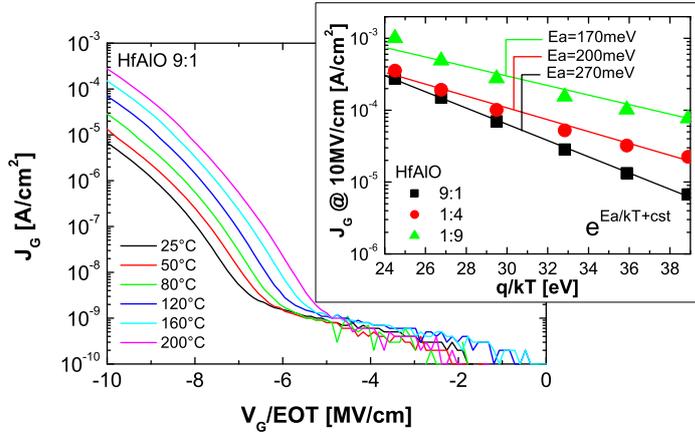


FIG. 3.19: Caractéristiques $J_G - V_G/EOT$ d'un empilement HTO-4nm/ $HfAlO$ -3nm/HTO-4nm activé en température. Insert : Extraction de l'énergie d'activation à 10MV/cm E_a tel que : $J_G = e^{-\frac{E_a}{kT} + cst}$.

$HfCl_4 : Al(CH_3)_3$	E_a
1 (HfO_2)	340meV
9 : 1 (HfO_2)	260meV
1 : 4 (HfO_2)	200meV
1 : 9 (HfO_2)	170meV
0 (Al_2O_3)	130meV

FIG. 3.20: Énergies d'activation extraites des tracés d'Arrhenius pour les différentes compositions de $HfAlO$.

$$\Gamma_{exp} = \frac{J_G}{T^4 \cdot \exp\left(-\frac{\phi_t}{kT}\right)} \quad (3.5)$$

- F [V/m] : Champ dans le *High-κ*.
 T [K] : Température.
 ϕ_t [J] : Profondeur du piège.
 J_G [A/cm^2] : Densité de courant expérimentale

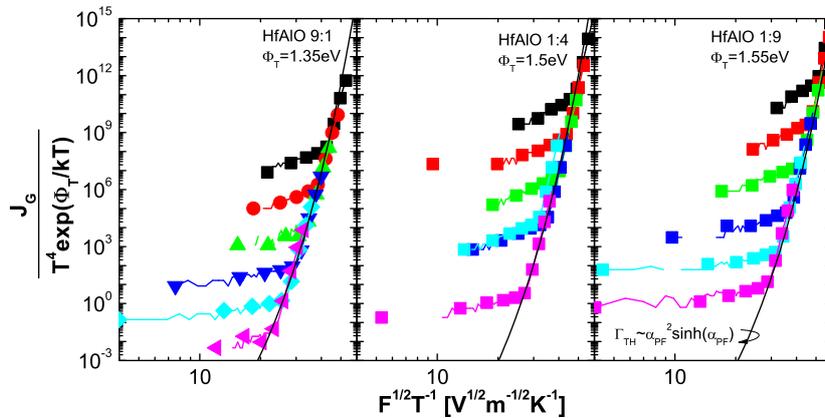


FIG. 3.21: Diagrammes de Hill correspondant aux caractéristiques $J_G - V_G/EOT$ de l'empilement HTO-4nm/ $HfAlO$ -3nm/HTO-4nm pour des températures variant entre 25°C et 200°C.

Ratio de cycles de dépôt $HfCl_4 : Al(CH_3)_3$	Profondeur de piège
1 (HfO_2)	$\phi_t=1,1eV$
9 :1	$\phi_t=1,35eV$
1 :4	$\phi_t=1,5eV$
1 :9	$\phi_t=1,55eV$

TAB. 3.4: Profondeurs des pièges assurant une conduction de type Poole-Frenkel extraites à partir des diagrammes de Hill pour chaque composition de $HfAlO$.

Pour chaque composition, on peut trouver une valeur de ϕ_t pour laquelle les caractéristiques se superposent à fort champ, ce qui traduit une conduction de type Poole-Frenkel, assistée par un piège au niveau énergétique ϕ_t . Les profondeurs de piège pour chaque composition de $HfAlO$ sont résumées dans le tableau 3.4.

On constate que plus la teneur en HfO_2 diminue, plus la profondeur de piège augmente. Notons que la profondeur de piège extraite a été trouvée indépendante de l'épaisseur de $HfAlO$. On constate que les caractéristiques se superposent avec la courbe théorique de conduction par Poole-Frenkel pour un matériau de type semi-cristallin au sens de Hill [Mot71, Hes86].

Les mécanismes de conduction à travers une seule couche de $HfAlO$ sont apparus comme plus complexes et grandement dépendants de l'électrode injectante (canal ou grille de contrôle) et de l'épaisseur du $HfAlO$. Nous avons décidé de nous intéresser aux tensions négatives, correspondant à l'injection depuis la grille de contrôle.

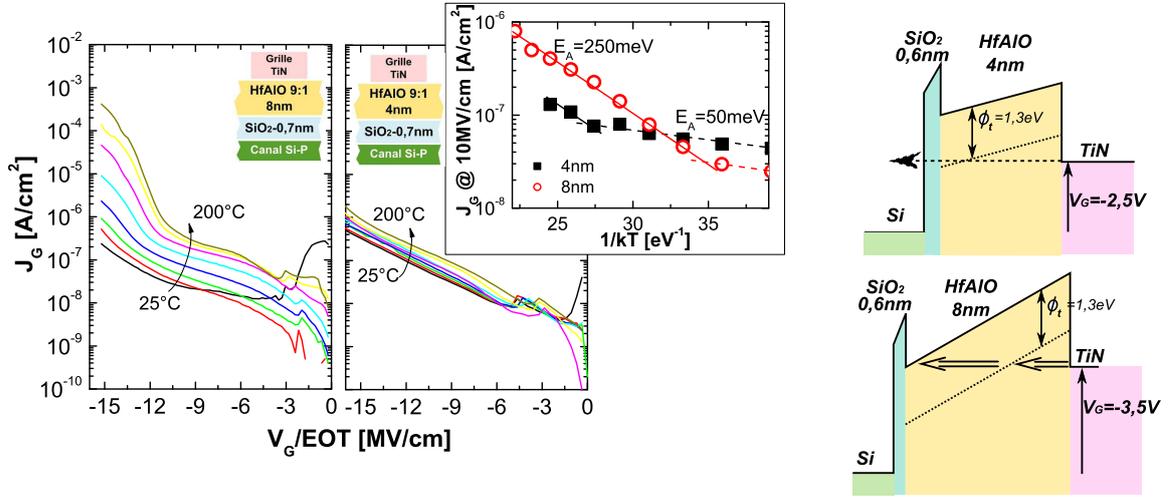
Tout d'abord, la figure 3.23a montre une faible dépendance du courant de fuite en fonction de la température dans le cas de 4nm d' $HfAlO$. En effet, la densité de courant n'augmente que d'une décade de 25°C à 200°C pour une épaisseur de $HfAlO$ de 4nm alors qu'elle atteint 4 décades pour le 8nm.

Pour mieux comprendre ce mécanisme, nous avons tracé le diagramme de bandes des deux structures sur la figure 3.23b. En s'appuyant sur les mesures électriques, nous avons estimé le SiO_2 interfacial de 0,6nm et la permittivité du $HfAlO$ -9:1 de $18\epsilon_0$.

Dans le cas d'un $HfAlO$ riche en Hf , nous avons aussi considéré une profondeur de piège de $\sim 1,3eV$, comme pour le HfO_2 pur [Ler04] correspondant à des lacunes d'oxygène [Gav06].



FIG. 3.22: Structure capacitive monocouche étudiée pour la dépendance en épaisseur et en température.



(a) Caractéristiques $J_G - V_G/EOT$ d'un monocouche $HfAlO$ pour différentes températures. Insert : Extraction de l'énergie d'activation E_a tel que : $J_G = e \frac{E_a}{kT} + cst$ pour différentes épaisseurs.

FIG. 3.23: Activation en température des $HfAlO$, impact de l'épaisseur.

Nous avons effectué la même analyse pour les $HfAlO$ riches en Al avec les défauts du Al_2O_3 [Li 06, Joh02]. Dans tous les cas, on observe que le piège est plus activé pour les *High-κ* épais à champ électrique donné. En effet pour les couches fines, l'énergie des pièges dans le volume du *High-κ* est supérieure au niveau d'injection, donc les pièges participent plus difficilement à la conduction. Pour les couches épaisses, le niveau des pièges est plus bas que le niveau d'injection et favorise la conduction assistée par pièges.



On peut conclure que plus l'aluminate d'hafnium est riche en Hf , plus il est activé en température et moins ses pièges sont profonds d'après la normalisation de Hill. On peut donc s'attendre à des courants de fuite plus importants à faible champ pour l' $HfAlO$ riche en Hf . De plus, la profondeur des défauts joue un rôle très important sur la conduction dans les monocouches épaisses. En effet, en fonction de leurs positions ils peuvent, ou non, participer à la conduction et modifier les caractéristiques des empilements.

3.2.4 Propriétés de piégeage des $HfAlO$

Dans cette partie nous nous intéresserons aux propriétés de piégeage d'empilements interpoly intégrant des *High- κ* . Pour les évaluer, nous observons le décalage des caractéristiques $C_G - V_G$ quand on applique une tension de stress (cf figure 3.25a). Ainsi nous obtenons la figure 3.25b qui représente le décalage de la tension de bande plate en fonction du temps de stress. On observe un décalage de V_{FB} vers les tensions positives, caractéristique d'un piégeage d'électrons.

Il apparaît que ΔV_{FB} augmente avec la concentration en *Hf* à V_G fixée. Cette tendance est aussi observée à $V_G/EOT = 10MV/cm$ ¹¹ (figure 3.26). Les propriétés de piégeage pourraient être dues à l'augmentation de la cristallinité (figure 3.25c) [Mol06, Buc06]. En effet, un matériau cristallin, plus ordonné, pourrait être plus sensible à des défauts (liaisons pendantes, lacunes d'oxygène, présence d'O - H, stœchiométrie, etc.).



FIG. 3.24: Structure capacitive tricouche étudiée.

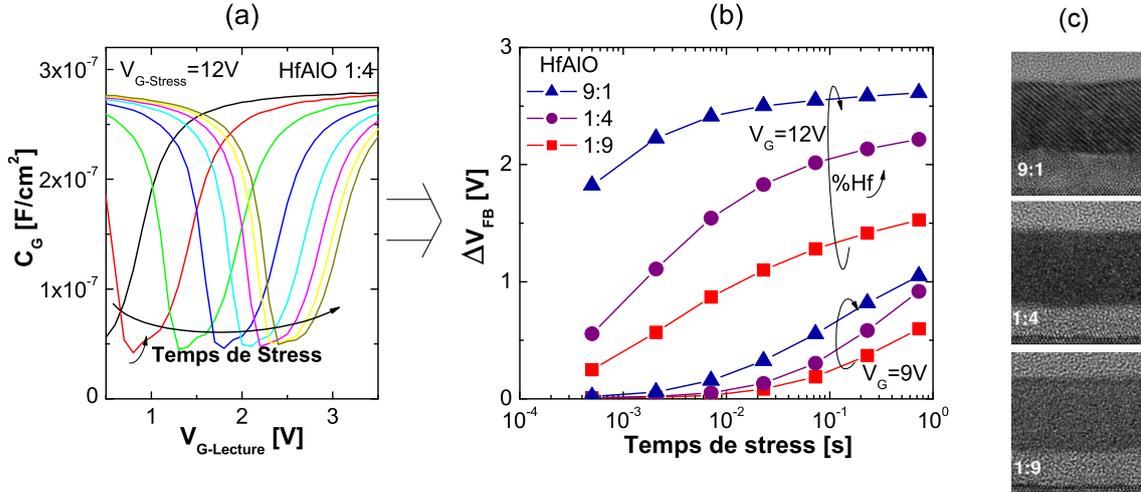


FIG. 3.25: a) L'ensemble des caractéristiques $C_G - V_G$ d'un empilement HTO-4nm /*High- κ* / HTO-4nm lorsqu'il subit des stress de tension de 12V. b) Caractéristiques de programmation $\Delta V_{FB} - Temps_{Stress}$ extraits du décalage des $C_G - V_G$ pour différentes compositions de $HfAlO$ pour $V_G = 9V$ et 12V. c) Rappel des photographies MET des $HfAlO$ présentées en 3.2.1.2.

¹¹À $V_G/EOT = 10MV/cm$, le champ dans l'oxyde *tunnel* est le même pour l'ensemble des dispositifs. On s'assure ainsi d'avoir les mêmes conditions d'injection et de ne comparer que les propriétés de piégeage du matériau *High- κ* .

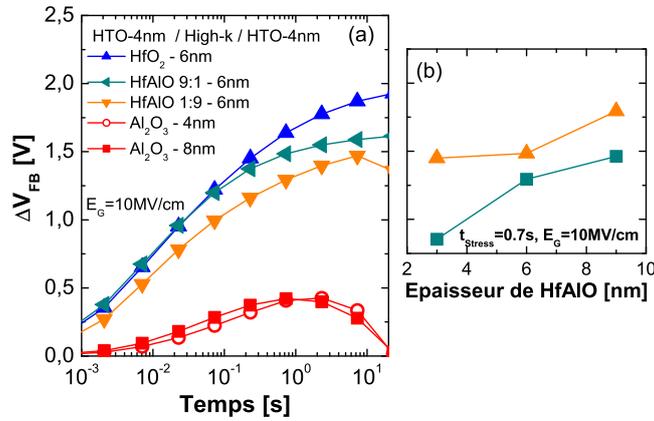


FIG. 3.26: a) Caractéristiques de programmation $\Delta V_{FB} - Temp_{Stress}$ d'empilements HTO-4nm /*High-κ*/ HTO-4nm pour différents types de *High-κ*. b) Décalage de la tension de bande plate pour différentes concentrations de *HfAlO* pour $V_G/EOT = 10 \text{ MV/cm}$ et $t_{Stress} = 0,7 \text{ s}$.

En s'appuyant sur ces mesures, nous avons extrait la charge piégée durant le stress. Deux types de normalisation ont été appliqués (cf. équation 3.2), une normalisation surfacique où l'on a considéré la charge à l'interface HTO-Bas et *HfAlO*, et une normalisation volumique. On observe une très bonne normalisation surfacique quelle que soit l'épaisseur de la couche de *High-κ* en regard de la normalisation volumique. Il semble donc que l'on piège principalement à l'interface *HTO-Bas/HfAlO* ou dans les premiers nanomètres du matériau, et que la contribution volumique du *High-κ* est négligeable pour les champs et les épaisseurs considérés.

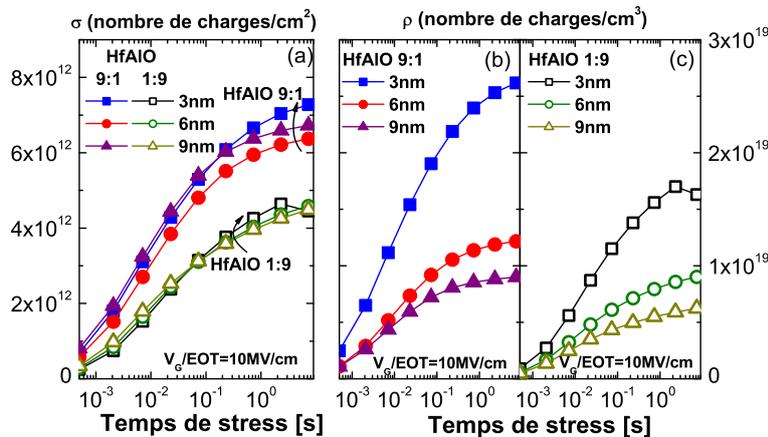


FIG. 3.27: Densité de charges piégées pour deux modes de normalisation : a) la charge est localisée à l'interface *SiO₂- High-κ*; b) et c) la charge est localisée dans le volume du *High-κ*. Empilement considéré HTO-4nm / *High-κ*/ HTO-4nm.

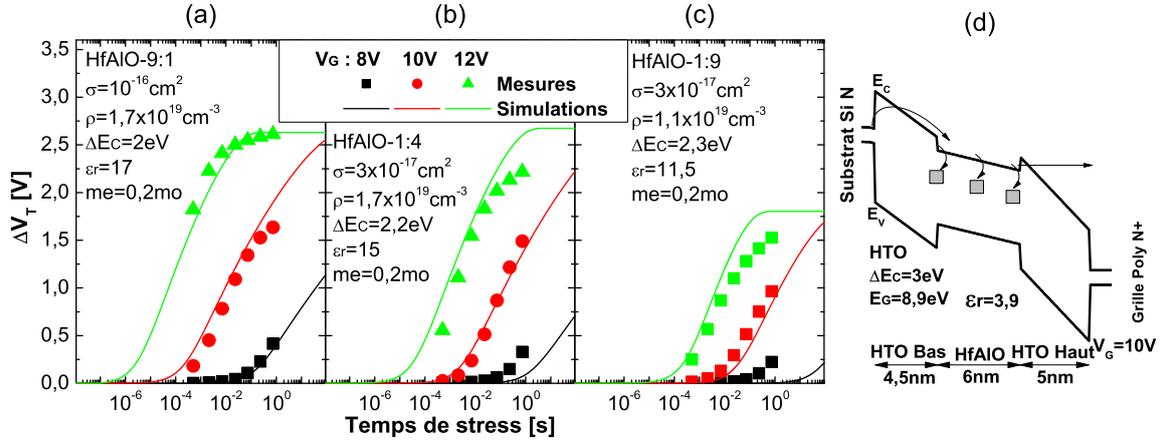


FIG. 3.28: Simulations des caractéristiques de piégeage d'empilements tricouche à base de a) $HfAlO$ -9:1, b) $HfAlO$ -1:4 et c) $HfAlO$ -1:9. d) Diagramme de bandes de l'empilement simulé.

Nous avons utilisé le modèle de piégeage présenté dans la partie 2.3.2 pour comprendre d'où provenaient les différentes cinétiques de piégeage entre les $HfAlO$. La figure 3.28 regroupe plusieurs simulations d'empilements $HTO/HfAlO/HTO$ pour les trois compositions précédemment étudiées. Nous avons ajusté les paramètres des matériaux (densité de pièges, section de capture et profondeur de pièges) pour obtenir une bonne corrélation entre les mesures et les simulations. On observe que les principales différences entre les $HfAlO$ viennent d'une :

- Section de capture plus importante du $HfAlO$ -9:1.
- Densité de pièges plus faible pour le $HfAlO$ -1:9.
- Réduction du champ dans l' HTO -Bas avec la diminution de la permittivité du $HfAlO$.

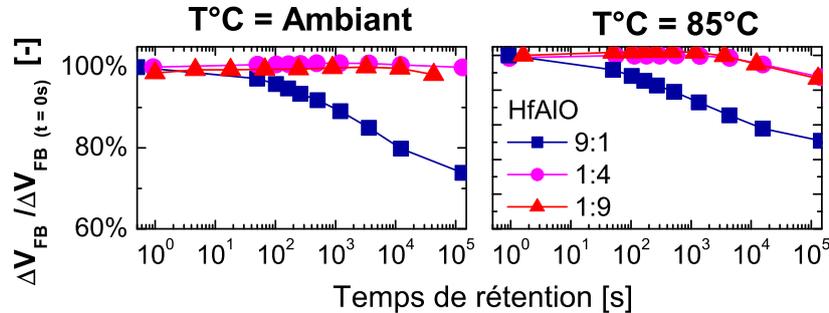


FIG. 3.29: Caractéristiques de rétention, $\Delta V_{FB} / \Delta V_{FB}@t_{Ret} = 0s - Temps$, d'empilements HTO -4nm / $HfAlO$ -6nm / HTO -4nm pour différentes compositions de $HfAlO$. Condition de programmation initiale fixée à $\Delta V_{FB} = 1,5 \text{ V}$.

Enfin, nous nous intéressons à la dynamique de dépiégeage des charges que l'on a pu injecter dans notre empilement OHO. La figure 3.29 représente la rétention à température ambiante et à 85°C pour différentes compositions de *HfAlO*. On observe une perte de charge beaucoup plus importante dans le cas du *HfAlO* riche en *Hf* alors que les *HfAlO*-1:4 et 1:9 gardent presque l'intégralité à 25°C après 10⁵s. On observe aussi une forte énergie d'activation avec une augmentation de la perte de charge d'un facteur deux à 85°C. Cette tendance est aussi observée dans la littérature pour des mémoires de type SONOS intégrant des *High- κ* en couche de piégeage [Tan04]. Cela peut provenir de la section de capture plus forte du *HfAlO*-9:1 et des pièges moins profonds, observée en écriture, entraînant un coefficient d'émission plus fort (cf équation 2.52).



En conclusion de cette partie, nous avons observé un piégeage plus important dans le cas des *HfAlO*-9:1 dû à une section de capture plus importante entraînant une moins bonne rétention de ces charges. De plus, il est apparu que les charges piégées étaient localisées près des interfaces *HTO/HfAlO* et que la contribution du volume serait négligeable.

3.3 Impact de la nitruration sur les *HfAlO*

Comme nous l'avons vu plus haut (section 3.2), les *HfAlO* présentent des pièges et des défauts souvent associés à des lacunes d'oxygène [Tak04, Gav05, Gav06]. Plusieurs approches existent pour compenser ces lacunes :

- l'incorporation d'azote [Yu 07, Jun02], de phosphore [Tse06] ou d'oxygène par des recuits nitrurants ou oxydants pour modifier le volume du matériau.
- des préparations de surfaces à base d'ozone (O_3) [Doh03] ou de plasmas nitrurants pour agir sur les défauts d'interface.

Dans cette partie, nous allons nous intéresser aux effets des recuits nitrurants sur les propriétés des *HfAlO*. Dans ce cadre, plusieurs types de traitements thermiques ont été appliqués. Il faut distinguer deux familles :

1. Les recuits après dépôt (en anglais : PDA pour Post Deposition Annealing), ont été effectués à la suite du dépôt du *High- κ* pour le stabiliser et/ou ajuster ses propriétés. Dans cette étude, deux types de recuits post-dépôt ont été réalisés :
 - Le recuit nitrurant : NH_3 à 750°C pendant 30min, visant à incorporer de l'azote dans le *High- κ* .
 - Le recuit standard : N_2 à 750°C pendant 1min.
2. Les recuits haute température ou recuits Source/Drain, sont effectués après l'implantation des Source/Drain pour activer les dopants. Dans notre technologie, ils sont donc effectués après la réalisation complète de l'empilement de grille. L'étude que nous allons présenter a été réalisée sur des capacités. Les recuits à haute température simulent ces recuits d'activation :
 - N_2 à 950°C pendant 15s.
 - ou N_2 à 1050°C en spike.

3.3.1 Impact sur les propriétés physico-chimiques

Avant de commencer l'étude électrique, il est important de bien comprendre comment le matériau réagit au recuit nitrurant. Aussi, voici un résumé de l'étude morphologie qui a été effectuée au Léti [Gra07]. On notera que l'on étudie ici des couches de *High-κ* de 6nm déposées sur substrat de Si à 300°C. L'oxyde natif est estimé à 0,7nm.

3.3.1.1 Incorporation de l'azote dans le film

La figure 3.30 représente le profil Auger de la concentration en azote dans la couche pour différents types de *High-κ*, du HfO_2 à l' Al_2O_3 . On suppose que le taux d'abrasion est le même pour toutes les analyses. Cela nous informe sur l'incorporation et la localisation de l'azote dans l'empilement après nitruration. Aussi, on observe que :

- l'alumine présente une incorporation très faible. Seules quelques traces d'azote (<5%) sont observées de façon uniforme.
- Quand la concentration en hafnium commence à augmenter du $HfAlO$ -1 : 9 au $HfAlO$ -1 : 2, la concentration en N augmente aussi, jusqu'à 15% pour le $HfAlO$ -1 : 2. Cela suggère que l'hafnium joue un grand rôle dans l'incorporation de l'azote.
- Quand la concentration continue à augmenter ($HfAlO$ -9 : 1 et HfO_2), l'azote n'est plus présent dans le volume. Il est totalement à l'interface avec une importante concentration de 20-25%.

Des mesures AR-XPS (Angle Resolved X-ray Photo-electron Spectroscopy) ont aussi été réalisées [Gra07] pour étudier la nature des liaisons : $Hf-N$ et $Al-N$. Elles confirment les résultats observés grâce à l'Auger, ainsi :

- Pour les *Hf*-riches, l'azote est uniquement présent à l'interface.
- Pour les *Al*-riches, l'intensité du pic $Hf-N$ augmente avec la profondeur et aucune

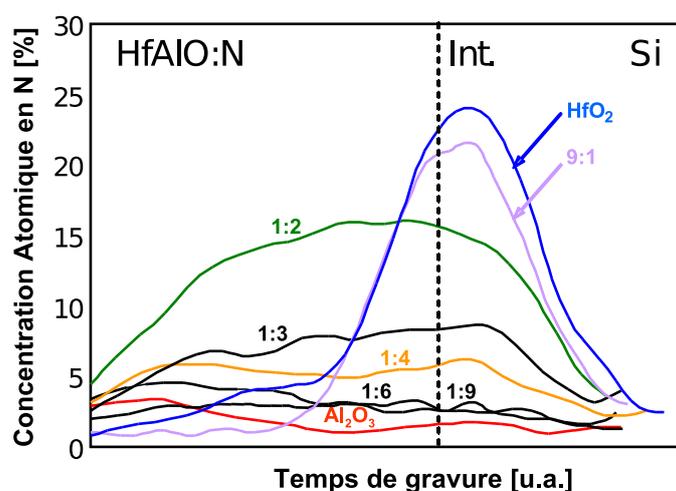


FIG. 3.30: Profil Auger de la concentration en nitrure dans un empilement bicouche : SiO_2 interfacial-0,6nm + *High-κ*-6nm (HfO_2 , Al_2O_3 , et $HfAlO$ avec différentes concentrations).

liaison $Al - N$ n'est détectée.

- Pour les $HfAlO$ -1:4, les analyses AR-XPS suggèrent une incorporation non uniforme de l'azote qui augmente avec la profondeur (comme observée sur les profils Auger figure 3.30). De plus, lorsque l'incorporation d'azote augmente, les liaisons $Hf - N$ dominent très légèrement sur les liaisons $Al - N$ mais les deux sont présentes.



En conclusion, pour les Hf -riches, les résultats Auger montrent une incorporation de l'azote à l'interface $High-κ/Si$ et les mesures AR-XPS mettent en évidence que l'azote est relié à l'hafnium. Pour les Al -riches, l'Auger a montré une très faible incorporation de l'azote dans le volume et les mesures AR-XPS ont montré seulement la présence de liaisons $Hf - N$, suggérant que les liaisons avec l'hafnium sont plus favorables que celles avec l'alumine dans ce cas. Au contraire, pour un $HfAlO$ intermédiaire, on observe les deux types de liaison. On peut noter qu'il semble que l'incorporation d'azote ne se fasse plus dans le volume du $HfAlO$ lorsque il est sous une forme polycristalline (cf. figure 3.4).

3.3.1.2 Influence sur le gap optique

Le gap optique a été extrait pour les différents échantillons par ellipsométrie (figure 3.31). Comme nous l'avons déjà vu sur la figure 3.3, le gap des $HfAlO$ diminue avec l'augmentation de la concentration en Hf . On peut noter que les recuits haute température (simulant les recuits d'activation source/drain), modifient peu le gap de la couche non nitrurée. Le recuit nitrurant après dépôt impacte plus lourdement le matériau :

- Pour l'alumine, son gap optique reste presque inchangé après nitruration.
- Lorsque la concentration en hafnium augmente, le gap diminue après nitruration. La diminution est d'autant plus forte que la concentration en Hf est grande. Cela peut

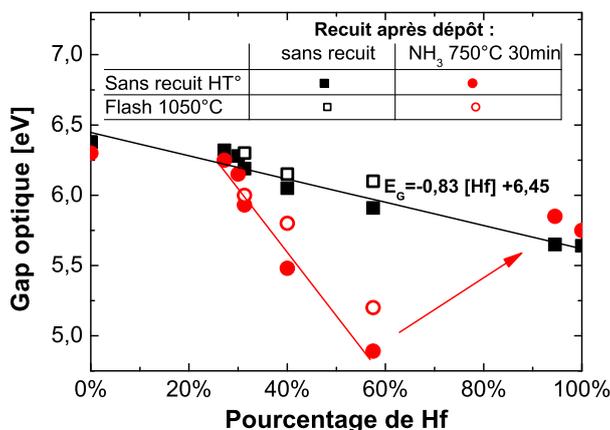


FIG. 3.31: Gap optique des $HfAlO$, impact de la concentration en Hf , de la nitruration et des recuits haute température.

être interprété comme une modification intrinsèque due à l'incorporation d'azote, hypothèse en accord avec les mesures Auger de la figure 3.30.

- Aussi, pour les fortes concentrations en *Hf*, où l'incorporation de l'azote ne se fait plus dans le volume, la nitruration ne modifie plus le gap optique.



La nitruration réduit le gap d'autant plus que l'on incorpore de l'azote dans le volume du *High-κ*. Au-delà d'une concentration limite en *Hf* (pour les *Hf*-rich), le gap ne varie plus après nitruration. Cela peut être mis en relation avec la faible incorporation d'azote dans le volume lorsque le matériau est cristallin (N principalement à l'interface).

3.3.1.3 Cristallinité

La cristallinité a aussi été étudiée. En effet, la nitruration augmente la température de cristallisation, comme le *Al* dans les *HfAlO* (cf. [Ho 02, Wie03, Zhu01] et figure 3.29). L'objectif est de trouver le lien entre incorporation d'azote et cristallinité des couches. Aussi, des mesures ATR-FTIR (Attenuated Total Reflection Fourier Transform InfraRed spectroscopy) (figure 3.32) ont été effectuées sur nos échantillons morphologiques avec des longueurs d'ondes de 600-1300 cm^{-1} . De plus, des photographies MET (figure 3.33) ont été réalisées sur des empilements à base de *HfAlO*.

Pour les échantillons sans recuit après dépôt, on observe un décalage de la raie *Al – O* (950 cm^{-1}) vers la raie *Hf – O* quand la concentration en *Hf* augmente (figure 3.32-a). Pour les échantillons nitrurés, on observe un décalage significatif de la raie *Al – O* vers les longueurs d'ondes plus basses dû à l'incorporation de l'azote. En effet, l'environnement n'est pas le même, donc les liaisons ne répondent pas aux mêmes fréquences : les liaisons *Al – O – N* ont été détectées à $920\text{-}940 \text{ cm}^{-1}$.

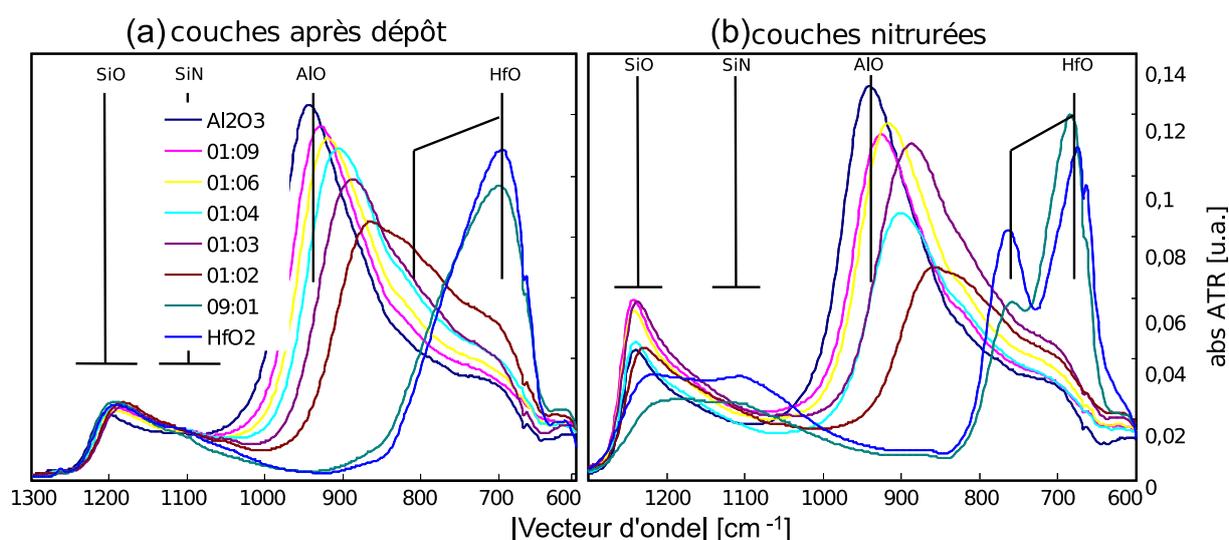
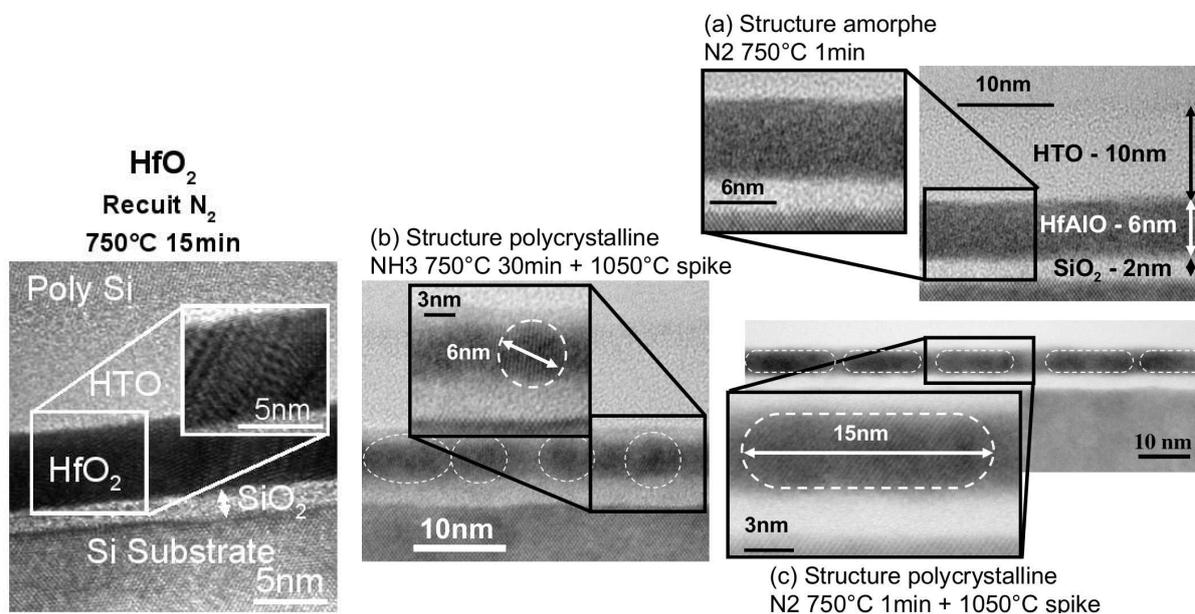


FIG. 3.32: Spectres ATR des *HfAlO* (a) après dépôt et (b) après nitruration.

Les *HfAlO* riches en *Al* ne montrent pas de cristallisation avant ou après nitruration. Alors qu'une cristallisation est observée pour le *HfAlO*-9:1 et le *HfO*₂ après la nitruration par la double bosse caractéristique. Néanmoins, il est impossible de conclure sur l'état cristallin du *HfAlO*-9:1 et *HfO*₂ après dépôt et donc sur l'impact de la nitruration sur la cristallisation seulement avec le spectre ATR-FTIR. On peut noter que la nitruration des *Hf*-riches génère l'apparition de liaisons SiN qui traduisent la nitruration du *SiO*₂ interfacial. Des analyses ATR-FTIR complémentaires (non montrées ici) ont été effectuées, le *HfO*₂ est apparu cristallin après le recuit à 750°C que ce soit sous atmosphère *NH*₃ ou *N*₂, les raies *Hf* – *O* étant moins intenses dans le cas du *HfO*₂ nitruré [Roc03, Gra07]. L'incorporation de N semble limiter la cristallisation [Yeo04].

On peut remarquer que l'analyse de l'interface montre des liaisons *Si* – *O* stables après le dépôt *High-κ*. La nitruration augmente l'intensité de ces liaisons pouvant s'expliquer par la transformation de l'oxyde chimique (*SiO*₂ interfacial) en oxyde thermique idéal.

Des photographies MET ont été effectuées sur des empilements tricouches : *SiO*₂-2nm/*High-κ*-6nm/HTO-10nm. Une première série de photos MET a été réalisée pour les empilements intégrant du *HfO*₂. Ces photos ont été effectuées sur les empilements ayant vu plusieurs types de recuits (nitrurant, non-nitrurant, recuit de simulation d'activation des sources/drains, etc). La figure 3.33a montre que le *HfO*₂ est polycristallin après le recuit *N*₂ - 750°C. Il en est de même pour tous les autres recuits (non montrés ici), en effet



(a) Photographies MET de la section d'un empilement incluant du *HfO*₂.

(b) Photographies MET de la section d'un empilement incluant de *HfAlO*-1:2 avec différents recuits : (a) recuit après dépôt *N*₂ à 750°C, (b) recuit après dépôt *NH*₃ 750°C avec un recuit final à 1050°C spike, (c) recuit après dépôt *N*₂ à 750°C avec un recuit final à 1050°C spike.

FIG. 3.33: Photographies MET d'empilements à base d'*HfO*₂ et d'*HfAlO*-1:2 avec différents recuits.

leur budget thermique étant supérieur à celui du recuit N_2 , la cristallinité du HfO_2 ne peut qu'augmenter. Une deuxième série de photos MET a été faite pour les empilements intégrant du $HfAlO$ -1:2. La figure 3.33b montre que le $HfAlO$ reste amorphe après le recuit à 750°C sous N_2 (figure 3.33b-a). Après le recuit haute température, l'échantillon commence à cristalliser, il se forme des cristaux de 15nm de diamètre (figure 3.33b-b). Mais le recuit nitrurant combiné au recuit haute température présente des grains cristallins de diamètre 6nm dans une matrice amorphe. Ces observations suggèrent que l'incorporation d'azote diminue la cristallinité des $HfAlO$, comme l'ont suggéré les analyses ATR-FTIR.

En conclusion de cette analyse morphologique, on observe que l'incorporation volumique de l'azote augmente avec la concentration en hafnium pour les $HfAlO$, d'où la faible incorporation dans l'alumine. Cette incorporation augmente la température de cristallisation, réduisant la taille des grains cristallins après les hautes températures. Néanmoins, passé une certaine concentration de Hf , le $HfAlO$ devient cristallin et l'azote ne s'incorpore plus dans le volume, on observe alors une nitruration de l'interface.

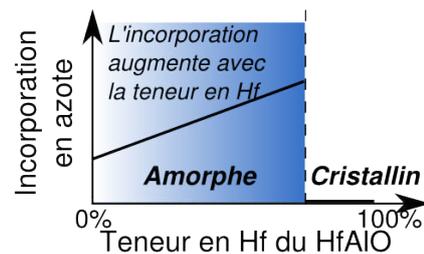


FIG. 3.34: Représentation de l'incorporation en azote en fonction de la concentration en Hf .

3.3.2 Effets du recuit NH_3 sur les propriétés de piégeage de l' HfO_2 et de l' $HfAlO$ -1:2

Dans cette partie, nous nous intéressons à l'impact de la nitruration sur les propriétés de piégeage du HfO_2 et du $HfAlO$ -1:2. L'empilement étudié est un tricouche SiO_2 -2nm/ $High-\kappa$ -6nm/HTO-10nm. Pour quantifier le piégeage, nous nous intéressons au décalage de la tension de bande plate (V_{FB}), extrait à partir des $C-V$ (quasi-statique), en fonction du temps de stress lorsque l'on applique une forte tension sur la grille. La figure 3.36 nous montre le décalage du V_{FB} indiquant un piégeage d'électrons dans la couche de $High-\kappa$ avec le temps de stress. La cinétique de piégeage est aussi grandement modifiée par les différents recuits.

Pour le HfO_2 , on observe que le recuit NH_3 réduit la cinétique de piégeage alors que les recuits haute température l'augmentent. Les études morphologiques ont montré que la nitruration n'incorpore pas de l'azote dans le volume du HfO_2 mais seulement à l'interface. Il apparaît donc, que l'optimisation de l'interface entre oxyde *tunnel* et $High-\kappa$ est un facteur important pour le contrôle du piégeage.

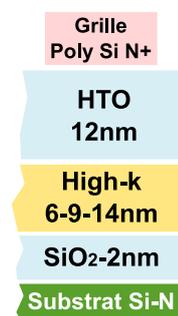


FIG. 3.35: Structure d'étude du piégeage.

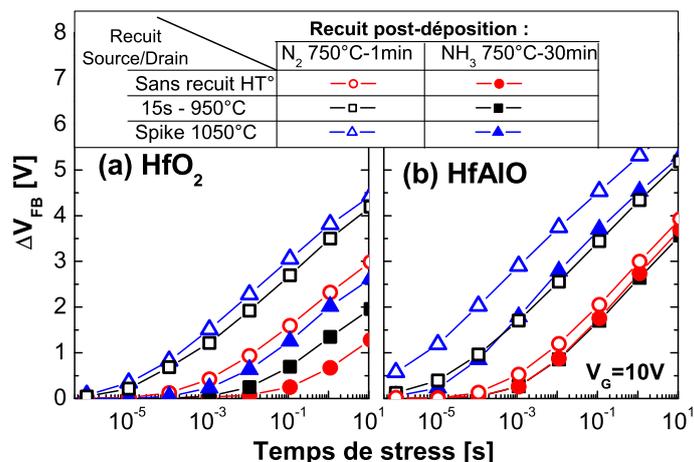


FIG. 3.36: Caractéristiques de programmation $\Delta V_{FB} - Temps_{Stress}$ d'empilements a) HTO-4nm / HfO_2 -6nm / HTO-4nm ou b) HTO-4nm / $HfAlO$ -1:2-6nm / HTO-4nm ; ayant subi ou non une nitruration (recuit 750°C-30min sous NH_3) et une simulation de recuit source/drain (1050°C-spike ou 950°C-15s).

Pour le $HfAlO$ -1:2, on observe aussi une importante augmentation de la cinétique de piégeage pour les dispositifs ayant subi les recuits haute température. De même, le recuit nitrurant protège de l'augmentation du piégeage dû au recuit d'activation. Il est possible de mettre en relation les propriétés de piégeage avec la structure cristalline des matériaux. En effet, les photos MET de la figure 3.33 nous ont montré que le $HfAlO$ -1:2 était amorphe après recuit N_2 puis devient polycristallin après le recuit d'ac-

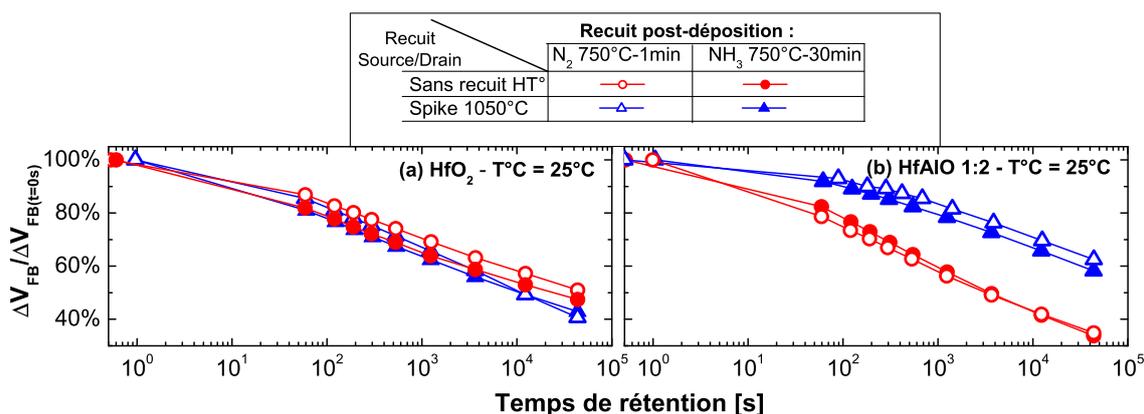


FIG. 3.37: Caractéristiques de rétention à 25 °C (Pourcentage de charges perdues en fonction du temps de rétention) d'empilements HTO-4nm / *High- κ* -6nm / HTO-4nm (soit HfO_2 , soit $HfAlO$) ayant subi ou non une nitruration (recuit 750°C-30min sous NH_3) et une simulation de recuit source/drain (1050°C-spike ou 950°C-15s). $\Delta V_{FB(t=0s)} = 2V$.

tivation spike. Il est possible que la forme cristalline fasse apparaître des défauts dus à l'organisation structurée et une stochiométrie imparfaite. Cela expliquerait l'augmentation des propriétés de piégeage avec les recuits d'activation. L'impact de la nitruration reste incertain :

- Soit l'incorporation d'azote permet de passiver ces lacunes dans le réseau cristallin du $High-\kappa$.
- Soit la nitruration augmentant avec la température de cristallisation, c'est la diminution de la taille des grains dans la matrice amorphe du matériau (figure 3.33) qui explique la diminution du nombre de pièges.

Il est fort probable que les deux mécanismes interviennent, expliquant l'augmentation du piégeage avec les recuits haute température et la diminution du piégeage à la fois pour le HfO_2 et le $HfAlO$ -1:2.

Néanmoins, il faut vérifier que les cinétiques de piégeage ne sont pas dûes à la nitruration ou à la modification de l'oxyde *tunnel* avec les recuits.

La figure 3.37 nous présente la rétention des électrons piégés dans nos empilements. Pour le HfO_2 , on remarque que les différents traitements thermiques ne modifient pas la cinétique de perte de charges. Cela nous indique qu'il n'y a pas eu de modification de l'oxyde *tunnel* après les recuits haute température. En effet, une recroissance, même faible de l'oxyde *tunnel*, réduit de façon importante la perte de charges dans le $High-\kappa$. De plus, cette même recroissance réduirait la dynamique de piégeage. Aussi, les propriétés de piégeage ne sont pas dues à la variation de l'épaisseur de l'oxyde *tunnel*, mais bien à un changement des propriétés de capture du matériau $High-\kappa$ lui-même. Enfin, les photographies MET de la figure 3.33 ne nous avaient montré aucune recroissance de l'oxyde *tunnel* entre les différents recuits.

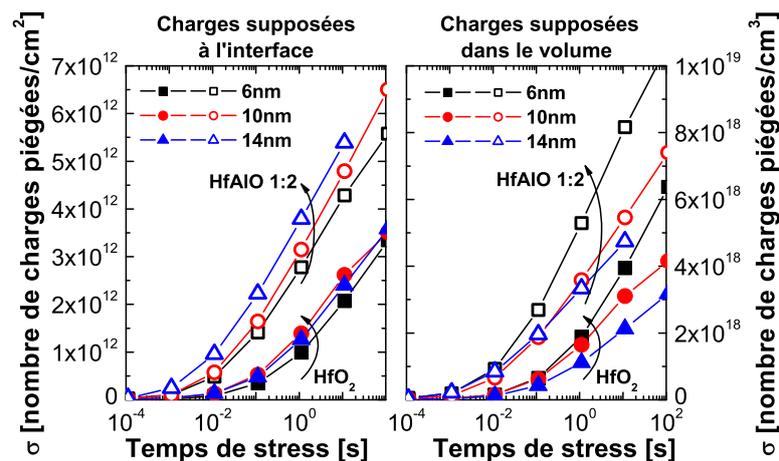


FIG. 3.38: Densité de charges piégées pour deux modes de normalisation : a) la charge est localisée à l'interface SiO_2 - $High-\kappa$; b) la charge est localisée dans le volume du $High-\kappa$. Empilement considéré HTO-4nm / $High-\kappa$ / HTO-4nm avec un recuit après dépôt NH_3 750°C et un recuit final à 950°C 15s. $V_G/EOT = 6MV/cm$.

Avec le $HfAlO$ -1:2 la tendance est même inversée : ce sont les dispositifs qui piègent le plus qui ont la meilleure rétention. Il est donc impossible d'attribuer l'augmentation du piégeage à une diminution de l'épaisseur de l'oxyde *tunnel*.

Afin de localiser physiquement les pièges dans l'empilement, nous avons étudié des capacités de différentes épaisseurs de *High-κ*. Sur la figure 3.38, deux normalisations ont été effectuées pour extraire la charge piégée pour les différentes épaisseurs, correspondant aux deux localisations possibles :

- Des charges surfaciques, localisées à l'interface oxyde *tunnel* et *High-κ*.
- Des charges volumiques uniformément réparties dans le volume du *High-κ*.

La figure 3.38 nous montre une excellente normalisation surfacique pour les deux *High-κ* alors que la normalisation volumique reste médiocre. On peut donc en conclure, au premier ordre, que la charge est localisée à l'interface $SiO_2/High-κ$ ou dans les premières couches du volume du matériau.



En conclusion, pour le HfO_2 , l'azote incorporé à l'interface par nitruration réduit le piégeage d'électrons. Pour les $HfAlO$, la nitruration réduit le piégeage en augmentant la température de cristallisation et en limitant le taux de cristallinité de la couche de *High-κ*. De plus, il apparaît que la charge piégée est principalement localisée à l'interface $SiO_2/High-κ$.

3.4 Impact du traitement thermique sur l'alumine

L'alumine est l'un des matériaux *High-κ* les plus attractifs pour les applications mémoires. Son grand gap, son faible courant de fuite à faible champ et sa constante diélectrique ($\epsilon_r \approx 9$) lui permettent d'assurer un couplage suffisant, tout en assurant son rôle d'isolant, notamment en rétention. Aussi, il fait l'objet d'un fort engouement de la part des industriels à la fois pour les applications à grille flottante et à couche de piégeage, la plus célèbre étant la TANOS de Samsung [Lee03, Par06, Lee06].

L'objectif de cette partie est de donner quelques éléments de compréhension sur ce matériau, notamment l'impact de la cristallinité sur les propriétés physico-chimiques et électriques de l'alumine. Aussi nous présenterons une partie des résultats de la caractérisation physico-chimique ainsi que des résultats de simulations atomistiques réalisées au Léti. Ces résultats permettront de mieux comprendre l'impact des défauts ou lacunes sur les résultats électriques et ainsi améliorer les caractéristiques mémoires telles que la programmation et la rétention.

3.4.1 Les propriétés physico-chimiques

Dans cette partie nous allons étudier les propriétés physico-chimiques de l'alumine, tels que le stress en tension, la densité, la structure cristalline et la teneur en hydrogène, en fonction de différents recuits après dépôt. Les échantillons étudiés sont des dépôts pleine plaque d'alumine 15nm déposés par ALCVD à 350 ° C (cf section 3.2.1.1). Les différents types de recuits après dépôt réalisés ont des températures comprises entre 700 ° C et 1050 ° C. De plus, deux types d'atmosphère ont été étudiés : O_2 ou N_2 .



3.4.1.1 Étude du stress, de la densité et de l'épaisseur

Les mesures de stress¹², d'épaisseur et de densité¹³ montrent clairement deux types de population (figure 3.39), liés au budget thermique du recuit :

- Faible budget thermique (recuits 700 ° C), alumine amorphe : épaisseur nominale de 14-15nm et densité faible de $3g/cm^3$.

Pour le stress, on note une influence de l'atmosphère O_2 ou N_2 sur le recuit 700 ° C. En effet, le recuit sous N_2 correspond à une alumine amorphe [Kra05]¹⁴ alors que l'alumine ayant vu le recuit sous O_2 présente une valeur de stress typique d'une couche cristallisée [Kra05]. On peut donc penser que l'alumine recuit à 700 ° C sous N_2 est amorphe alors que le recuit sous O_2 a généré des amorces de cristallisation, augmentant ainsi le stress.

- Fort budget thermique, alumine cristalline : diminution de l'épaisseur à $\sim 12,5nm$, augmentation de la densité à $3,5g/cm^3$ et augmentation rapide du stress lors de la cristallisation puis diminution progressive avec les recuits. On peut attribuer la diminution du stress avec l'augmentation de la température de recuit par des phénomènes de relaxation, que l'augmentation de la température favorise.

Le recuit spike à 1050 ° C constitue un cas particulier. On pense que certains phénomènes de relaxation, diffusion, réarrangement cristallographique n'ont pas le temps de se produire pour ce type de recuit rapide.

¹²Le stress en tension mécanique est obtenu par une mesure de flèche en utilisant la formule de Stoney.

¹³Les mesures d'épaisseur et de densité ont été réalisées en XRR (réflectométrie X équipement Jordan Valley JVX).

¹⁴On notera que l'on ne retrouve pas exactement les mêmes résultats que [Kra05]. Notre alumine correspondrait plutôt à une alumine épaisse d'après ces mesures. En effet, [Kra05] mesure le stress à une température donnée lors de cycles chauffant/refroidissant, et trouve (pour les faibles épaisseurs) que le stress augmente avec l'augmentation de la température et décroît presque de la même valeur durant le refroidissement : il observe un processus réversible. Dans notre cas, nous mesurons le stress après le recuit, à température ambiante c'est pourquoi nous ne retrouvons pas le caractère réversible.

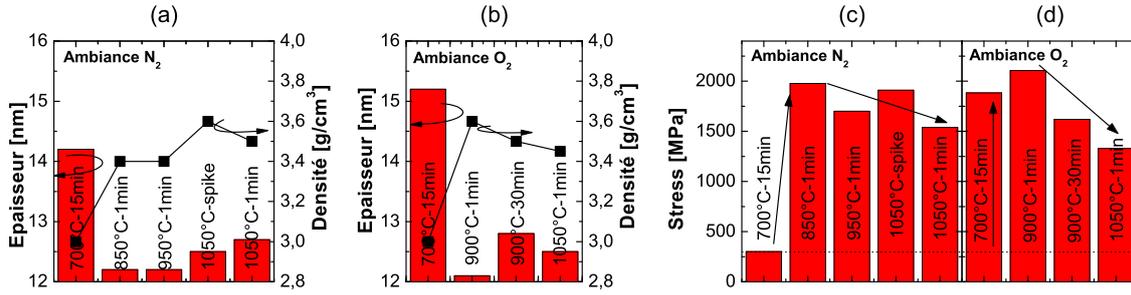


FIG. 3.39: (a-b) Densité et épaisseur physique de l' Al_2O_3 en fonction du recuit post-déposition. (c-d) Stress en tension dans l' Al_2O_3 en fonction du recuit post-déposition.

3.4.1.2 Mesures FTIR-ATR

La figure 3.40a montre l'absorption en FTIR-ATR de l'alumine 15nm suivant les différents recuits. Le premier pic situé vers $900 - 950cm^{-1}$ correspond à la liaison $Al - O$. Le second pic situé vers $1200 - 1250cm^{-1}$ correspond à la liaison $Si - O$ ¹⁵.

Concernant le pic Al-O, on peut extraire deux informations : la position du pic et sa largeur à mi-hauteur (appelée Full Width Half Maximum FWHM). En effet, la largeur à mi-hauteur est liée à l'ordre dans la couche : ainsi plus les liaisons sont ordonnées et plus la largeur à mi-hauteur du pic sera petite. Il est aussi possible d'aller plus loin dans l'exploitation des mesures ATR. En effet, dans [Lef07], il distingue deux phases cristallines de l'alumine grâce à la position du pic $Al - O$:

- la phase α se situe vers $920cm^{-1}$.
- la phase γ vers $930cm^{-1}$ et présente un épaulement vers $780cm^{-1}$ ¹⁶.

Cette analyse met en évidence les deux populations liées au budget thermique du recuit correspondant à l'état amorphe ou cristallin :

l'alumine amorphe (correspondant aux recuits à $700^\circ C$), possédant :

- une largeur à mi-hauteur du pic $Al - O$ important, indiquant un fort désordre dans la couche.
- un pic $Si - O$ faible : très faible recroissance de l'interface $Si - SiO_2$.

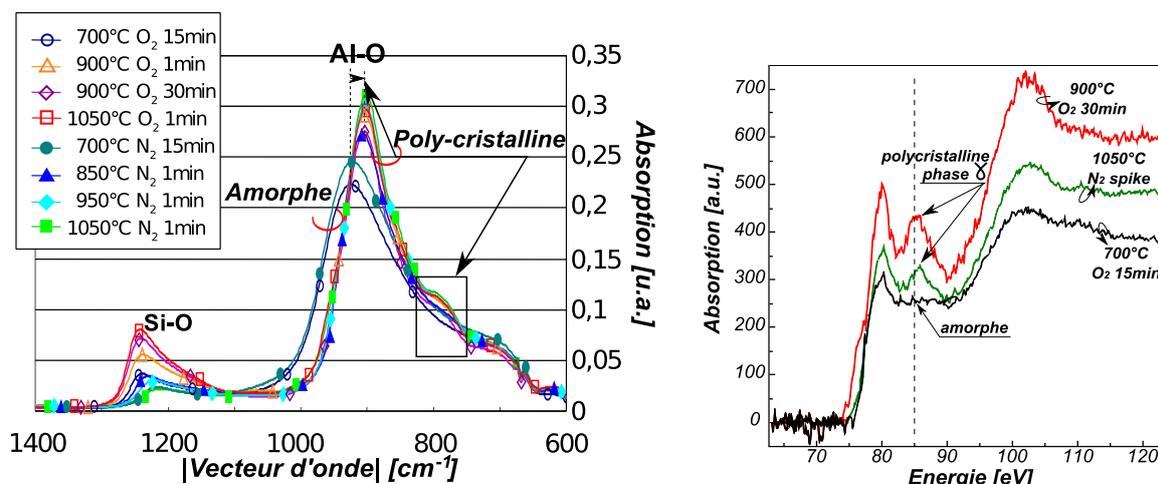
l'alumine cristallisée (correspondant aux recuits de 850 à $1050^\circ C$), possédant :

- une largeur à mi-hauteur du pic $Al - O$ fine, caractéristique d'une augmentation de l'ordre dans la couche.
- la présence du deuxième mode de la phase γ vers $780cm^{-1}$ ¹⁷.
- l'augmentation du pic $Si - O$ due à la recroissance de la silice interfaciale.

¹⁵l'aire du pic est proportionnelle au nombre de liaisons et donc proportionnelle à l'épaisseur de la couche.

¹⁶Ainsi en regardant la dérivée seconde des spectres FTIR-ATR il est possible de mettre en évidence le deuxième mode de la phase γ vers $780cm^{-1}$.

¹⁷via l'analyse de la dérivée seconde.



(a) Mesures FTIR-ATR, nature cristalline de l' Al_2O_3 15nm.

(b) Mesures EELS, phase cristalline de l' Al_2O_3 .

FIG. 3.40: Identification de la nature et de la phase cristalline de l' Al_2O_3 en fonction du recuit post-déposition.

Pour vérifier la cristallinité de nos échantillons et confirmer les phases identifiées en FTIR-ATR, nous avons réalisé sur trois échantillons une analyse EELS (figure 3.40b). Pour l'alumine recuite à 900 °C et à 1050 °C, on observe la présence d'un pic à 85eV, caractéristique de la phase γ ou phase θ (appelée gamma désordonnée), ce qui confirme nos résultats FTIR-ATR.

Enfin des mesures ellipsométriques ont permis d'extraire le gap de l'alumine en fonction de son recuit. Le tableau 3.5 regroupe le gap optique de ces différents échantillons. On observe une augmentation de 0,5eV du gap pour l'alumine cristalline. Cela peut être mis en relation avec la densification de l'alumine. En effet, les modèles 1D de calcul du gap¹⁸ confirment qu'une diminution de la maille cristalline et une densification de la couche, entraînent une augmentation du gap.

	Gap
Après dépôt	6,5eV
PDA 700 ° C	6,5eV
PDA 1050 ° C	7eV

TAB. 3.5: Gap de l'alumine en fonction du budget thermique.

¹⁸Par exemple des modèles de Kronig-Penney.

3.4.1.3 Concentration d'hydrogène dans l'empilement mémoire

Une analyse SIMS (Secondary Ion Mass Spectrometry)¹⁹ a été effectuée sur quatre empilements mémoires TANOS (Al_2O_3 -15nm/ Si_3N_4 -5nm/ SiO_2 -4nm/ Si massif) après avoir subi les quatre recuits suivants :

- 700 °C O_2 15min.
- 900 °C O_2 30min.
- 1050 °C N_2 spike.
- 1050 °C N_2 1min.

L'objectif est d'étudier la concentration en hydrogène dans l'empilement en fonction des différents recuits post-déposition. La figure 3.42 représente une mesure SIMS permettant de quantifier la concentration en hydrogène dans l'alumine. On observe que l'augmentation du budget thermique réduit la concentration en hydrogène dans l'alumine, due à la désorption de l' H_2O , mais aussi dans le nitrure.

La figure 3.41 représente les mesures MIR réalisées sur de l'alumine ayant subi différents recuits. L'étuvage²⁰ des échantillons permet de réduire l'eau²¹ présente sur la surface des plaques²². Les liaisons $O-H$ présentent un pic d'absorption à $3500cm^{-1}$ [Tsy90].

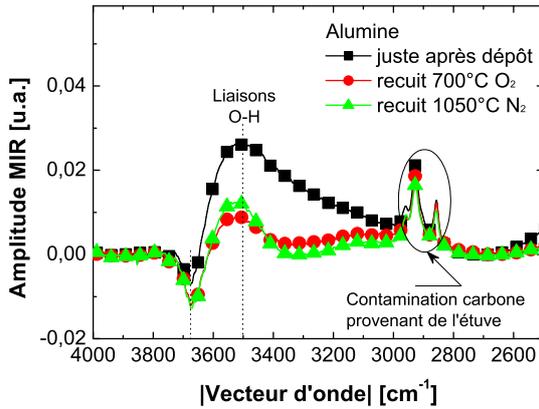


FIG. 3.41: Mesures MIR réalisées sur différents échantillons Al_2O_3 déposés sur Si, permettant de quantifier les liaisons $O-H$. Plusieurs recuits ont été effectués. Les échantillons ont été étuvés avant mesure pour éviter la présence d'eau sur les plaques.

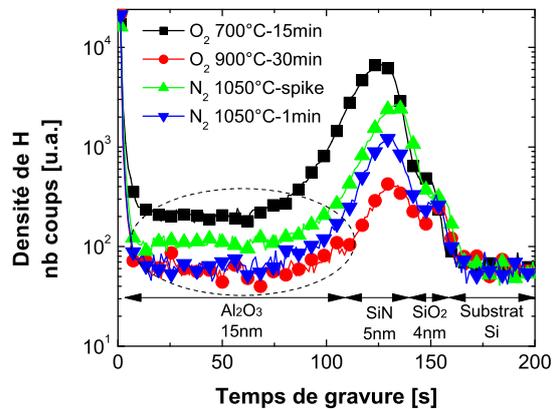


FIG. 3.42: Mesures SIMS de la densité d'hydrogène dans un empilement Al_2O_3 -16nm/ Si_3N_4 -6nm/ SiO_2 -4nm/ Si en fonction du recuit post-déposition effectué sur Al_2O_3 .

¹⁹Méthode de mesure qui consiste à bombarder la surface de l'échantillon à analyser avec un faisceau d'ions, pulvérisant la surface de l'échantillon. Une partie de la matière pulvérisée est ionisée et accélérée vers un spectromètre de masse permettant de mesurer la composition élémentaire, isotropique ou moléculaire de la surface de l'échantillon.

²⁰Étuvage : 48h à 100C.

²¹L'eau présente un pic d'absorption à $3200cm^{-1}$

²²On peut noter l'apparition de pics pour $3700cm^{-1}$. Cela est du à un pic présent sur la plaque de référence SiO_2 qui apparaît dans les valeurs négatives pour les échantillons d'alumine. Il est intéressant

Ainsi, on observe une forte concentration de liaisons $O - H$ pour une alumine obtenue juste après dépôt. Cette concentration diminue fortement après recuit. Néanmoins, nous n'observons pas de différence notable entre le recuit à $700\text{ }^\circ\text{C}$ et celui à $1050\text{ }^\circ\text{C}$.

Ainsi, on observe une diminution de la concentration en hydrogène pour les recuits à $1050\text{ }^\circ\text{C}$, ce qui n'est pas le cas des liaisons $O - H$. On peut donc en déduire que pour l'alumine recuit à $700\text{ }^\circ\text{C}$, l'hydrogène n'est pas présente que sous forme de $O - H$ et peut exister sous d'autres formes : en substitution ($Al - H$) ou en interstitiel.



En résumé de ces analyses physico-chimiques, l'alumine de 15nm déposée par ALD à $350\text{ }^\circ\text{C}$ est apparue cristalline pour les recuits à $900\text{ }^\circ\text{C}$ et à $1050\text{ }^\circ\text{C}$. De plus l'augmentation du budget thermique a montré une augmentation du gap, une densification du matériau et une diminution du stress en tension après cristallisation. Enfin, nous avons mis en évidence que les recuits haute température réduisaient la concentration en hydrogène dans l'alumine sans modifier le nombre de liaisons $O - H$.

3.4.1.4 Simulation atomistique

Nous venons de voir que les recuits réalisés juste après dépôt peuvent impacter la concentration en hydrogène ou en liaisons $O - H$ dans l'alumine. Or, leurs présences peuvent être la cause de défauts ou de pièges [Xio07] pouvant entraîner des courants de fuite et un piégeage parasite dans l'alumine. Aussi, avant d'évaluer l'impact des recuits sur les propriétés électriques de l'alumine, nous allons nous intéresser à l'impact de défauts sur des états électriquement actifs dans le gap de l'alumine. Pour cela, nous nous appuyerons sur des résultats de simulations atomistiques sur la phase γ de l'alumine, réalisées au CEA-Léti.

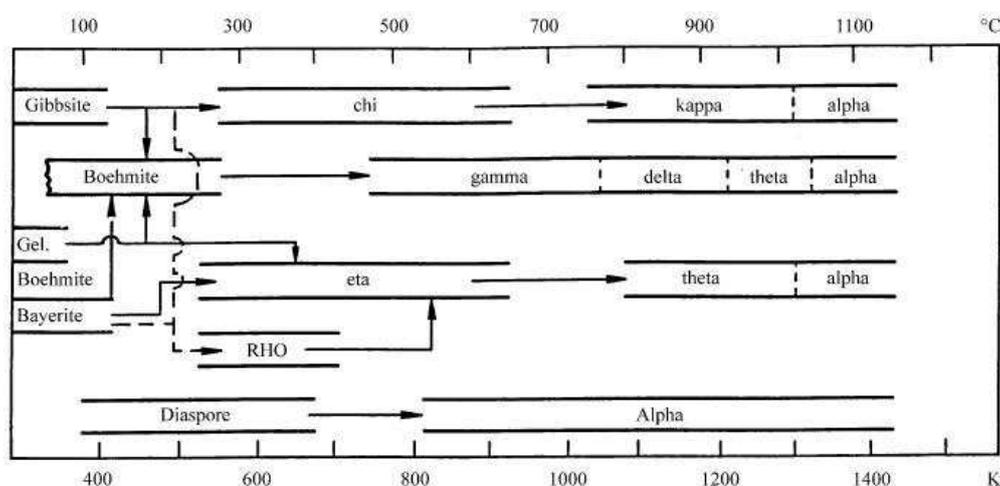


FIG. 3.43: Diagramme d'états de l'alumine [San09].

de constater que [Tsy90] associe cette pulsation au $O - H$ surfacique.

Les techniques de simulations atomistiques sont nombreuses et complexes, et sont l'objet de thèses entières. Ici, nous ne présenterons que des simulations en DFT²³. Elles permettent, à partir de la structure électronique de la matière (atome, structure, liaison...), d'obtenir la densité d'états du système. Ainsi la première chose à déterminer est la structure atomique du système, ici l'alumine.

L'alumine existe sous plusieurs formes. La figure 3.43 représente le diagramme de phases de l'alumine extrait de [San09]. On notera que la phase α , plus connue sous le nom de safire, est la forme limite de l'alumine. Elle n'est pas atteinte dans les conditions des procédés de la microélectronique. Les phases les plus probables, pour les températures de dépôt et de recuit, sont les phases κ et γ ²⁴ [Lef07]. En effet, l'alumine- α présente un gap de $\sim 9,2eV$ pour diminuer à $\sim 6,9eV$ en phase γ , des lacunes d'oxygène le réduisant à $\sim 6,3eV$ ce qui est comparable avec les mesures que nous avons obtenues (cf tableau 3.5). De plus, les mesures EELS, présentées dans la figure 3.40b, montrent la présence de la phase γ dans notre alumine après cristallisation. Ainsi, les simulations effectuées sont basées sur une alumine sous phase γ . L'étude s'est intéressée aux lacunes d'oxygène, à leur passivation avec l'hydrogène, et aux hydrogènes interstitiels.

La phase γ de l'alumine est donc une phase de transition métastable de l'alumine. La phase gamma est obtenue à partir d'alumine amorphe et reste présente jusqu'à des températures de recuit de 1200 ° C. Diverses structures sont proposées dans la littérature pour l'alumine γ , celle-ci ayant en effet un degré élevé de désordre et présente des schémas de diffraction pour le moins difficiles à interpréter²⁵.

La structure, qui a été choisie pour réaliser la phase γ , est de type spinelle modifiée. La structure utilisée s'appuie sur l' Al_2MgO_4 , les Mg occupent les sites à coordination tétraédrique²⁶, et les Al les sites à coordination octaédrique²⁷ dans un réseau cubique faces centrées d'oxygène. Ainsi, en spinelle, en remplaçant le Mg par l' Al , les atomes

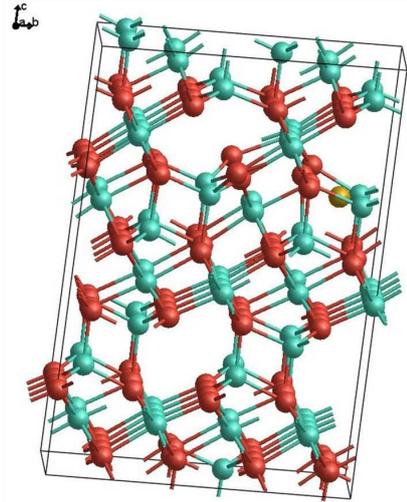


FIG. 3.44: Al_2O_3 en phase γ obtenue à partir d'une structure spinelle modifiée : supercellule de 160 atomes [Men05], avec l'oxygène en rouge, l'aluminium en turquoise, avec l'ajout d'un hydrogène en interstitiel en or.

²³DFT : *Density Functional Theory*.

²⁴On notera que la phase γ semble souvent de façon historique/confuse recouvrir les phases γ , δ , θ .

²⁵Néanmoins, des analyses XRD ont montré une bonne compatibilité entre une structure spinelle et la figure de diffraction pour l'alumine γ .

²⁶Sites à coordination tétraédrique : l'atome est au centre d'un tétraèdre formé par les atomes qui l'entourent, il a donc 4 liaisons.

²⁷Sites à coordination octaédrique : l'atome est au centre d'un octaèdre formé par les atomes qui l'entourent, il a donc 6 liaisons.

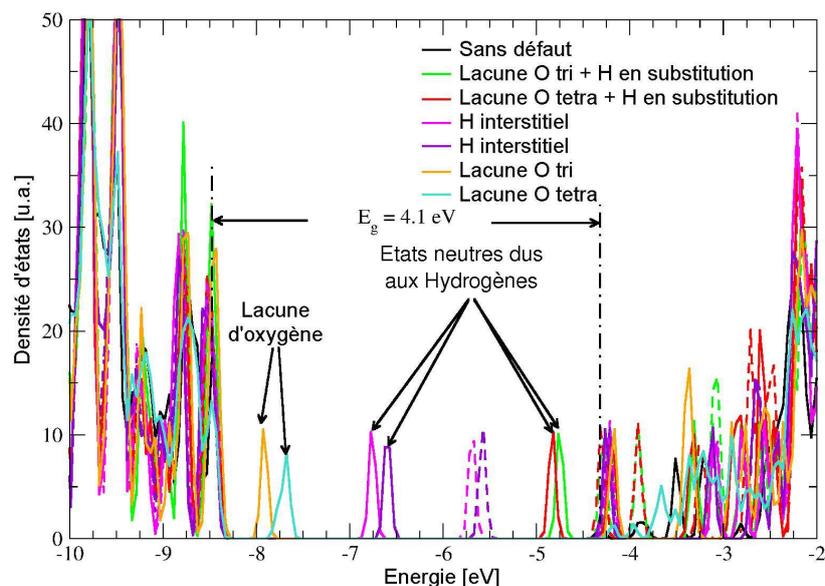


FIG. 3.45: Défauts actifs dans le gap d'une alumine sous phase γ . Les traits pleins représentent les états pleins et en pointillés ce sont les états disponibles. Les lacunes d'oxygène génèrent des états profonds à 1eV de la bande de valence. Les atomes d'hydrogène en interstitiel créent des états en milieu de gap. L'hydrogène en substitution ne passive pas les lacunes d'oxygène et crée deux pics près de la bande de conduction. Le gap de l'alumine sous phase γ est sous-estimé à 4,1eV (obtenu par DFT/LDA).

oxygènes sont trois ou quatre fois coordonnés (c'est-à-dire qu'ils sont reliés à trois ou quatre Al) et les atomes aluminiums sont quatre ou six fois coordonnés (c'est-à-dire qu'ils sont reliés à 4 ou 6 O). Néanmoins, on obtient de l' Al_3O_4 , ce qui n'est pas satisfaisant. Il ne faut donc remplacer qu'une fraction des Mg par des Al . Là est toute la difficulté, ce qui explique qu'il n'existe à ce jour que très peu de modèles de $Al_2O_3-\gamma$, type spinelle modifiée. Celui utilisé au CEA-Léti s'appuie sur le travail de [Men05].

La figure 3.44 correspond à la structure cristalline du $Al_2O_3-\gamma$ utilisé pour la simulation en DFT²⁸. La figure 3.45 représente la densité d'états obtenue par ces simulations pour différentes configurations :

- Sans défaut, en noir. On observe un gap de 4,1eV : il est important de souligner que la méthode de calcul utilisée, la DFT, sous-estime le gap. Des calculs en GW²⁹ [San09] donnent un gap à $\sim 6.9eV$ et des lacunes d'oxygène le réduisant à $\sim 6,3eV$ ³⁰
- Avec des lacunes d'oxygène, soit en site trivalent³¹ (en orange), soit en site tétravalent³² (en turquoise) : dans les deux cas, on note l'apparition d'un pic à $\sim 1eV$ de la bande de valence. On peut penser que ces états pleins, trop bas en énergie,

²⁸Avec utilisation du code Siesta [San97], les pseudopotentiels de Troullier-Martins et l'approximation LSDA.

²⁹Méthode permettant d'avoir une meilleure image du gap mais plus coûteuse en temps de calcul.

³⁰Cela est comparable avec les mesures ellipsométriques allant de 6,5eV à 7eV, cf tableau 3.5.

³¹Site trivalent : l'oxygène a trois Al comme premiers voisins.

³²Site tétravalent : l'oxygène a quatre Al comme premiers voisins.

ne peuvent pas participer au courant d'électrons *tunnel* mesuré dans les mémoires TANOS [Liu09].

- Avec des lacunes d'oxygène compensées par des hydrogènes en substitution, pour les sites trivalents (en vert) et pour les sites tétravalents (en rouge) : ces défauts créent des états pleins très proches de la bande de conduction et l'on peut noter que leurs états libres sont légèrement au dessus de la bande de conduction calculée. Or, comme le gap est sous-estimé en DFT, il est possible que ces états soient juste sous la bande de conduction de l'alumine. Ils pourraient, ainsi, participer à la conduction. Les hydrogènes, en substitution de lacunes d'oxygène, ne sont donc pas à écarter.
- Avec des atomes d'hydrogène en interstitiel (violet et magenta, correspondant à deux positions d'interstitiels stables) : on observe deux états dans le gap correspondant aux états pleins et aux états disponibles. Leurs positionnements, suffisamment proches de la bande de conduction ($\sim 0,5eV$ et $\sim 1eV$), laissent penser qu'ils pourraient fortement participer au piégeage dans l'alumine ou à des mécanismes de conduction de type TAT ou Poole-Frenkel en température.

On notera que les $O-H$ ne sont pas forcément stables et induisent des H interstitiels. Les $O-H$ et $O-$ restant stables, ne semblent pas induire d'états dans le gap. De plus des simulations effectuées avec des défauts d'hydrogène sans comportement interstitiel, donc uniquement un comportement $O-H$, ne génèrent pas de pic dans le gap de l'alumine.



En conclusion, on peut donc penser que ce ne sont pas les $O-H$ qui créent des pièges dans le gap de l'alumine, mais plutôt les hydrogène interstitiels. Néanmoins, on peut s'interroger sur la stabilité des hydrogènes en site interstitiel durant les recuits. De plus, l'alumine n'est pas purement cristalline, ainsi, dans une structure polycristalline ou amorphe, les joints de grains et les interfaces sont des zones privilégiées de défauts pour l'hydrogène, pouvant donner les mêmes pièges actifs que ceux que nous avons présentés.

3.4.2 Étude des caractéristiques électriques

3.4.2.1 Étude du courant de fuite

Nous avons étudié la conduction dans une alumine cristalline de $8nm$ d'épaisseur pour déceler d'éventuelles conductions assistées par pièges. Ainsi, la figure 3.46a représente la densité de courant à différentes températures. On observe une activation. Pour identifier un mode de conduction, nous avons tracé le diagramme de Hill pour ces différentes températures, cf figure 3.46b. On obtient une bonne normalisation pour une profondeur de pièges de $\sim 0,8eV$. Cette valeur est en accord avec les simulations atomistiques présentées précédemment. En effet, des défauts d'hydrogène en interstitiel ont créé des états dans le gap à $\sim 1eV$ de la bande de conduction.

Il apparaît donc que l'alumine cristalline aurait une conduction de type Pool-Frenkel et que les défauts participants pourraient être dus à des atomes d'hydrogène en interstitiel.

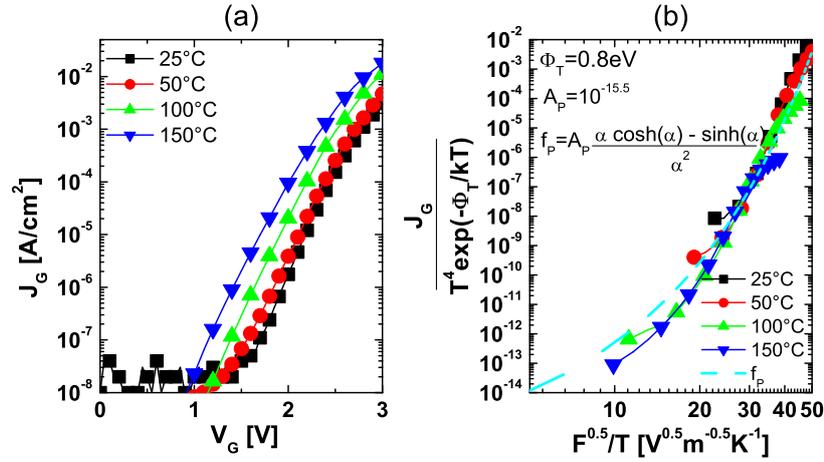


FIG. 3.46: (a) Caractéristiques $J_G - V_G$ pour différentes températures et (b) diagrammes de Hill pour une alumine cristalline de 8nm.

3.4.2.2 Étude des propriétés de piégeage

Nous avons vu précédemment que des atomes d'hydrogène pouvaient générer des états dans le gap de l'alumine. De plus, la concentration en hydrogène, mesurée par SIMS, diminue avec l'augmentation de la température du recuit après dépôt. Nous allons donc nous intéresser aux propriétés de piégeage de l'alumine pour voir s'il existe une corrélation entre piégeage, recuit et concentration en hydrogène.

Nous avons effectué des mesures de piégeage sur de l'alumine déposée sur une pleine plaque de silicium et ayant subi des recuits sous atmosphère N_2 juste après dépôt :

- Recuit 15min à 700 ° C.
- Recuit 1min à 850 ° C.
- Recuit 1min à 950 ° C.
- Recuit flash à 1050 ° C.
- Recuit 1min à 1050 ° C.

Ces mesures ont été réalisées grâce à une bille de mercure qui jouait le rôle de grille. Le piégeage a été quantifié en effectuant des mesures $C - V$. En réalisant des balayages successifs, en partant de tensions de plus en plus fortes, on observe un décalage de la tension de bande plate du dispositif, caractéristique d'un piégeage de charges négatives dans l'alumine (cf figure 3.47a). La figure 3.47b représente le décalage de la tension de bande plate en fonction de la tension maximum du balayage, normalisé par l'EOT de l'empilement (V_G/EOT).

On observe que le piégeage diminue avec l'augmentation de la température des recuits. On obtient ainsi une bonne corrélation entre les mesures de piégeage, la densité d'hydrogène mesurée dans l'alumine et les états générés dans le gap obtenus en simula-

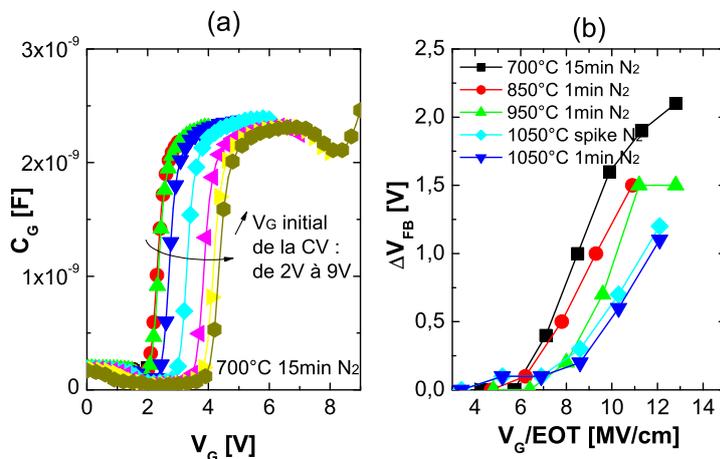


FIG. 3.47: (a) Caractéristiques $C_G - V_G$ permettant d'effectuer des mesures de piégeage de l'alumine recuite à 700°C 15 min sous N_2 . Caractéristiques de piégeage de couches d'alumine $\sim 15\text{nm}$ ayant subi différents recuits post-déposition (b) sous atmosphère N_2 .

tion atomistique. On peut donc penser que le piégeage dans l'alumine est principalement dû aux défauts d'hydrogène qu'il contient. Aussi un recuit à haute température permet de réduire significativement ce piégeage parasite.



En conclusion, il semble qu'une alumine cristalline soit préférable en terme de piégeage et de courant de fuite. En pratique, une fois intégrée dans une technologie mémoire, il est fort probable que l'alumine soit cristalline en fin de fabrication, notamment au fort recuit d'activation des dopants. Néanmoins, il est préférable de la stabiliser tout de suite après dépôt par un recuit haute température : soit 900°C 15min sous O_2 , soit 1050°C 1min sous N_2 .

3.5 Silicate d'hafnium : $HfSiON$

L'utilisation de *High- κ* en monocouche [Van05, Van06b] ou en tricouche³³ [Gov03, Lai07] permet d'augmenter la densité de courant d'injection à forte tension³⁴ sans dégradation théorique des propriétés d'isolation. Cependant cette intégration est limitée par le piégeage parasite dans la couche *High- κ* ainsi que les courants de fuite assistés par pièges qui entraînent une dégradation de la fiabilité du dispositif. Ainsi, l'ingénierie du matériau *High- κ* pour réduire la densité de pièges est nécessaire pour de telles applications. En particulier, il a été montré que le silicate d'hafnium nitruré ($HfSiON$) présentait des propriétés de piégeage réduites en comparaison à l'oxyde d'hafnium (HfO_2) ou les silicates d'hafnium ($HfSiO$) [Buc06] ainsi qu'un courant de fuite plus faible [Akb04]. Ainsi nous nous intéressons ici aux propriétés de conduction de couches à base de $HfSiON$ pour application *tunnel*.

L'objectif est donc d'obtenir sur ces empilements des caractéristiques $I_G - V_G$ présentant une sensibilité plus forte au champ électrique par rapport au SiO_2 . En d'autres termes, il faut que le courant de grille soit :

- élevé à fort champ, ce qui correspond à une forte injection de charges dans la couche de piégeage pendant le régime de programmation de la mémoire.
- réduit à faible champ, ce qui correspond à peu de pertes de charges pendant le régime de rétention de la mémoire.

3.5.1 Analyses matériaux

Le $HfSiON$ de cette étude est obtenu à partir d'un $HfSiO$ déposé par MOCVD, dont la concentration en Hf est de 60%. La nitruration est effectuée par un recuit NH_3 de 30min. Différentes températures de nitruration ont été étudiées : 700 °C, 750 °C et 800 °C.

Une étude poussée sur des couches de 2,5nm a été réalisée au Léti en vue de leur intégration dans les filières MOS, notamment un contrôle des épaisseurs déposées a été réalisé : mesures ellipsométriques et photographies TEM. Ainsi, les résultats matériaux présentés ici proviennent de cette étude. Pour les applications mémoires non-volatiles, une épaisseur de 5nm est préférable. Ainsi, nous avons réalisé la couche de $HfSiO$ en deux passes de 2,5nm avant nitruration. Les observations MET réalisées sur des empilements mémoires complets (figure 4.48) montrent une bonne cohérence des épaisseurs et aucune discontinuité de la couche de $HfSiON$ n'est observée.

Dans cette partie, nous présentons quelques données de la littérature sur le $HfSiO$ et le $HfSiON$, tout en les confrontant avec nos propres mesures matériaux.

La figure 3.48 présente le gap ainsi que les hauteurs de barrière du $HfSiO$ en fonction

³³Empilement de type VARIOT.

³⁴Pendant l'écriture et l'effacement de la mémoire.

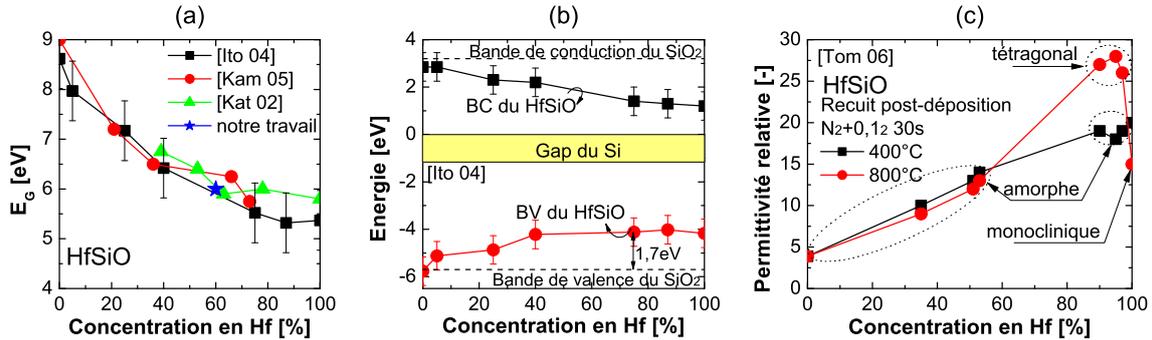


FIG. 3.48: (a) Gap optique, (b) niveaux de la bande de conduction (BC) et de valence (BV) (c) permittivité des HfSiO extraits de différentes publications [Kat02, Ito04, Kam05, Tom06] et des travaux réalisés au Léti.

de sa composition en Hf extraits de plusieurs publications [Kat02, Ito04, Kam05]. On observe une diminution du gap avec l'augmentation de la concentration en Hf partant de 9 eV (gap du SiO_2) jusqu'à 5,5 eV (gap du HfO_2). Dans notre cas, la composition est de 60% en Hf , ce qui correspond d'après ces références à un gap de 6 eV environ, cela est en parfait accord avec les mesures ellipsométriques (figure 3.50a) réalisées sur nos échantillons. On notera que le niveau de la bande de valence reste constant lorsque la concentration en Hf est supérieure à 40% (cf figure 3.48b).

La figure 3.48c, extraite de [Tom06], représente la permittivité en fonction de la composition en Hf . On observe que la permittivité dépend grandement de l'état cristallin du HfSiO . Les mesures de [Tom06] montrent une très nette augmentation de la permittivité pour une structure polycristalline (tétraogonale) comparée à une phase amorphe.

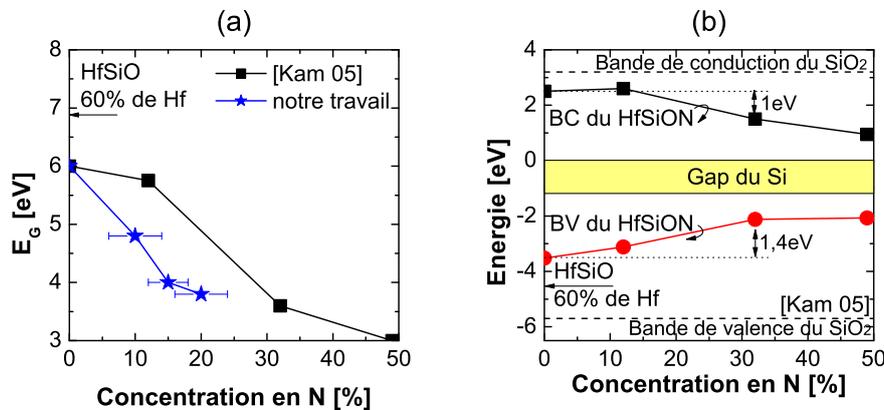


FIG. 3.49: (a) Gap optique et (b) niveau de la bande de conduction (BC) et de valence (BV) des HfSiON extraits de [Kam05] et des travaux réalisés au Léti.

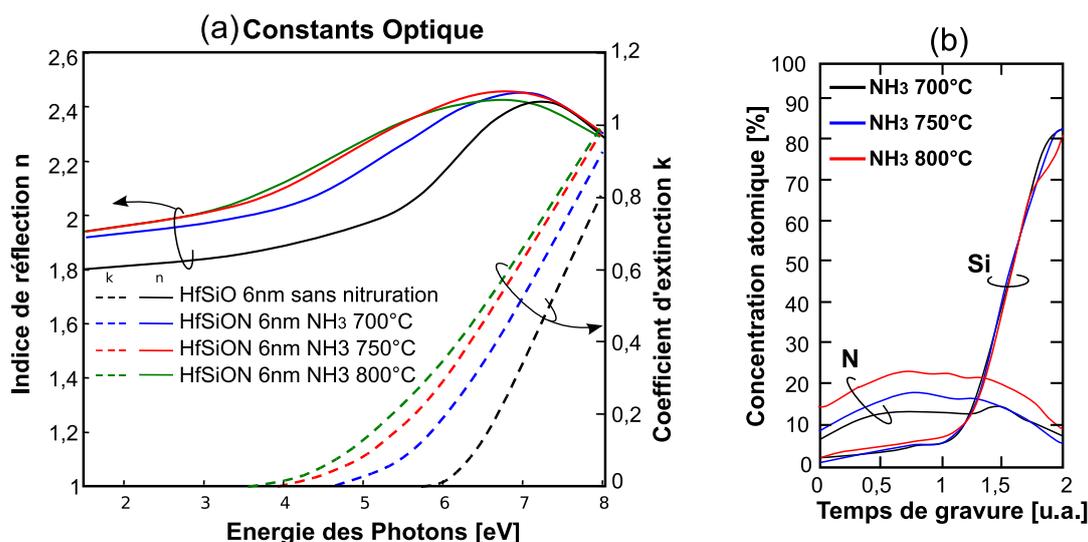


FIG. 3.50: (a) Mesures ellipsométriques du gap optique et (b) Mesures de la concentration atomique en azote par mesures SIMS pour différentes températures de nitruration.

Cette tendance est confirmée en simulation atomistique [Fis07]. L'objectif de la nitruration est d'augmenter la température de cristallinité du $HfSiO$ pour le maintenir amorphe [Vis02]. Aussi, nous retiendrons une permittivité de 15 pour le $HfSiO$ à 60% de Hf avant nitruration.

La figure 3.49 présente le gap et les hauteurs de barrière des $HfSiON$, en fonction de la quantité d'azote incorporée. On observe que le gap diminue jusqu'à 3eV pour une concentration de 50% d'azote. D'après [Kam05] cette diminution impacterait la hauteur de barrière côté bande de conduction et bande de valence. Cette forte réduction des hauteurs de barrière, malgré une diminution des pièges et de la conduction assistée par pièges [Akb04], peut être pénalisante à faible champ due à l'augmentation des courants *tunnel*.

Les silicates déposés au Léti sont nitrurés par un recuit NH_3 30min réalisé in-situ. La figure 3.50b regroupe les mesures SIMS des $HfSiON$ ayant subi des nitrurations à différentes températures. Il ressort que l'incorporation d'azote dépend beaucoup de la température. De plus, ce type de nitruration (nitruration four), du fait des températures élevées, entraîne une nitruration de l'interface $Si-SiO_2$ due à la diffusion des espèces d'azote.

3.5.2 Étude du courant de fuite

Pour comparer les courants de fuite, nous avons étudié des structures capacitives réalisées sur des substrats n³⁵ et p³⁶. Les deux types de structures ont été réalisés : les bicouches et les tricouches.

- Le SiO_2 a été obtenu par croissance, plusieurs épaisseurs ont été réalisées : 0,8nm, 1nm et 2nm.
- Le $HfSiON$ a été obtenu comme décrit dans la partie 3.5. Les mesures électriques d'épaisseurs³⁷ nous ont permis d'extraire une EOT pour le $HfSiON$ de 1nm environ, soit une constante diélectrique de 17,7 pour une épaisseur physique de 5nm.
- Pour les tricouches, du HTO a été déposé. La comparaison des EOT entre les plaques avec $HfSiON$ encapsulé et non encapsulé permet de remonter à une épaisseur de HTO de 1,6-1,7nm environ³⁸.
- Un recuit haute température (950 ° C-15s) a été réalisé afin de simuler le budget thermique³⁹ lors de la réalisation d'une cellule mémoire.

Dans ce paragraphe nous présentons les caractéristiques de courant de fuite de grille. Les mesures ont été effectuées sur des capacités de $100\mu m \times 100\mu m$ en tests systématiques. A noter que des tests manuels ont également été réalisés sur ces échantillons. Les résultats sont cohérents avec les tests systématiques. De plus, les courbes sont reproductibles dans la gamme de tensions balayées.

3.5.2.1 Structure bicouche $SiO_2/HfSiON$

La figure 3.51a présente les caractéristiques $J_G - V_G/EOT$ pour les empilements bicouches $SiO_2/HfSiON$ pour différentes températures de nitruration, avec une épaisseur de SiO_2 interfaciale de 2nm. On observe qu'une augmentation de la température de nitruration entraîne une augmentation du courant de fuite, plus particulièrement à $V_G < 0$. On peut attribuer ce comportement à la réduction du gap avec l'incorporation d'azote (cf figure 3.49), augmentant plus le courant d'électrons de la grille vers le substrat car en contact direct avec le $HfSiON$.

La figure 3.51b présente les caractéristiques $J_G - V_G/EOT$ pour les empilements bicouches $SiO_2/HfSiON$ avec une épaisseur de SiO_2 interfaciale de 0,8nm à 2nm. La température de nitruration est dans tous les cas de 800 ° C. À faible champ, on observe une diminution du courant de fuite lorsque l'épaisseur du SiO_2 interfacial augmente, simplement due à une diminution de la transparence avec l'augmentation de l'interface. À fort champ, une superposition des caractéristiques apparaît. Pour les tensions positives, elle

³⁵Pour les tensions positives : injection d'électrons du substrat vers la grille.

³⁶Pour les tensions négatives : injection d'électrons de la grille vers le substrat.

³⁷Mesures C(V) en tenant compte de la déplétion substrat (grille poly-N+ supposée métallique).

³⁸Ainsi, le retard à nucléation du HTO sur $HfSiON$ peut être estimé à 1,5nm environ. En effet, la valeur de HTO sur plaques témoins était de 3,2nm.

³⁹Recuit d'activation des jonctions.

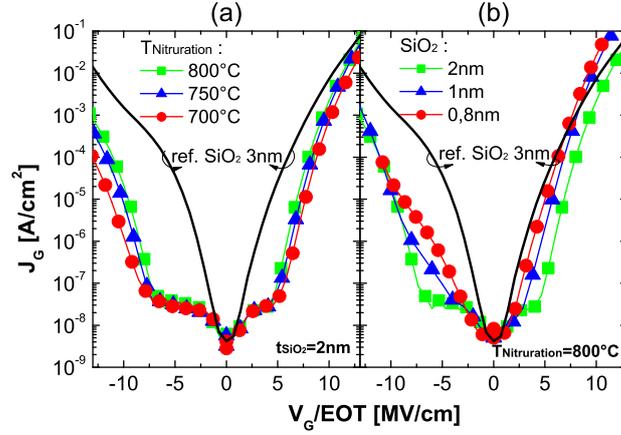


FIG. 3.51: Caractéristiques $J_G - V_G/EOT$ d'un bicouche $SiO_2/HfSiON$. (a) Impact de la température de nitruration. (b) Impact de l'épaisseur de SiO_2 .

n'apparaît pas, car la conduction Fowler-Nordheim intervient à plus de 20MV/cm pour le bicouche SiO_2 -2nm/ $HfSiON$ -5nm. On observe que la référence de SiO_2 -3nm possède un courant de fuite plus important à faible champ. À fort champ, les bicouches possédant des oxydes interfaciaux de 0,8nm et 1nm ont des courants injectants supérieurs à la référence. Aussi, en première approche, les bicouches $SiO_2/HfSiON$ présentent une meilleure sensibilité au champ électrique qu'un oxyde *tunnel* de SiO_2 de 3nm.

Pour comparer les performances de ces empilements une fois intégrés dans une mémoire non-volatile, les courbes sont normalisées de la manière suivante : on suppose que l'empilement est intégré dans une cellule mémoire à grille flottante continue avec un interpoly de 10nm. Ainsi, lorsqu'une tension V_{GC} est appliquée sur la grille de contrôle de la cellule mémoire, la différence de potentiel aux bornes de l'empilement *tunnel* est V_{GF} (cf équation 1.3). Ici, nous n'avons considéré aucune charge sur la grille flottante ($Q_{GF} = 0$)⁴⁰.

Ainsi, le courant tracé en figure 3.52a correspond au courant d'injection dans l'empilement *tunnel*, soit le débit de charges injectées dans la couche de piégeage si l'on néglige le courant de fuite à travers l'interpoly. Ce niveau de courant traduit donc les performances de l'empilement considéré, une fois intégré dans une cellule mémoire.

Il apparaît :

- À $V_G > 0$, les empilements à base de $HfSiON$ présentent une meilleure sensibilité au champ électrique, par rapport à un monocouche de SiO_2 . En effet, par rapport à 3nm de SiO_2 , le courant à faible champ⁴¹ est plus faible, alors que le courant à fort champ⁴² est plus important.

⁴⁰Ceci n'est valable que pour les temps courts, lorsque la charge dans la grille flottante de la mémoire est encore faible.

⁴¹Correspondant au régime de rétention de charge.

⁴²Correspondant au régime d'écriture.

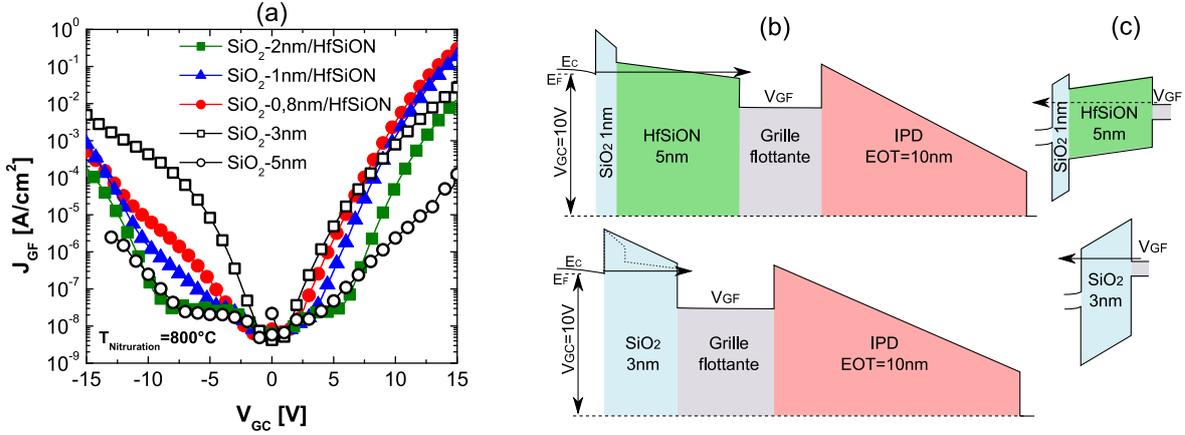


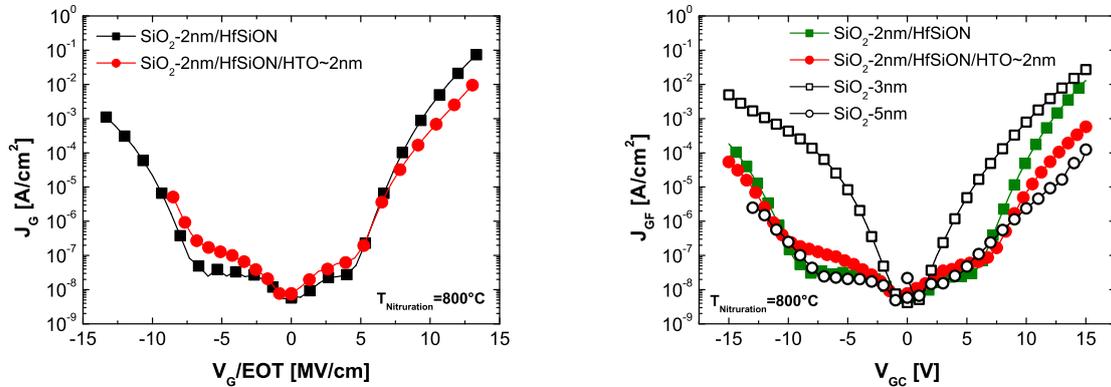
FIG. 3.52: (a) Densité de courant traversant l'empilement *tunnel*, intégré dans une cellule mémoire avec un interpoly de 10nm d'EOT, en fonction de la tension de grille de contrôle. Différents types d'empilements *tunnel* : monocouche de SiO_2 ou bicouche $SiO_2/HfSiON$. (b) Bande de conduction de la structure considérée à $V_{CG} = 10V$. Le courant tracé correspond donc au courant à travers l'empilement *tunnel* pour une certaine tension V_{CG} appliquée aux bornes de la cellule mémoire. (c) Zoom sur l'empilement *tunnel* à $V_{CG} = -10V$.

- À $V_G < 0$, le constat est moins clair. En effet, le côté asymétrique du bicouche impose une injection d'électrons côté $HfSiON$ pour les tensions négatives diminuant la sensibilité au champ (figure 3.52c). À faible champ, les empilements à base de $HfSiON$ isolent mieux que le SiO_2 3nm, mais à fort champ, ils présentent également des niveaux de courant inférieurs. En revanche, par rapport à du SiO_2 5nm, on mesure tout de même des courants plus forts à fort champ et du même ordre de grandeur à faible champ.

De manière plus générale, on peut noter une plus forte sensibilité au champ électrique des empilements à base de $HfSiON$ par rapport au SiO_2 . Cette amélioration de la sensibilité au champ électrique mesurée pour les empilements à base de $HfSiON$ peut être expliquée de manière qualitative par l'effet de *barrière en crête* mise en évidence dans le schéma de la figure 3.52b. En conclusion, les empilements bicouches à base de $HfSiON$ semblent très prometteurs pour l'intégration dans le stack *tunnel* des mémoires non volatiles à grille flottante.

3.5.2.2 Structure tricouche $SiO_2/HfSiON/HTO$

Nous avons également étudié des structures tricouches, où le $HfSiON$ est encapsulé par une couche de HTO de $\sim 2nm$. La figure 3.53a montre que le tricouche présente un courant très réduit à fort champ par rapport au bicouche. De plus les tricouches présentent un niveau de courant de fuite à bas champ plus important.



(a) Caractéristiques $J_G - V_G/EOT$ d'empilements bicouches $SiO_2/HfSiON$ et tricouches $SiO_2/HfSiON/HTO$.

(b) Courants à travers ces empilements en considérant un interpoly de 10nm d'EOT.

FIG. 3.53: Dans les deux cas, l'épaisseur de l'oxyde interfacial est de 2nm. Le $HfSiON$ est nitruré à 800 ° C.

Enfin la figure 3.53b présente le courant à travers les différents empilements, les courbes étant normalisées en supposant un interpoly de 10nm d'EOT. Il apparaît ici que le gain du $HfSiON$ est moindre sur les tricouches en comparaison avec les bicouches, en particulier à bas champ où le courant de fuite est plus important. En conclusion, ces mesures sembleraient indiquer que la structure tricouche est moins performante en tant qu'empilement *tunnel* en comparaison aux structures bicouches, essentiellement dû au courant de fuite élevé à faible tension.

Néanmoins, la structure tricouche ne devrait pas trop réduire l'injection à fort champ (cf figure 3.54). De plus, il est surprenant d'avoir un courant de fuite plus fort à faible champ avec un tricouche : sa couche de HTO devrait réduire la conduction à faible champ. En effet, le HTO joue le rôle de barrière pour les électrons piégés dans la grille flottante ce qui permettrait une meilleure rétention (cf figure 3.54c&f). Le principal défaut des structures tricouches, pour les applications *tunnel*, serait leur plus grande capacité de piégeage, ce qui pourrait fortement détériorer l'endurance. En effet, le "confinement" de la couche de $High-\kappa$ entre deux oxydes est plus favorable à la capture d'électrons dans ces empilements.



Les empilements bicouches $SiO_2 / HfSiON$ présentent une sensibilité accrue au champ électrique par rapport au SiO_2 , ce qui signifie qu'ils montrent une meilleure injection de courant à fort champ, idéale en régime de programmation de la mémoire, et une meilleure isolation à faible champ, pour assurer une bonne rétention. Ceci les rend particulièrement intéressants pour les applications en tant que *stack tunnel*.

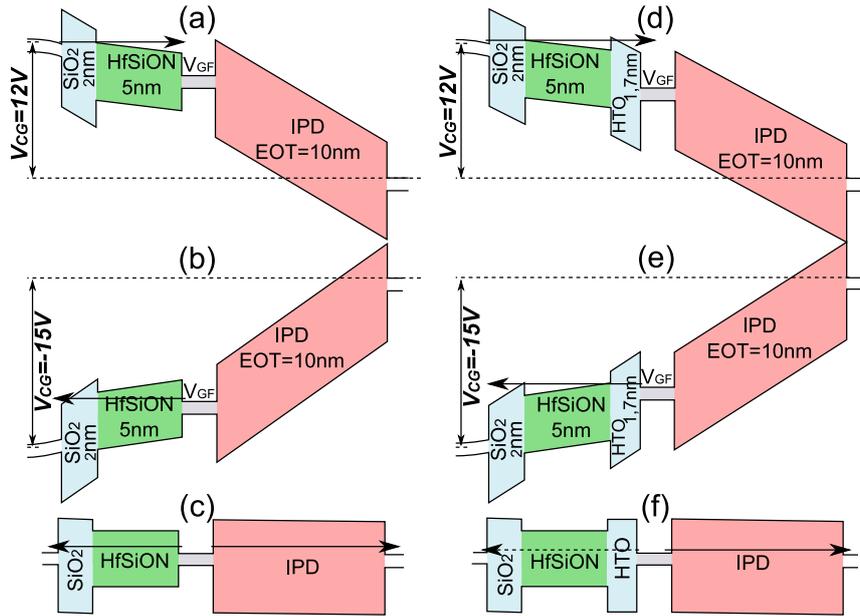


FIG. 3.54: Diagrammes de bandes d'une mémoire à grille flottante pour un empilement *tunnel* bicouche à (a) $V_{GC} = 12V$, (b) $V_{GC} = -15V$ et (c) $V_{GC} = 0$; ainsi que pour une structure tricouche $V_{GC} = 12V$ (d), $-15V$ (e) et 0(f).

3.6 Évaluation des performances mémoires

Dans cette section, à partir des mesures électriques présentées précédemment dans ce chapitre, nous allons évaluer les performances des mémoires à grille flottante qui intègreraient des diélectriques d'interpoly à base de *High-κ*.

3.6.1 Simulation de la fenêtre de programmation

Pour valider l'intérêt de l'intégration des *HfAlO* comme diélectrique d'interpoly dans les futures cellules mémoires Flash, nous avons simulé les performances de programmation grâce à un modèle analytique [De 01] de conservation de la charge. La variation de la charge piégée dans la grille flottante pendant l'écriture, l'effacement ou la rétention est donnée par l'équation 3.6.

$$\frac{dQ_{GF}}{dt} = J_{out} - J_{in} \quad (3.6)$$

Q_{GF} : La charge stockée dans la grille flottante.

J_{in} et J_{out} : Les densités de courant entrant et sortant de la grille flottante vers la grille ou le substrat, soit :

$$J_{in} = J_{in}^{sub} + J_{in}^{grille}$$

$$J_{out} = J_{out}^{sub} + J_{out}^{grille}$$

Les potentiels dans l'empilement sont calculés grâce aux équations 1.1 et 3.7.

$$V_{GF} = \frac{C_{IPD}}{C_{IPD} + C_{Tun}} \cdot V_{GC} + \frac{Q_{GF}}{C_{IPD} + C_{Tun}} \quad (3.7)$$

Pour simuler les caractéristiques d'écriture (figure 3.56), le courant d'écriture à travers l'oxyde *tunnel*, $J_{in} = J_{in}^{sub}$, est calculé suivant l'équation 2.8 grâce à un formalisme WKB (cf équation 2.16). Le courant de fuite à travers l'interpoly, J_{out}^{grille} , a été extrait des mesures expérimentales sur la figure 3.16. Les autres courants, J_{in}^{grille} et J_{out}^{sub} ont été considérés comme négligeables lors de l'écriture de la cellule mémoire. Il est donc possible de déterminer la charge injectée dans la grille flottante, donnée par l'équilibre des courants entrant et sortant.

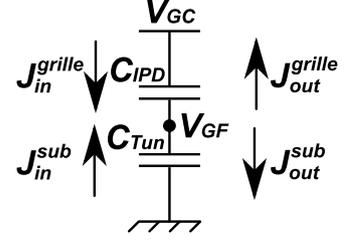


FIG. 3.55: Schéma électrique d'une cellule mémoire élémentaire.

Il est aussi possible d'obtenir une expression analytique de ΔV_T en négligeant le courant J_{out} et en considérant un courant J_{in} de type Fowler-Nordheim simplifié extrait de [Pan95] (équation 3.8). Ainsi, on obtient l'expression du ΔV_T de l'équation 3.9.

$$J_{in} = A \cdot F^2 \exp \frac{-B}{F} \quad (3.8)$$

$$A = \frac{m_e^{Si} \cdot q^3}{8\pi \cdot h \cdot m_e^{ox} \cdot \phi_b}$$

$$B = \frac{8\pi \sqrt{2m_e^{ox}} \phi_b^{3/2}}{3 \cdot h \cdot q}$$

ϕ_b : Écart entre la bande de conduction du *Si* et de l'oxyde *tunnel*.

m_e^{Si} : Masse effective des électrons du silicium dans la direction de l'injection.

m_e^{ox} : Masse effective des électrons dans l'oxyde tunnel.

$$\Delta V_T = V_P - \frac{1}{\alpha_G} \cdot \frac{B \cdot t_{ox}}{\ln \left(\frac{A \cdot B \cdot tps_P}{t_{ox} \cdot (C_{tun} + C_{IPD})} + \exp \left(\frac{B \cdot t_{ox}}{\alpha_G \cdot V_P} \right) \right)} \quad (3.9)$$

t_{ox} : Épaisseur de l'oxyde *tunnel*.

V_P : Tension de programmation.

tps_P : Temps de programmation.

Différentes simulations d'écriture sont reportées sur la figure 3.56. Deux types d'architecture ont été considérés :

- Une architecture standard pour la cellule mémoire, correspondant au nœud technologique de dimensions caractéristiques supérieures à $40nm$. Dans ce cas la grille de contrôle recouvre les flancs de la grille flottante (cf schéma des figures 1.14 et 3.56-a)[Par04, ITR07]. On considère un oxyde *tunnel* de $8nm$, une grille flottante en polysilicium et un diélectrique d'interpoly OHO ($HTO-4nm/HfAlO-9:1-9nm/HTO-4nm$). À partir de ces paramètres, on obtient un coefficient de couplage de $\alpha_G \sim 0,7$.

- Une architecture Flash NAND de $32nm$. On considère dans ce cas que la grille de contrôle ne recouvre pas les flancs de la grille flottante, de par le faible espace entre deux cellules adjacentes (cf figure 1.14). Aussi pour maintenir un coefficient de couplage de $\alpha_G \sim 0,7$, nous avons considéré un monocouche de $12nm$ de $HfAlO-9:1$ comme interpoly présentant une EOT réduite de $3,2nm$ (cf schéma de la figure 3.56-b).

Il apparaît que l'on atteint un ΔV_T de $3V$ pour une tension de stress de $17V$ et un temps d'écriture de $10\mu s$ pour les deux architectures. Ainsi, nous montrons l'intérêt des *High-κ* dans la course à la miniaturisation et à la forte densité mémoire tout en maintenant les performances d'écriture. En effet, les matériaux *High-κ* permettent de s'affanchir du recouvrement des flancs de la grille flottante, rendu difficile dans les technologies avancées.

On remarque aussi que l'expression analytique se superpose parfaitement avec la simulation de la caractéristique de programmation sauf en saturation. Cela nous indique que le courant de fuite à travers nos diélectriques interpoly n'intervient qu'en saturation. Il peut donc être négligé pour les temps d'écriture courts ($< 10ms$) pour ces tensions et ces épaisseurs.

En se basant sur le modèle analytique (équation 3.9), il est possible d'extraire l'équation 3.10 : expression analytique de la tension de programmation. On notera que le courant de fuite à travers l'interpoly a été négligé. À partir de cette équation, nous avons tracé un réseau de caractéristiques qui déterminent la tension de programmation en fonction de l'EOT du diélectrique d'interpoly (cf figure 3.57). Les figures 3.57a&c correspondent à un dispositif avec débordements latéraux alors que pour les figures 3.57b&d il n'y a pas de

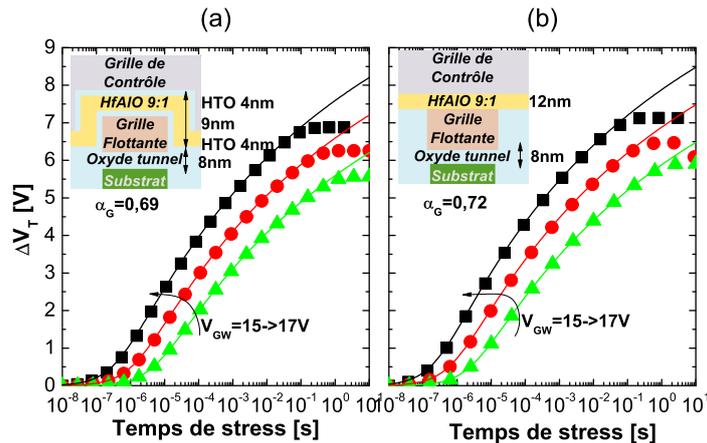


FIG. 3.56: Simulations de la caractéristique d'écriture pour une mémoire flash pour les nœuds technologiques (a) $63nm$ avec $8nm$ d'oxyde *tunnel*, un interpoly tricouche HTO- $4nm$ / $HfAlO-9:1-9nm$ /HTO- $4nm$ et une grille flottante continue avec mur latéral ($\alpha_G = 0,63$) (b) $35nm$ avec $8nm$ d'oxyde *tunnel* et $12nm$ $HfAlO-9:1$ en interpoly. Symboles : simulations basées sur les courants entrant et sortant (cf équation 3.6), lignes : expression analytique de l'équation 3.9.

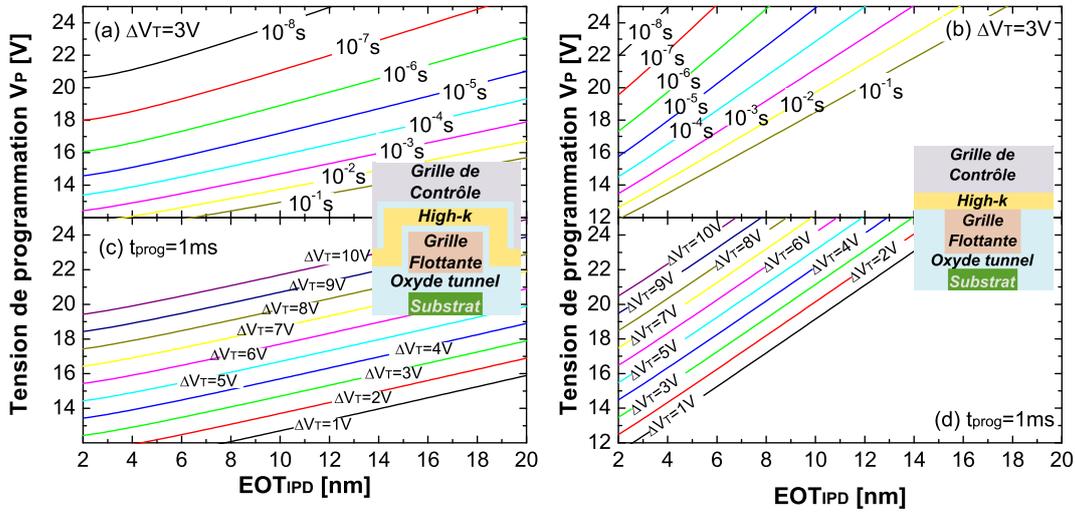


FIG. 3.57: Tension de programmation d'une cellule mémoire Flash NAND en fonction de l'EOT de l'interpoly pour obtenir a-b) un $\Delta V_T = 3V$ pour plusieurs temps d'écriture, c-d) temps d'écriture de $1ms$ pour plusieurs ΔV_T . $t_{tunnel} = 8nm$.

débordements latéraux. Les figures 3.57a&b donnent la tension de programmation pour obtenir un $\Delta V_T = 3V$ pour différents temps d'écriture. Alors que pour les figures 3.57c&d, le temps d'écriture est fixé à $1ms$, et l'on explore plusieurs ΔV_T .

$$V_P = \Delta V_T + \frac{\left(1 + \frac{S_1}{S_{IPD}} \cdot \frac{EOT_{IPD}}{t_{ox}}\right) \cdot t_{ox} \cdot B}{\ln\left(\frac{A \cdot B \cdot tps_P}{\varepsilon_{SiO_2} \cdot \left(1 + \frac{S_{IPD}}{S_1} \cdot \frac{t_{ox}}{EOT_{IPD}}\right)}\right)} \quad (3.10)$$

$\frac{S_{IPD}}{S_1}$: le rapport des surfaces en regard de la grille flottante de l'interpoly et du canal, soit pour une structure classique avec débordement latéral de la grille de contrôle : $\frac{S_{IPD}}{S_1} = 3$ et pour une structure plane sans débordement latéral : $\frac{S_{IPD}}{S_1} = 1$.

La figure 3.57 met en évidence une réduction de plusieurs volts de la tension de programmation due à la réduction de l'EOT de l'interpoly grâce l'utilisation d'empilements à base de *High-κ*. On observe les différences de comportement avec ou sans débordements latéraux :

- Les débordements latéraux donnent une dépendance faible de la tension de programmation en fonction de l'EOT du diélectrique d'interpoly. Il n'est donc pas pertinent de réduire fortement cette épaisseur : le gain en tension serait faible en comparaison à la perte en rétention.
- L'absence de débordements latéraux augmente fortement cette dépendance. Ainsi, il est intéressant de réduire l'EOT du diélectrique d'interpoly (intégration de *High-κ*) pour diminuer les tensions d'écriture. Mais plus encore, cela est indispensable pour assurer des tensions de programmation comprises entre 15V et 17V[ITR07].

On notera que dans le cas où l'on néglige le courant de fuite à travers l'interpoly, la dépendance entre tension de programmation et ΔV_T est de $1V$ pour $1V$.

Comme le montre [Via09b], on notera que pour une mémoire à couche de piégeage nitrure (SONOS, TANOS, ...) la dépendance du ΔV_T en fonction du V_G d'écriture n'est pas de $1V$ pour $1V$ comme pour les mémoires à grille flottante continue, mais plutôt de $0,6/0,7V$. Elle est dégradée en raison du courant à travers l'oxyde bloquant (J_{out}^{e-}) qui limite l'accumulation de charges dans la couche de piégeage [Pra06, Wan05a].

De plus, d'un point de vue pratique, cette hypothèse⁴³ est d'autant plus juste que la tension de programmation est faible, que le temps est court et que le ΔV_T est réduit. Ainsi, elle n'est pas réaliste pour les EOT d'interpoly trop faibles ou trop fortes :

- Dans un cas, les fuites dues à sa faible épaisseur seraient trop fortes.
- Dans l'autre, le faible couplage entre grille de contrôle et grille flottante serait trop faible, générant un champ important et un courant fort à travers l'interpoly.

Dans le cas où cette hypothèse ne pourrait être vérifiée, on observe une saturation du ΔV_T pour les mémoires à grille flottante continue. Il serait donc impossible d'extraire certains ΔV_T sur nos graphiques pour les faibles et forts EOT_{IPD} .



L'intégration d'interpoly à base de *High-κ* présente un fort intérêt pour les mémoires avancées. Elle permet de maintenir les tensions de programmation malgré la suppression des débordements latéraux de la grille de contrôle. Néanmoins, pour les applications industrielles, il est important d'évaluer l'impact de l'intégration des *High-κ* sur la rétention. En effet, les courants de fuite à faible champ assistés par pièges peuvent impacter grandement la rétention de nos dispositifs.

3.6.2 Simulation de la rétention

Les défauts et pièges présents dans les *High-κ* peuvent dégrader fortement la rétention [Mol06]. Aussi, en s'appuyant sur les mesures de courants présentées dans les figures 3.16, 3.17 et 3.18, nous avons évalué l'impact de l'intégration de *High-κ* dans le diélectrique d'interpoly de mémoires à grille flottante continue. Pour cela, nous avons linéarisé les courants mesurés près de leur zone de fonctionnement en rétention (cf équation 3.11).

$$J_{IPD} = e^{A \cdot F_{IPD} + B} \quad (3.11)$$

$F = \frac{V_G - V_{GF}}{EOT_{IPD}}$ le champ dans l'IPD calculé à partir de l'équation 1.3 en fonction de la charge stockée.

En intégrant l'équation 3.6 et en considérant que $J_{out}^{grille} = J_{IPD}$, on obtient l'expression de la charge stockée en fonction du temps de rétention (cf équation 3.12) ainsi que le temps de rétention ($t_{Ret}(p)$) pour un certain pourcentage de perte de charge : p (cf équation 3.13).

$$Q_{GF} = \frac{-1}{\gamma} \ln (e^{-\gamma \cdot Q_{GF}^{initial}} - t \cdot \gamma e^{\beta}) \quad (3.12)$$

⁴³Négliger le courant de fuite à travers l'interpoly.

$$t_{Ret}(p) = \frac{e^{-\gamma \cdot Q_{GF_{initial}}} - e^{-\gamma \cdot p \cdot Q_{GF_{initial}}}}{\gamma \cdot e^{\beta}} \quad (3.13)$$

$$\gamma = \frac{A}{(C_{IPD} + C_{Tun}) \cdot EOT_{IPD}}$$

$$\beta = \frac{A \cdot V_G (\alpha - 1)}{EOT_{IPD}} + B$$

Ainsi, en s'appuyant sur nos mesures expérimentales réalisées sur des capacités (figures 3.16, 3.17 et 3.18) et l'équation 3.13, nous pouvons extrapoler les temps de rétention pour différents types d'empilements d'interpoly. La figure 3.58 présente ces extrapolations pour différents pourcentages de perte de charge : 10%, 20% et 30%.

On observe une amélioration de la rétention avec l'augmentation de la concentration en *Al* pour les *HfAlO*. Ainsi, la rétention est améliorée grâce à l'augmentation du gap des matériaux *High-κ*. De plus, l'augmentation de l'épaisseur du *High-κ* permet une légère augmentation de la rétention.

Un autre point important est l'activation en température du courant de fuite dans les empilements à base de *High-κ* (cf figure 3.19). De la même façon, nous nous intéressons au temps de rétention en fonction de la température. Ici nous étudions des monocouches de

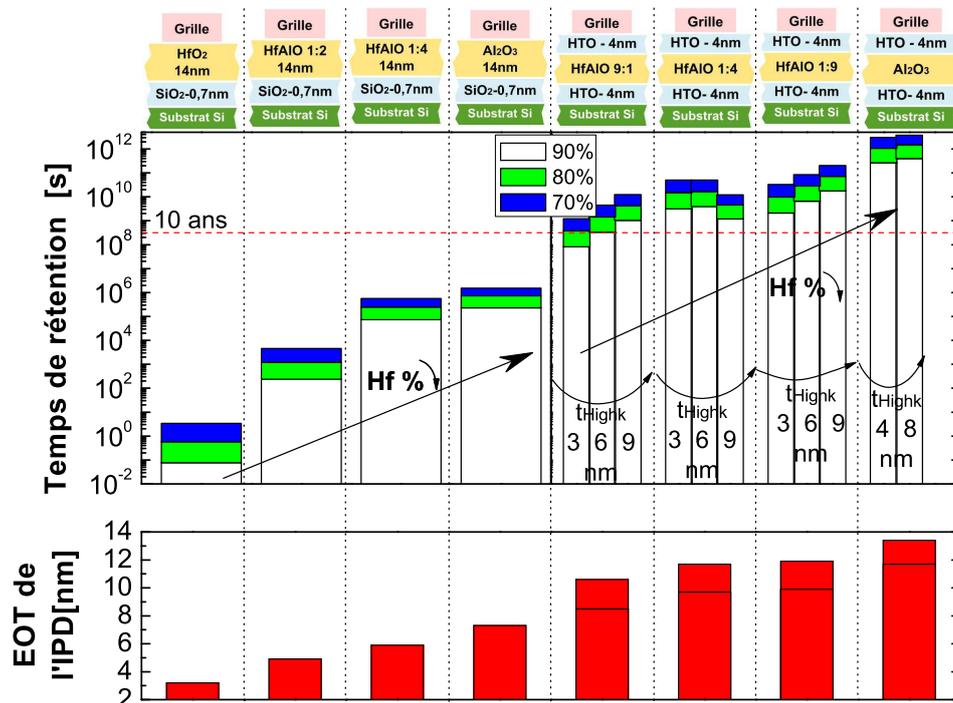


FIG. 3.58: Simulations du temps de rétention d'une cellule mémoire flash intégrant différents *High-κ* en interpoly. Le courant de fuite durant la rétention est extrait des $J_G - V_G$ présentées dans le chapitre 3. Le temps de rétention a été obtenu pour différents pourcentages de charge perdue 10%, 20% et 30%. Le ΔV_T initial est de 8V. Aucune fuite dans l'oxyde tunnel (8nm) n'a été considérée.

High-κ comme diélectrique d'interpoly. Le tracé d'Arrhenius, figure 3.59, nous montre une augmentation de l'activation en température avec l'augmentation de la concentration en *Hf* pour les *HfAlO*. Cette tendance correspond aux mesures de courant en température (cf tableau 3.20). Néanmoins, les énergies d'activation dépendent du champ électrique auquel elles sont extraites, ce qui peut expliquer les différences entre celles obtenues ici et celles extraites dans le tableau 3.20.

En conclusion, la figure 3.60 résume la problématique entre tension de programmation et rétention. Elle représente le temps de rétention simulé, extrait pour une perte de charge de 10% d'un $\Delta V_T = 6V$, en fonction de la tension qui faudrait appliquer pour atteindre un ΔV_T de 6V en 1ms (cf équation 3.10).

Ainsi, l'utilisation d'un tricouche (*HTO/High-κ/HTO*) améliore significativement la rétention mais dégrade aussi grandement l'EOT de l'empilement donc le coefficient de couplage et les propriétés de programmation et d'effacement. L'utilisation *HfAlO* en tricouche permet la réduction de l'EOT mais au détriment de l'isolation. Les monocouches d' Al_2O_3 apparaissent aussi comme des candidats intéressants grâce à leurs faibles courants de fuite à faible champ et leur faible activation en température.

On retiendra que les 10 ans de rétention sont atteints avec les tricouches pour des applications à grille flottante, alors que les monocouches ont des difficultés pour atteindre cette spécification. Il est donc impératif d'utiliser les *High-κ* dans des structures soit bicouche (*HTO* suffisamment épais côté grille flottante) soit tricouche pour ces applications. Néanmoins, pour les mémoires à couche de piégeage (par exemple *SANOS*) l'utilisation de l'alumine monocouche est très envisageable car l'émission est fortement contrôlée par la zone de piégeage. Ce point sera discuté dans le chapitre 4.

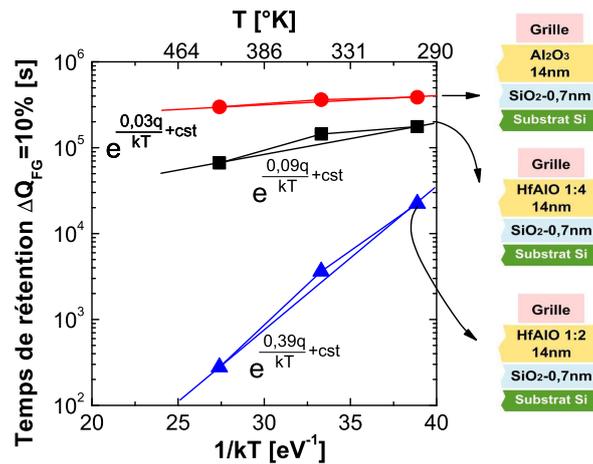


FIG. 3.59: Simulations du temps de rétention d'une cellule mémoire flash intégrant différents monocouches de *High-κ* en interpoly en fonction de la température de rétention. Le courant de fuite durant la rétention est extrait des $J_G - V_G$ présentées dans le chapitre 3. Le temps de rétention a été extrait pour une perte de charge de 10% d'un ΔV_T initial de 8V. Aucune fuite dans l'oxyde *tunnel* (8nm) n'a été considérée.

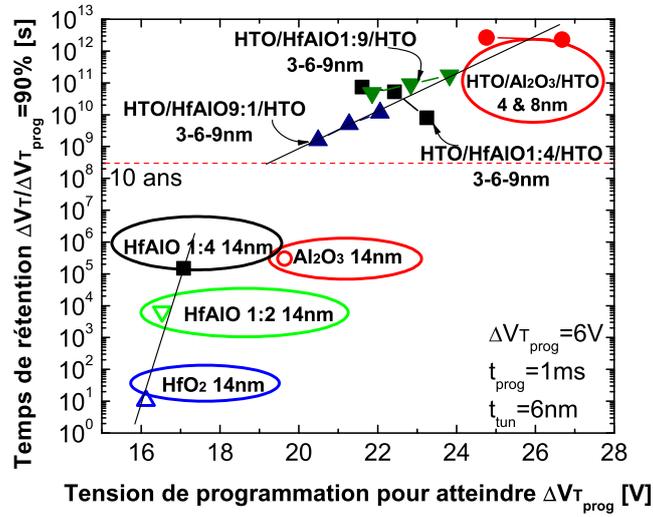


FIG. 3.60: Simulations du temps de rétention d'une cellule mémoire flash à grille flottante continue intégrant différents interpolys de *High- κ* fonction de la tension de programmation. Le temps de rétention a été obtenu pour une perte de charge de 10% d'un ΔV_T initial de 6V. La tension de programmation permettrait d'atteindre $\Delta V_T = 6V$ en 1ms (cf équation 3.10). Aucune fuite dans l'oxyde *tunnel* (6nm) n'a été considérée.

3.7 Conclusion

Dans ce chapitre nous avons étudié les propriétés des matériaux *High- κ* en vu de leur intégration dans des mémoires non-volatiles. Nous avons mis en évidence l'intérêt des aluminates d'hafnium grâce au contrôle de leurs propriétés physiques par l'ajustement de la concentration en *Hf* (cf tableau 3.6).

	Augmentation du % de <i>Hf</i>
Gap	↘
Propriétés de couplage	↗
Température de cristallisation	↘
Charges fixes négatives	↘
Piégeage	↗
Courant à fort champ	↘
Courant à faible champ	↗
Activation en température	↗

TAB. 3.6: Résumé des propriétés des *HfAlO*.

Pour ces matériaux nous avons aussi mis en évidence des charges fixes négatives et un piégeage aux interfaces *HfAlO/SiO₂*. De plus, dans une structure tricouche, l'impact de l'épaisseur du *High- κ* (entre 3nm et 9nm) ne modifie pas les courants à V_G/EOT

constant. De même, la diminution de l'épaisseur du *HTO* côté anode augmente légèrement la conduction à fort champ, alors que celle côté cathode augmente le courant à faible champ.

La nitruration des aluminates d'hafnium a aussi été étudiée. Nous avons mis en évidence qu'elle avait un rôle important sur la température de cristallisation et le piégeage dans les *HfAlO* (cf tableau 3.7). Pour le *HfO₂*, l'azote incorporé à l'interface réduit le piégeage d'électrons. Alors que pour les *HfAlO*, la nitruration réduit le piégeage en augmentant la température de cristallisation et en limitant le taux de cristallinité de la couche de *High-κ*. La nitruration permet donc de réduire le piégeage dans les *High-κ* à base de *Hf*.

	Impact de la nitruration sur		
	<i>Al₂O₃</i>	<i>HfAlO-1:2</i>	<i>HfO₂</i>
Incorporation taux d'azote	volumique faible	volumique important - 16%	à l'interface important - 25%
Gap optique	=	↘	=
Cristallinité avant recuit	Amorphe	Amorphe	Cristallin
après recuit spike	=	Limite la cristallinité ↘ taille de grain	=
Propriétés de piégeage		↘	

TAB. 3.7: Résumé de l'impact de la nitruration sur les *HfAlO*.

Nous avons aussi étudié l'impact du recuit sur l'alumine et mis en évidence la nécessité d'un recuit à forte température (alumine cristalline) pour assurer une bonne compatibilité avec une grille métallique, en particulier *TiN* (cf tableau 3.8).

	<i>Al₂O₃</i>	
	Amorphe	Cristallin
Densité	↘	↗
Stress	↘	↗
Épaisseur physique	↗	↘
Concentration en H	↗	↘
Gap	↘	↗
Compatibilité avec le <i>TiN</i>	↘	↗

TAB. 3.8: Résumé des propriétés de l'*Al₂O₃*.

Enfin les silicates d'hafnium ont été étudiés en vu de leur intégration en tant qu'oxyde *tunnel*. Les structures bicouches sont apparues comme de bons candidats de par leur très bonne sensibilité au champ (cf tableau 3.9).

	Augmentation de la concentration en Hf	Augmentation de la T ° C de nitruration
Gap	↘	↘
Température de cristallisation	↘	↗
Sensibilité au champ	$SiO_2 < \text{Tricouche} < \text{Bicouche}$	

TAB. 3.9: Résumé des propriétés des $HfSiON$.

Enfin, la suppression des débordements latéraux entre la grille de contrôle est la grille flottante nécessite l'intégration de matériaux *High- κ* . Les empilements tricouches permettent une bonne rétention tout en maintenant des tensions de programmation acceptables. Les empilements monocouches permettent une forte réduction des tensions de programmation mais présentent des difficultés à atteindre les 10 ans de rétention. En vu de nos résultats, on peut donc penser qu'un empilement bicouche $HTO/HfAlO-1:4$ serait le plus performant des applications à grille flottante.

Chapitre 4

Intégration des matériaux à forte permittivité électrique dans les mémoires non-volatiles

Dans le chapitre précédent, nous avons présenté quelques résultats morphologiques et électriques sur les *High- κ* en vue d'une bonne compréhension de ces matériaux avant leur intégration dans des mémoires non-volatiles. Dans ce chapitre, nous nous intéresserons à leur comportement une fois intégrés dans des dispositifs mémoires et nous allons essayer de relier les propriétés matériaux obtenues dans le chapitre 3 avec les performances des dispositifs mémoires étudiés. Nous verrons dans une première partie l'intégration des matériaux *High- κ* comme diélectrique d'interpoly pour des applications à nanocristaux. Ensuite, nous étudierons les mémoires à couche de piégeage en nitrure intégrant de l'alumine en oxyde bloquant, notamment l'impact des épaisseurs. Enfin, les oxydes *tunnel*, *SiO₂/HfSiON*, ont été évalués dans des dispositifs intégrant les améliorations mis en place dans les parties précédentes.

Sommaire

4.1	<i>High-κ</i> dans les mémoires à nanocristaux	157
4.1.1	Procédé de fabrication des mémoires à nanocristaux	157
4.1.2	Caractéristiques de programmation des mémoires à nanocristaux	159
4.1.3	Caractéristiques de rétention et d'endurance	160
4.2	Mémoire à base de nitrure avec oxyde bloquant <i>High-κ</i> . . .	162
4.2.1	Impact de l'épaisseur de l'oxyde <i>tunnel</i>	163
4.2.2	Impact de l'oxyde bloquant	165
4.2.2.1	Caractéristiques d'écriture et d'effacement	165
4.2.2.2	Caractéristiques d'endurance	169
4.2.2.3	Caractéristiques de rétention	170
4.2.3	Impact de l'épaisseur de l'alumine et du nitrure	171
4.2.3.1	Caractéristiques d'écriture et d'effacement	172

4.2.3.2	Caractéristiques d'endurance	176
4.2.3.3	Caractéristiques de rétention	178
4.2.4	Impact des traitements thermiques sur l'alumine	184
4.2.4.1	Caractéristiques d'écriture et d'effacement	185
4.2.4.2	Caractéristiques de rétention	187
4.2.5	Impact de la grille de contrôle	188
4.2.5.1	Étude des grilles en polysilicium	188
4.2.5.2	Étude des grilles métalliques en TiN, TaN ou TaAlN .	190
4.3	Mémoire à base de nitrure avec oxyde <i>tunnel</i> en $SiO_2/HfSiON$	195
4.3.1	Caractéristiques d'écriture et d'effacement	196
4.3.2	Caractéristiques de rétention	199
4.4	Conclusion	201

4.1 Intégration de *High-κ* en tant que diélectrique d'interpoly pour les mémoires à nanocristaux

Dans cette partie, nous allons étudier l'intégration des matériaux *High-κ*, en tant que diélectrique d'interpoly, dans des mémoires à nanocristaux de silicium. Les mémoires à nanocristaux sont une alternative intéressante dans le suivi de la *road map*. L'utilisation d'une zone de piégeage discrète permet une meilleure réduction de l'oxyde *tunnel*, une plus grande robustesse au SILC (Stress Induced Leakage Current) et un faible couplage parasite entre cellules adjacentes [Mol07]. En effet, contrairement à une mémoire à grille flottante continue, un défaut dans l'oxyde *tunnel* n'affecte que les charges en regard et ne vide pas toute la charge stockée. (cf figure 4.1). L'une des difficultés de ces mémoires est le faible couplage entre les nanocristaux et la grille de contrôle. L'aspect discret des nanocristaux rend inutile les débordements latéraux de la grille de contrôle. Aussi, sans l'intégration de *High-κ* dans le diélectrique d'interpoly :

- Le coefficient de couplage est très faible.
- Le courant de fuite à travers le diélectrique d'interpoly durant l'écriture et l'effacement rendent la programmation très délicate voir impossible ($J_{IPD} \sim J_{Tun}$).

Ainsi, l'utilisation d'*High-κ* permet d'augmenter le coefficient de couplage et de réduire les fuites dans le diélectrique d'interpoly ($J_{IPD} \ll J_{Tun}$). En effet, les empilements à base de *High-κ* présentent une meilleure isolation à fort champ et un bien meilleur couplage qu'un simple *HTO* (cf figures 3.16&3.17).

4.1.1 Procédé de fabrication des mémoires à nanocristaux

Les nanocristaux de silicium, réalisés dans les dispositifs que nous allons étudier, sont déposés par LPCVD (diamètre $\phi = 6 - 10nm$ et densité $d \approx 9 \cdot 10^{11}cm^{-2}$) sur $4nm$ d'oxyde thermique. Le dépôt des nanocristaux est réalisé en deux étapes :

- Une étape brève de nucléation, réalisée sous chimie silane.
- Une étape de croissance où l'on utilise du dichlorosilane (SiH_2Cl_2) pour faire croître sélectivement les nucléis sans former de nouveaux nanocristaux.

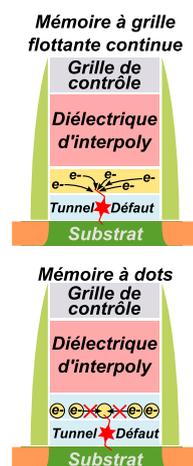


FIG. 4.1: Schémas montrant l'intérêt des mémoires à nanocristaux face au SILC.

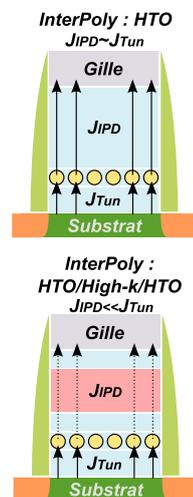


FIG. 4.2: Schémas des courants dans une mémoire à nanocristaux en fonction du couplage avec la grille de contrôle.

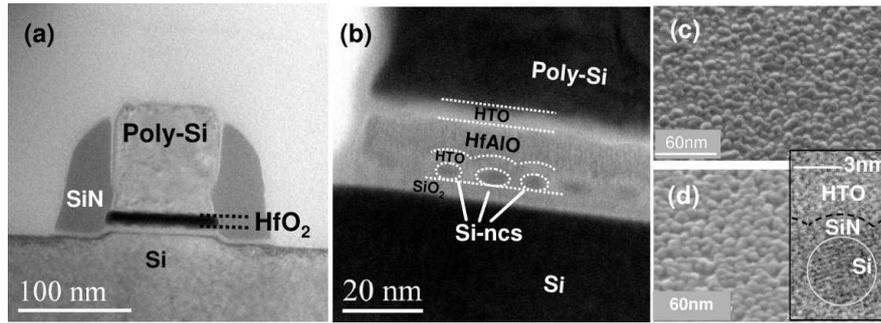


FIG. 4.3: (a-b) Photographies MET de la section de mémoires à nanocristaux intégrant des tris-couches *HTO/High-κ/HTO* comme diélectrique d'interpoly. Photographies MEB des nanocristaux de silicium sans c) et avec d) enrobage de nitrure. Insert : Photographie MET de la section de nanocristaux avec son enrobage de nitrure.

Cette technique permet de séparer la nucléation et la croissance des nanocristaux de silicium et ainsi qu'un meilleur contrôle de la taille, de la densité et de l'homogénéité des nanocristaux. Ces caractéristiques morphologiques impactent grandement les caractéristiques mémoires, d'où l'importance de les contrôler [Jac08, Jac07]. Après dépôt, ils ont été passivés par un recuit NH_3 à $750^\circ C$ pour éviter leur oxydation.

Ensuite, différents diélectriques interpoly sont réalisés. Ils sont composés d'une couche de $8nm$ de *High-κ* (HfO_2 , $HfAlO$, Al_2O_3 et Si_3N_4). Elle est située entre deux *HTO* de $4nm$ pour les empilements tricouches. Les EOT respectives des différents diélectriques d'interpoly tricouches sont $10nm$, $10,5nm$, $11,2nm$ et $12,7nm$.

Une grille, en polysilicium dopée in situ N^+ , a été déposée sur les diélectriques d'interpoly tricouches. Pour les diélectriques d'interpoly bicouches, une grille en *TiN* a été réalisée.

La figure 4.3 présente différentes photographies, obtenues par microscope électronique, des dispositifs que nous allons étudier. Nous pouvons observer les trois couches de l'oxyde bloquant : *HTO/High-κ/HTO*, ainsi que les nanocristaux de silicium présents dans une matrice de SiO_2 ou de nitrure.

La figure 4.4 présente les variantes technologiques étudiées ainsi que l'EOT, mesurée électriquement, de l'empilement complet.

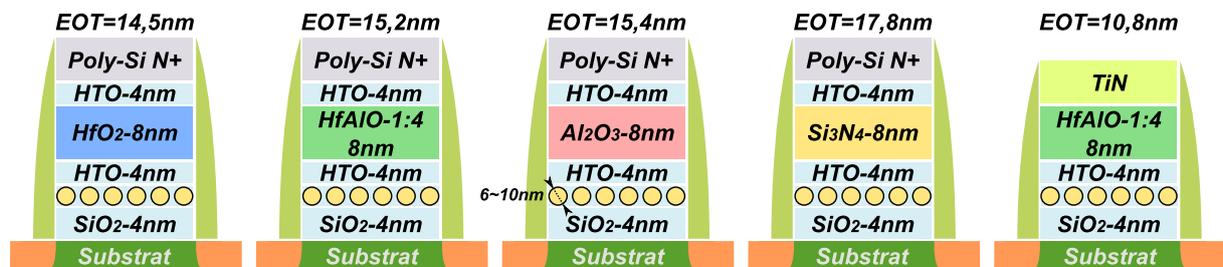


FIG. 4.4: Schémas des variantes technologiques de mémoires à nanocristaux avec l'EOT totale des différents empilements.

4.1.2 Caractéristiques de programmation des mémoires à nanocristaux

La figure 4.5 représente les caractéristiques d'écriture et d'effacement en Fowler-Nordheim pour différents types de diélectriques d'interpoly. Nous présentons, ici, des empilements tricouches *HTO-4nm/High-κ-8nm/HTO-4nm* et des empilements bicouches *HTO-4nm/High-κ-8nm*. Le réseau de caractéristiques regroupe plusieurs tensions d'écriture (14V/15V/16V/17V) et d'effacement (-12V/-13V/-14V/-15V).

Tout d'abord, il est important de constater qu'il est possible de programmer une mémoire à nanocristaux en Fowler-Nordheim grâce à l'intégration de *High-κ*. Cela est dû à l'augmentation du coefficient de couplage et de leurs faibles courants de fuite à fort champ (cf figures 3.16 et 3.17). En effet, des dispositifs avec un diélectrique d'interpoly en *HTO* ont aussi été étudiés (non montrés ici) mais leur faible coefficient de couplage et leur fort courant de fuite rendaient l'écriture et l'effacement impossibles.

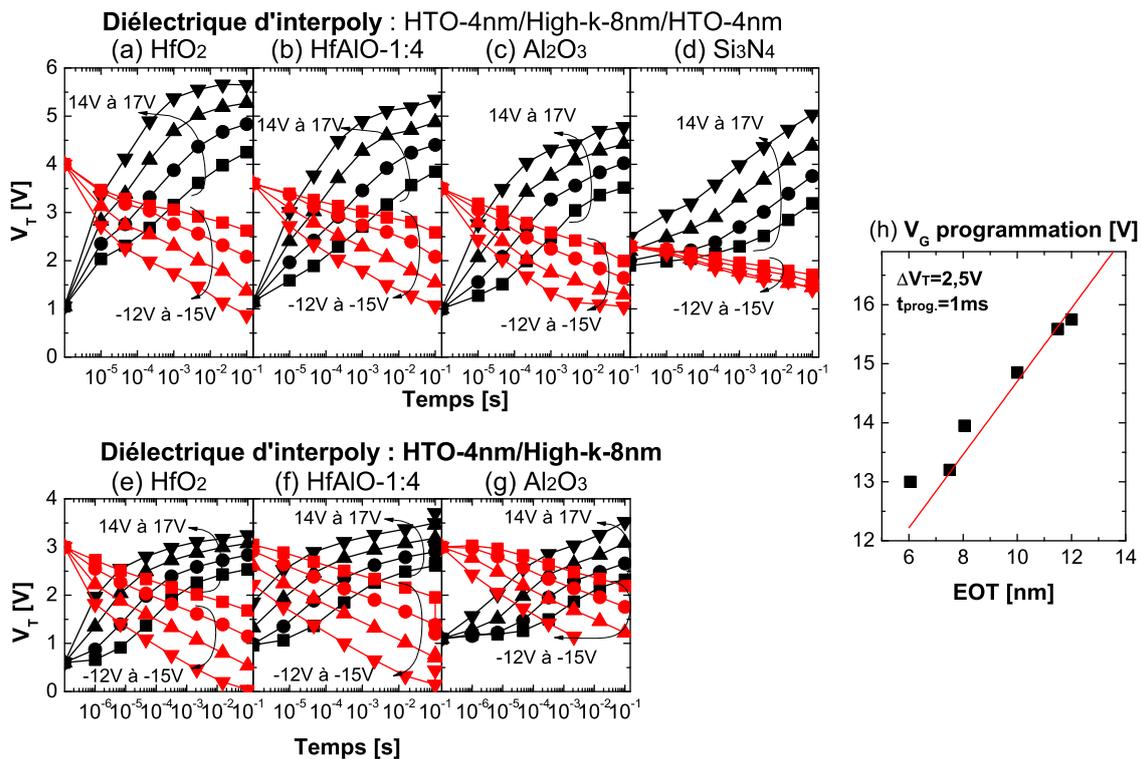


FIG. 4.5: Caractéristiques d'écriture et d'effacement de mémoires à nanocristaux de silicium pour différents diélectriques d'interpoly : tricouche *HTO-4nm/High-κ-8nm/HTO-4nm* (a) *HfO₂*, b) *HfAlO1:4*, c) *Al₂O₃* et d) *Si₃N₄*, et bicouches *HTO-4nm/High-κ-8nm* (e) *HfO₂*, f) *HfAlO1:4*, g) *Al₂O₃* ($W \times L = 0,5 \mu m \times 0,25 \mu m$). (h) Tensions de programmation extraites pour obtenir $\Delta V_T = 2,5V$ en $1ms$.

Il apparaît que l'utilisation de *High- κ* améliore fortement la fenêtre de programmation. Pour les diélectriques d'interpoly tricouches, plus la permittivité du *High- κ* est grande plus la fenêtre s'ouvre. En effet, la diminution de l'EOT permet une nette amélioration des performances d'écriture et d'effacement. La figure 4.5h met en évidence cet effet. Elle représente les tensions de programmation qu'il a fallu appliquer sur nos différents dispositifs pour obtenir un $\Delta V_T = 2,5V$ en $1ms$. On observe une décroissance linéaire de la tension de programmation avec l'EOT du diélectrique d'interpoly. La courbe en trait plein de la figure 4.5h correspond à l'équation 3.10 du chapitre 3, ainsi on remarque le même comportement que pour une mémoire à grille flottante continue.

Néanmoins, les caractéristiques de programmation des dispositifs intégrant des diélectriques bicouches ont une saturation plus rapide que ceux intégrant des tricouches. Cela peut provenir du plus fort courant de fuite à travers les bicouches. Il est aussi possible que le fort couplage entre la grille de contrôle et les nanocristaux impacte la saturation, car demandant une plus grande quantité d'électrons pour générer le même ΔV_T .

4.1.3 Caractéristiques de rétention et d'endurance

Les figures 4.7a et 4.7b représentent les caractéristiques de rétention de nos différents empilements à $25^\circ C$ et à $125^\circ C$. L'objectif est d'évaluer l'impact des diélectriques d'interpoly à base de *High- κ* sur la rétention des mémoires à nanocristaux. Il apparaît que les empilements à base d' HfO_2 présentent une perte de charge plus importante que pour l' Al_2O_3 . En effet, l'alumine possède une plus forte hauteur de barrière que l' HfO_2 ¹ entraînant un courant de fuite plus fort à faible champ.

Enfin, nous avons extrait le temps de rétention à 15% de charges perdues pour plusieurs températures². On n'observe aucune activation en température en dessous de $125^\circ C$. Ce bon comportement peut être dû à la différence de la zone de stockage des électrons : pièges ou bande de conduction, mais aussi à leur profondeur énergétique dans les nanocristaux. Dans ces derniers, les électrons sont stockés sur la bande de conduction du *Si* soit un écart de $\sim 3,1eV$ entre le niveau de l'électron et la bande de conduction du *SiO₂*, alors que pour une mémoire à couche de piègeage à base de nitrure, les électrons sont stockés dans le gap (cf figure 4.6). Ces pièges sont à une profondeur $\sim 1eV$ [Arr07], soit un écart $\sim 2,1eV$ avec la bande de conduction du *SiO₂*. Il est ainsi plus facile d'activer en température un électron piégé dans du nitrure qu'un électron stocké sur la bande de conduction d'un nanocristal.

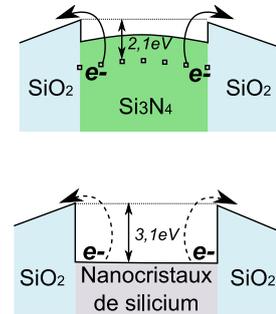


FIG. 4.6: Diagrammes des bandes de conduction à $V_G = 0V$ pour une mémoire SONOS et à nanocristaux de silicium.

¹Pour le HfO_2 $\Delta E_C \sim 1,5eV$ et pour l' Al_2O_3 $\Delta E_C \sim 2,8eV$, cf tableau 3.1.

²Défini comme le temps pour lequel l'on a perdu 15% de la charge initialement piégée.

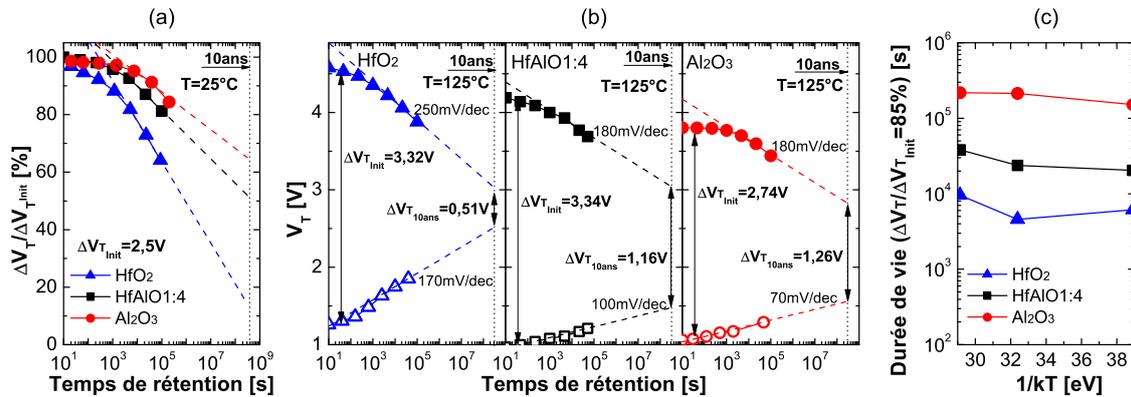


FIG. 4.7: Caractéristiques de rétention de mémoires à nanocristaux de Si pour des diélectriques d'interpoly tricouches $\text{HTO-4nm}/\text{High-}\kappa\text{-8nm}/\text{HTO-4nm}$ intégrant différents types de *High-κ*. a) Pourcentage de charge piégée en fonction du temps de rétention à 25°C . b) Fenêtre de programmation en fonction du temps de rétention à 125°C . c) Temps de rétention pour 15% de charge perdue en fonction de la température.

La figure 4.8 représente les caractéristiques d'endurance des différents empilements. On observe que le diélectrique d'interpoly à base de HfO_2 présente une fenêtre de programmation constante de $\sim 3\text{V}$ sur 10^5 cycles, avec une dérive de $\sim 1,5\text{V}$. Elle provient d'une faible dégradation de la pente sous le seuil³ mais principalement d'un décalage des $I_D - V_G$ dû à un piégeage parasite dans le HfO_2 . La forte dérive du V_T des dispositifs à base d' Al_2O_3 provient d'une forte dégradation des pentes sous le seuil.

Les HfAlO présentent un bon compromis entre l' HfO_2 et l' Al_2O_3 en terme de fenêtre de programmation, de dérive en endurance et en rétention. Les diélectriques d'interpoly à base de Si_3N_4 présentent une caractéristique d'un fort piégeage parasite dans le diélectrique d'interpoly. En effet, on observe un décalage des caractéristiques $I_D - V_G$, sans dégradation de la pente sous le seuil, typique d'un piégeage non-réversible dans l'empilement.



En conclusion, nous avons pu mettre en évidence l'intérêt de l'intégration des *High-κ* dans les mémoires à nanocristaux de silicium. Il est apparu que les diélectriques d'interpoly à base de HfO_2 présentaient de bonnes fenêtres de programmation ainsi qu'une bonne endurance en Fowler-Nordheim, alors que l' Al_2O_3 possède une perte de charge en rétention plus faible que le HfO_2 . Aussi les HfAlO présentent un bon compromis entre le HfO_2 et le Al_2O_3 en termes de fenêtre de programmation, de dérive en endurance et en rétention.

³Due à une dégradation de l'interface Si/SiO_2 générée par l'écriture et l'effacement. En effet, l'injection d'électrons à travers l'oxyde tunnel peut dégrader cette interface. Elle est d'autant plus grande que le champ d'injection (et donc la vitesse des électrons) est fort.

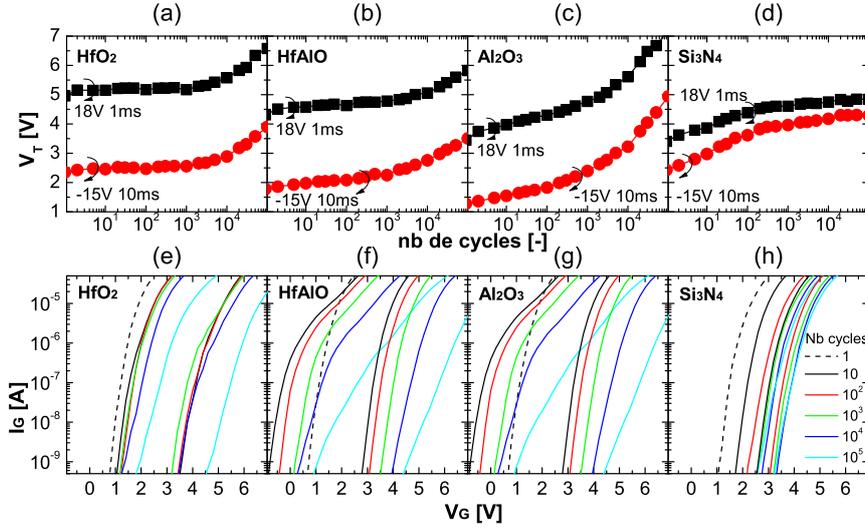


FIG. 4.8: (a-d) Caractéristiques d'endurance de mémoires à nanocristaux de Si pour des diélectriques d'interpoly tricouches HTO-4nm/high- κ -8nm/HTO-4nm intégrant différents types de *High-κ*. (e-h) Caractéristiques $I_D - V_G$.

4.2 Mémoire à base de nitrure avec oxyde bloquant *High-κ*

Dans cette partie, nous allons nous intéresser aux mémoires à base de nitrure, en particulier, celle intégrant de l'alumine comme oxyde bloquant. Ces empilements (TANOS⁴, SANOS⁵, SAONOS⁶ ...) font l'objet d'un fort engouement de la part de l'ensemble des acteurs du marché des mémoires de type NAND [ITR07]. Aussi, dans cette partie, nous allons nous intéresser à l'influence de chacune de ces couches :

- Oxyde *tunnel* : impact de son épaisseur sur la programmation et la rétention.
- Oxyde bloquant : intérêt d'intégrer de l'alumine ou un bi-couche *HTO/Al₂O₃*.
- Couche de piégeage : impact de son épaisseur sur l'efficacité de piégeage et de la rétention, influence du couple des épaisseurs *Si₃N₄/Al₂O₃*.
- Grille de contrôle : intérêt de l'intégration des grilles à fort travail de sortie.

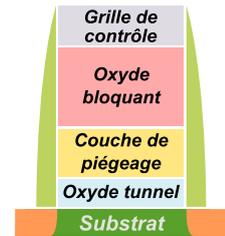


FIG. 4.9: Schéma d'une mémoire à couche de piégeage.

⁴TaN/Al₂O₃/SiN/SiO₂/Si

⁵Poly-Si/Al₂O₃/SiN/SiO₂/Si

⁶Poly-Si/Al₂O₃/SiO₂/SiN/SiO₂/Si

Les objectifs sont :

- D'évaluer les performances de programmation et la fiabilité de ces dispositifs.
- D'étudier la compatibilité des différentes couches et de comprendre les mécanismes physiques contrôlant ces mémoires.

Pour cela, nous utiliserons, le plus souvent possible, l'un des modèles de piégeage présenté dans le chapitre 2 (paragraphe 2.3.2 et 2.3.3).

4.2.1 Impact de l'épaisseur de l'oxyde *tunnel*

Dans un premier temps, nous allons étudier l'impact de l'épaisseur de l'oxyde *tunnel* sur les caractéristiques de piégeage et de rétention d'empilements SONOS. Une couche de nitrure de $6nm$ a été déposée sur l'oxyde *tunnel* par LPCVD. L'oxyde bloquant est un *HTO*⁷ de $12nm$. Les réseaux des caractéristiques d'écriture de ces empilements sont regroupés dans la figure 4.11.

La figure 4.12a compare les caractéristiques d'écriture de ces empilements. Il apparaît que les oxydes les plus fins donnent les vitesses de programmation les plus élevées. Cela est dû à deux effets :

- L'augmentation du champ dans l'oxyde *tunnel*, à V_G constant.
- Le changement du mode d'injection de Fowler-Nordheim à *tunnel direct*⁸.

On peut remarquer que, pour ces empilements, la tension de transition entre l'injection Fowler-Nordheim à *tunnel direct* s'effectue :

- à plus de $22V$ pour le *tunnel* à $2,5nm$,
- à $18V$ pour le $3,5nm$,
- à $14V$ pour le $5nm$.

Aussi, le comportement similaire des oxydes $2,5nm$ à $15V$ et $3,5nm$ à $18V$ est dû :

- Aux transparences de l'oxyde *tunnel* similaires pour les deux dispositifs (aires des barrières égales, cf figure 4.10).
- Au même mode de conduction (*tunnel direct*), donnant au courant d'injection la même dépendance en champ (cf figure 4.10).

On observe donc la même cinétique de piégeage.

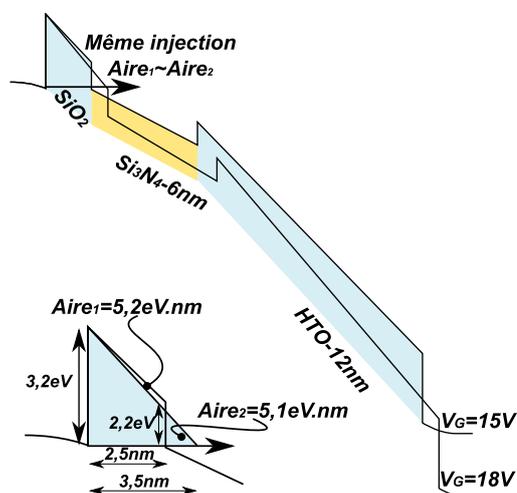


FIG. 4.10: Diagrammes des bandes de conduction d'une mémoire SONOS pour différentes tensions et épaisseurs d'oxyde *tunnel*.

⁷Le *HTO* est un oxyde de silicium obtenu par dépôt réalisé à haute température : *High Thermal Oxide*

⁸En injection *tunnel direct*, à champ constant, l'épaisseur intervient directement dans l'expression du courant (pas seulement pas l'intermédiaire du champ) ce qui n'est pas le cas en Fowler-Nordheim (cf figure 2.2).

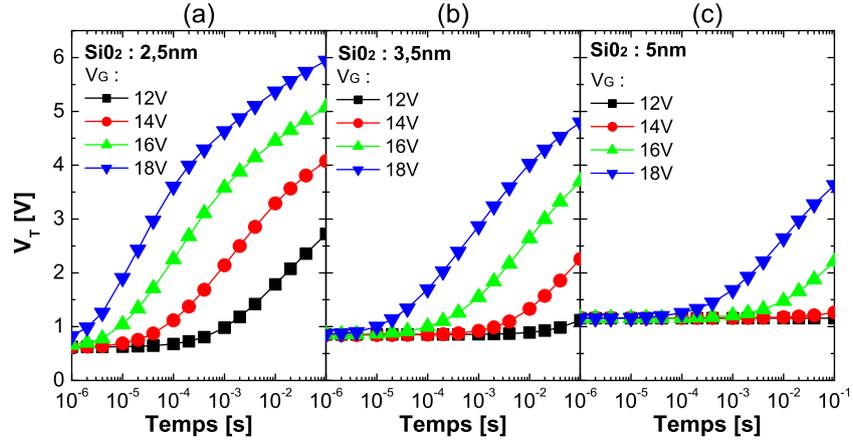


FIG. 4.11: Réseaux des caractéristiques d'écriture de SONOS intégrant différentes épaisseurs d'oxyde *tunnel* : a) 2,5nm, b) 3,5nm et c) 5nm. $HTO=12,5\text{nm}$ et $t_{\text{Si}_3\text{N}_4} = 6\text{nm}$.

De plus, on observe qu'une augmentation de 1-1,5nm entraîne une augmentation de 3V de la tension d'écriture pour les mêmes ΔV_T .

La figure 4.12b présente les caractéristiques de rétention de ces dispositifs à 125°C . Il apparaît qu'un oxyde *tunnel* de 3,5nm améliore nettement la rétention en comparaison à un oxyde *tunnel* 2,5nm. De plus, le 5nm n'améliore pas la rétention pour un $\Delta V_{T\text{initial}} \sim 3\text{V}$ (soit ici $V_{T\text{initial}} \sim 4\text{V}$) par rapport à l'oxyde *tunnel* de 3,5nm. Aussi, pour ces conditions de rétention, on peut conclure que la perte de charges s'effectue uniquement

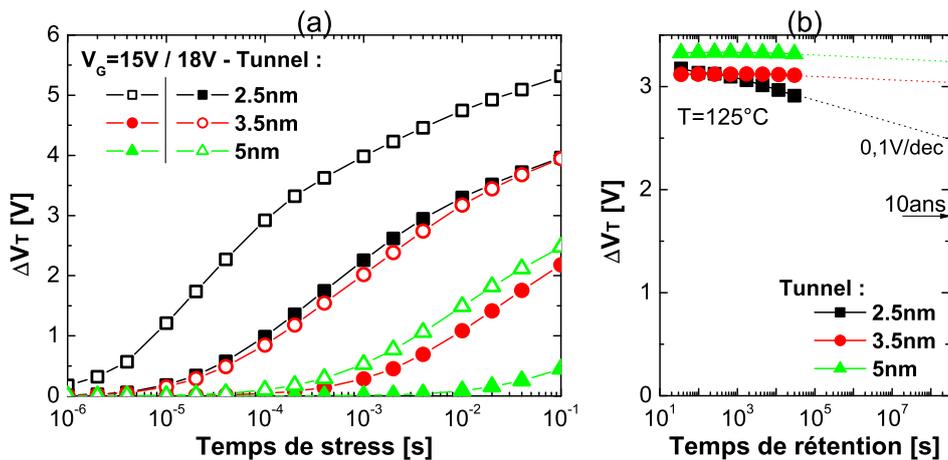


FIG. 4.12: a) Caractéristiques de programmation $\Delta V_T - T_{\text{Stress}}$ et b) caractéristiques de rétention à 125°C pour des SONOS intégrant différentes épaisseurs d'oxyde *tunnel* : 2,5nm ($V_{T\text{init}} = 0,7\text{V}$), 3,5nm ($V_{T\text{init}} = 0,9\text{V}$) et 5nm ($V_{T\text{init}} = 1,2\text{V}$). $HTO=12,5\text{nm}$ et $t_{\text{Si}_3\text{N}_4} = 6\text{nm}$.

par l'oxyde *tunnel* et que l'on peut considérer comme négligeable les fuites dans le *HTO*. Cela nous permettra, dans la partie 4.2.2, de détecter les fuites pour d'autres oxydes bloquants. En conclusion les oxydes *tunnel* de 3-3,5nm se présentent comme de bons compromis entre écriture et rétention.

4.2.2 Impact de l'oxyde bloquant

Dans cette section, nous nous intéressons à l'impact de l'oxyde bloquant sur les performances des mémoires de type SONOS. L'objectif est d'évaluer l'influence de l' Al_2O_3 (mémoires SANOS) par rapport au *HTO* (mémoires SONOS) sur les caractéristiques mémoires. Nous nous sommes aussi intéressés aux empilements bicouches *HTO/Al₂O₃* (mémoires SAONOS) permettant d'améliorer la rétention [Bre08].

La figure 4.13 regroupe les schémas des différents dispositifs étudiés dans cette partie. Des transistors de grandes dimensions ($W=L=1\mu m \rightarrow 10\mu m$) ont été fabriqués avec plusieurs types d'empilements de grille. L'oxyde *tunnel* fait 2,5nm. La couche de piégeage en Si_3N_4 de 6nm a été déposée en LPCVD à 625 ° C. Différents oxydes bloquants ont été intégrés :

- Un *HTO* de 12,5nm.
- Un Al_2O_3 de 16nm
- Un bicouche Al_2O_3 -8nm/*HTO*-4nm

Enfin, une grille en polysilicium N+ a été déposée.

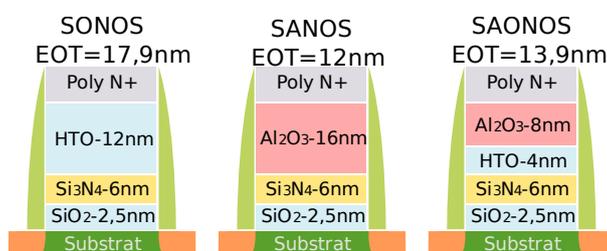


FIG. 4.13: Schémas des dispositifs étudiés pour évaluer l'impact de l'oxyde bloquant. L'EOT indiquée correspond à une mesure électrique de tout l'empilement mémoire.

4.2.2.1 Caractéristiques d'écriture et d'effacement

Dans cette partie, nous allons étudier l'influence de l'oxyde bloquant sur les dynamiques d'écriture et d'effacement. L'ensemble des dispositifs a été programmé par Fowler-Nordheim. La figure 4.14 regroupe les réseaux des caractéristiques de programmation de nos empilements. La comparaison de leurs performances d'écriture est effectuée sur la figure 4.15a. On observe que la diminution de l'EOT de l'oxyde bloquant augmente la vitesse de programmation. La figure 4.15b confronte les caractéristiques d'effacement. Il apparaît que le point mémoire utilisant un *HTO* comme oxyde bloquant ne semble pas

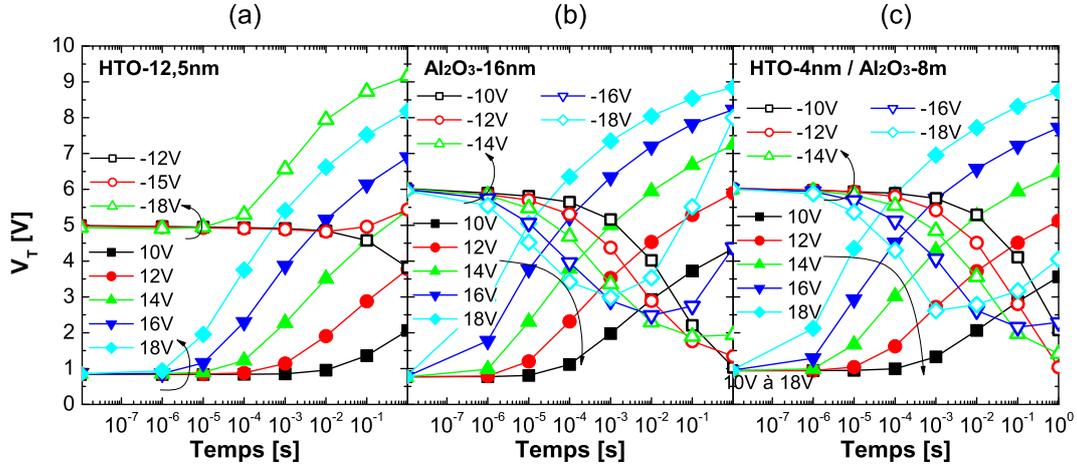


FIG. 4.14: Réseaux des caractéristiques d'écriture et d'effacement, $V_T - Temps_{Stress}$, a) de la SONOS (12,5nm/6nm/2,5nm), b) de la SANOS (16nm/6nm/2,5nm) et c) de la SAONOS (8nm/4nm/6nm/2,5nm).

pouvoir s'effacer. Cela est sans doute dû à son faible coefficient de couplage générant un champ électrique très similaire dans l'oxyde *tunnel* et l'oxyde bloquant, donnant des courants d'électrons entrant et sortant très similaires, qui empêchent l'effacement. Le monocouche d' Al_2O_3 présente un effacement plus rapide que le bicouche Al_2O_3/HTO . De la même façon, cela peut être attribué au meilleur couplage de l'alumine (l'EOT étant plus faible) entre le nitrure et la grille augmentant le champ électrique dans l'oxyde *tunnel*. Néanmoins, on observe la saturation de l'effacement pour des temps plus courts.

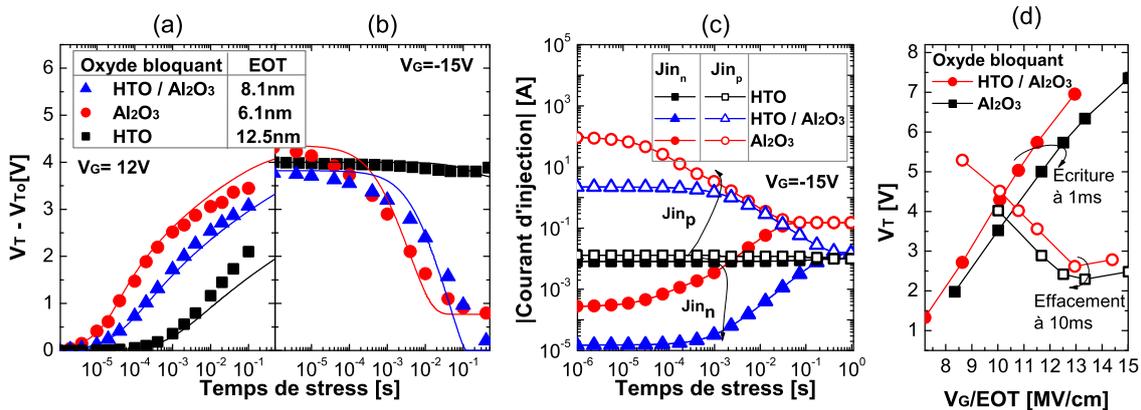


FIG. 4.15: Caractéristiques a) de programmation et b) d'effacement $\Delta V_T - Temps_{Stress}$ pour différents types d'oxyde bloquant : Al_2O_3 , HTO et HTO/ Al_2O_3 . c) Simulation des courants d'effacement (électrons venant de la grille de contrôle et trous venant du canal) pour les différents types d'oxyde bloquant. d) Caractéristiques $V_T - V_G/EOT$ [Boc08].

Pour mieux comprendre les mécanismes physiques impactant la programmation de nos dispositifs, nous avons utilisé le modèle présenté dans le chapitre 2 (cf paragraphe 2.3.2) sur nos caractéristiques d'écriture et d'effacement. En ajustant les paramètres (cf tableau 4.1), on obtient une bonne corrélation entre mesure expérimentale et simulation en écriture et en effacement. Pour l'écriture, la diminution de l'EOT de l'oxyde bloquant, avec l'intégration de l' Al_2O_3 , augmente le champ dans l'oxyde *tunnel* ainsi que le courant d'injection, donnant une cinétique d'écriture plus rapide. Pour l'effacement, nous allons détailler un peu plus.

La figure 4.15c représente les courants dans la structure simulée (cf figure 4.16) :

- J_{in_p} étant le courant de trous venant du canal et permettant l'effacement.
- J_{in_n} étant le courant d'électrons injectés depuis de la grille qui compense l'effacement (courant d'*Electron Back Tunnelling*).

On peut observer que pour une mémoire SO-NOS standard, le courant de trous est très faible, ce qui rend l'effacement très lent. De plus, comme le courant parasite d'électrons est proche de celui des trous, le mécanisme d'effacement sature très tôt⁹.

Dans le cas d'un oxyde bloquant à base d'alumine, on observe que la structure monocouche d' Al_2O_3 donne un courant de trous plus important que celle avec le bicouche HTO/Al_2O_3 entraînant une meilleure efficacité d'effacement. Ceci est principalement dû à un champ plus important dans l'oxyde *tunnel* grâce à une EOT plus faible. Néanmoins, le courant d'électrons venant de la grille, *Electron Back Tunnelling*, est lui aussi plus fort dans le cas du monocouche Al_2O_3 , provoquant une saturation plus rapide de l'effacement. Dans le cas du bicouche HTO/Al_2O_3 , le plus faible courant d'électrons, venant de la grille, est principalement dû à la transparence plus faible vue par les électrons de la grille. Cela s'explique par son EOT plus forte et à sa structure de bande en crête (cf figure 4.17).

La figure 4.15d représente le V_T en fonction de V_G/EOT après $1ms$ d'écriture et $10ms$ d'effacement. On peut assimiler cela à l'efficacité de l'écriture et de l'effacement entre bicouche et monocouche. En écriture, elle est meilleure pour le bicouche malgré un champ

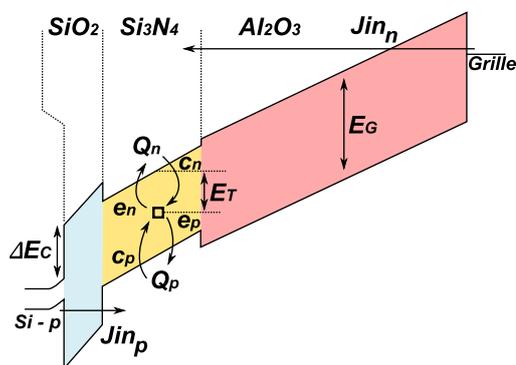


FIG. 4.16: Diagramme de bandes d'une mémoire SANOS et des mécanismes pris en compte pour la simulation.

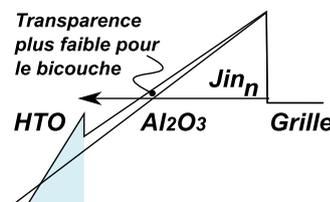


FIG. 4.17: Diagramme de bandes côté grille pour un diélectrique Al_2O_3 ou HTO/Al_2O_3 pour $V_G = -15V$. On observe une transparence plus faible pour l'oxyde bloquant bicouche.

⁹Le mécanisme effacement sature lorsque les courants de trous injectés depuis le canal et d'électrons venant de la grille deviennent égaux.

	<i>SiO₂</i>	<i>Si₃N₄</i>	<i>HTO</i>	<i>Al₂O₃</i>
Gap, E_G [eV]	8,5	5,1	9	6,4
ΔE_C [eV]	3,15	2	2,8	2,3
Permittivité, ϵ_r	3,9	8	4	9
Masse effective des électrons, m_e	0,5	0,5	0,4	0,4
Masse effective des trous, m_h	0,7	0,5	0,4	0,2
Densité de pièges, ρ [cm^{-2}]	$1,8 \cdot 10^{13}$			
Profondeur des pièges, ΔE_T [eV]	1,8			

TAB. 4.1: Paramètres utilisés dans le modèle pour l'approximation des résultats présentés sur la figure 4.15.

d'injection similaire. Nous avons vérifié que ce décalage n'est pas dû aux différents V_{T0} entre le dispositif bicouche ($V_{T0} = 0,94V$) et le dispositif monocouche ($V_{T0} = 0,77V$). Il est possible d'attribuer cela à une meilleure isolation de l'oxyde bloquant grâce à la plus grande hauteur de barrière du *HTO* par rapport à l'alumine. Cette meilleure isolation permettrait de maintenir les électrons injectés dans le nitrure et d'éviter qu'ils soient évacués vers la grille de contrôle.

En effacement, le monocouche d' Al_2O_3 présente une meilleure efficacité. La figure 4.18 met en évidence que cela ne peut pas être expliqué par le champ dans l'oxyde *tunnel*. Il est possible que le même ΔV_T nécessite plus de charges dans la couche de piégeage dans le cas du monocouche, dû à l'EOT de l'oxyde bloquant plus faible. Il serait plus facile d'évacuer ce plus grand nombre de charges, d'où un effacement plus efficace (cf équation 2.40).

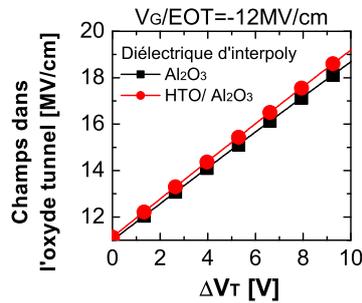


FIG. 4.18: Champs dans l'oxyde lors de l'effacement en fonction du ΔV_T pour différents oxydes bloquants à $V_G/EOT = -12V$.

4.2.2.2 Caractéristiques d'endurance

La figure 4.19a compare les caractéristiques d'endurance pour les différents types d'oxydes bloquants : Al_2O_3 -16nm ou bicouche HTO -4nm/ Al_2O_3 -8nm. On observe une importante fermeture de la fenêtre de programmation de la SANOS qui n'apparaît pas pour la SAONOS. Comme le montrent les $I_D - V_G$ de la figure 4.19b, cela peut être attribué à une forte dégradation de la pente sous le seuil des dispositifs avec monocouche en Al_2O_3 . La dégradation de la pente sous le seuil est principalement due à une dégradation de l'interface Si/SiO_2 . Les principales causes sont :

- Une grande quantité d'électrons traversant l'oxyde *tunnel*.
- Des électrons fortement énergétiques dans le canal ou qui traversent l'oxyde *tunnel*.

Comme nous l'avons vu dans la partie 4.2.2.1, les oxydes bloquants en Al_2O_3 présentent un coefficient de couplage plus fort (EOT de l'oxyde bloquant plus faible). Cela induit des champs plus forts dans l'oxyde *tunnel*¹⁰ et nécessite une plus grande quantité de charges injectées pour obtenir le même ΔV_T . Ces contraintes supplémentaires peuvent expliquer une plus grande dégradation de l'interface Si/SiO_2 dans le cas des oxydes bloquants en Al_2O_3 .

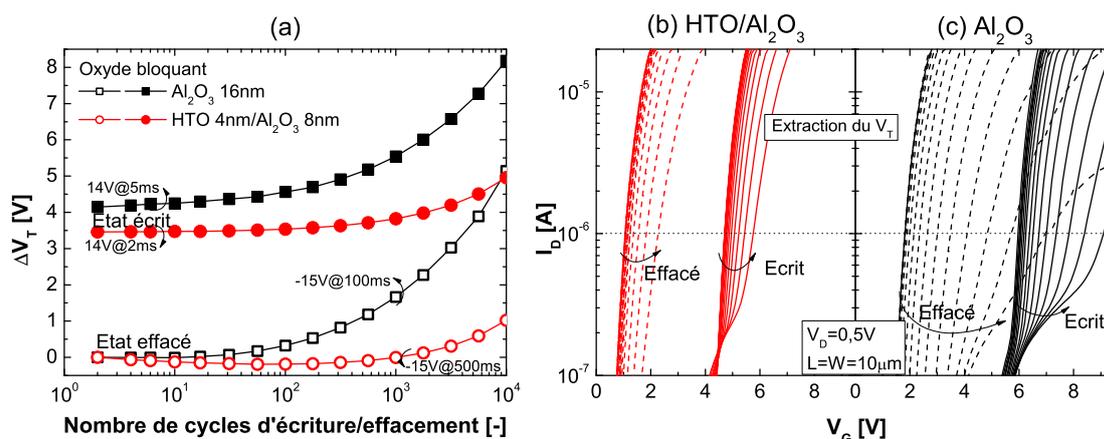


FIG. 4.19: (a) Caractéristiques d'endurance $V_T - nb_{cycle}$ pour différents types d'oxydes bloquants : Al_2O_3 et HTO/Al_2O_3 . Évolution des caractéristiques $I_D - V_G$ lors du cyclage des dispositifs intégrant (b) du HTO/Al_2O_3 ou (c) de l' Al_2O_3 en oxyde bloquant [Boc08].

¹⁰Pour un $\Delta V_T = 0V$, le champ dans l'oxyde *tunnel* est pour l'oxyde bloquant en :

- Al_2O_3 , de $\sim 11,1MV/cm$ pour l'écriture et de $\sim 11,5MV/cm$ pour l'effacement.
- HTO/Al_2O_3 , de $\sim 9,7MV/cm$ pour l'écriture et de $\sim 10MV/cm$ pour l'effacement.

4.2.2.3 Caractéristiques de rétention

Les figures 4.20a-c regroupent les caractéristiques de rétention pour les différents oxydes bloquants et pour des températures entre 25 °C à 200 °C. Les résultats de la partie 4.2.1 nous indiquent que la pente de 0,09V/dec de la cellule mémoire SONOS à 25 °C, représente la fuite à travers l'oxyde *tunnel*. Aussi, l'augmentation de la pente pour les SAONOS (0,11V/dec) et pour les SANOS (0,17V/dec) peut être attribuée aux fuites à travers l'oxyde bloquant. De plus, la figure 4.20d met en évidence la faible activation en température des SAONOS par rapport aux SANOS. Cela peut s'expliquer par l'activation en température de la conduction à travers l' Al_2O_3 (cf figure 3.46). Le *HTO* jouerait le rôle de barrière pour les électrons. On peut donc en conclure que la barrière supplémentaire apportée par le *HTO*, dans le cas du bicouche, permet une amélioration notable de la rétention.

Pour décorrélérer l'effet de l'EOT et l'impact du *HTO* sur la rétention, nous avons caractérisé la rétention de trois dispositifs :

- Un dispositif intégrant un oxyde bloquant bicouche SiO_2 -4nm/ Al_2O_3 -8nm, avec une EOT de 7,3nm.
- Deux dispositifs intégrant des oxydes bloquants en Al_2O_3 de 16nm et 20nm, permettant d'obtenir des EOT de ces couches comprises entre celles du bicouche.

La figure 4.21 montre que les deux mémoires intégrant un oxyde bloquant en alumine possèdent une moins bonne rétention de l'information que celle intégrant un bicouche, même avec une EOT de l'oxyde bloquant plus importante. Ainsi, l'utilisation d'un bicouche *HTO*/ Al_2O_3 permet une amélioration de la rétention.

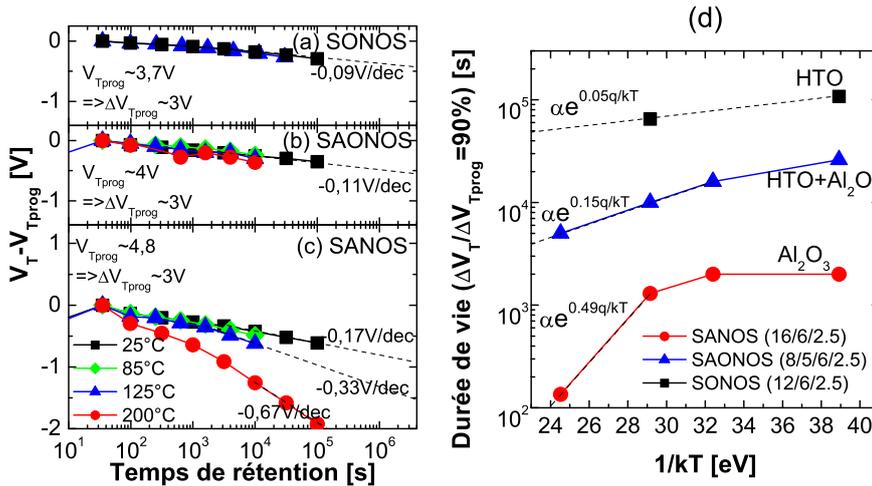


FIG. 4.20: Caractéristiques de rétention pour (a) un oxyde bloquant *HTO*-12,5nm ($V_{Tprog} = 3,6V$), (b) un oxyde bloquant bicouche *HTO*-5nm/ Al_2O_3 -8nm ($V_{Tprog} = 4,0V$) et (c) pour un oxyde en Al_2O_3 -16nm ($V_{Tprog} = 4,6V$). (d) Courbes d'activation en température pour ces différents oxydes bloquants ($\Delta V_{Tprog} \sim 3V$) [Boc08].

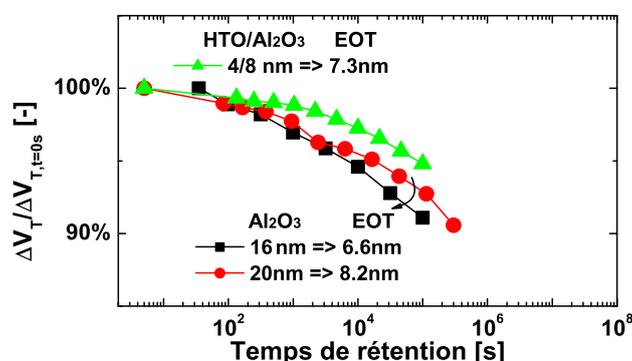


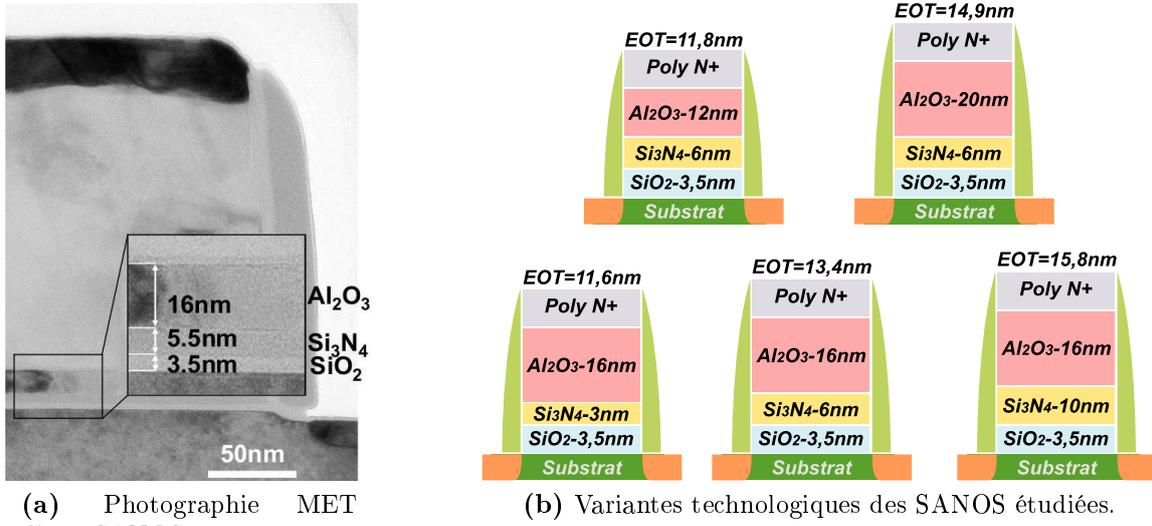
FIG. 4.21: Caractéristiques de rétention pour différents oxydes bloquants : un oxyde en Al_2O_3 -16nm, un oxyde en Al_2O_3 -20nm et un bicouche HTO -5nm/ Al_2O_3 -8nm. L'épaisseur de l'oxyde *tunnel* est de 3,5nm et celle du nitrure est de 6nm.



En conclusion, il apparaît que l'utilisation d'oxyde bloquant à base d' Al_2O_3 est intéressante, aux mémoires de type SONOS, pour assurer de bonnes performances mémoires. L'utilisation d'un bicouche HTO -5nm/ Al_2O_3 -8nm permet une importante amélioration des caractéristiques de rétention par rapport à une mémoire SANOS grâce au HTO jouant le rôle de barrière pour les électrons. De plus, en adaptant bien le bicouche à la structure globale de la mémoire, la SAONOS présente une bonne efficacité de piégeage et ainsi de bonnes caractéristiques d'endurance.

4.2.3 Impact de l'épaisseur de l'alumine et du nitrure

Dans cette partie nous allons nous intéresser à l'impact des épaisseurs de l'alumine et du nitrure sur les caractéristiques d'écriture, d'effacement, d'endurance et de rétention dans des mémoires SANOS. Aussi des transistors ont été réalisés, avec différents empilements de grille. L'oxyde *tunnel* mesure 3,5nm, proche des épaisseurs typiques des TANOS [Lee06]. Trois épaisseurs de nitrure (3nm, 6nm and 10nm) ont été déposées par LPCDV. L'oxyde bloquant, en Al_2O_3 , a été déposé par ALD (cf partie 3.2.1.1) à 350 ° C puis recuit à 700 ° C pendant 15min sous atmosphère O_2 . Trois épaisseurs ont été retenues : 12nm, 16nm et 20nm. Enfin, la grille de contrôle a été réalisée en polysilicium N+. La figure 4.22a est la photographie HR-TEM d'un empilement de cette étude. La figure 4.22b regroupe les différentes variantes technologiques que nous allons étudier.



(a) Photographie MET d'une SANOS.

(b) Variantes technologiques des SANOS étudiées.

FIG. 4.22: Photographies MET et variantes technologiques des transistors SANOS [Boc09].

4.2.3.1 Caractéristiques d'écriture et d'effacement

La figure 4.23 regroupe l'ensemble des réseaux des caractéristiques d'écriture et d'effacement de ces empilements SANOS sur lesquels nous allons nous appuyer pour réaliser cette étude. Les figures 4.24a&b comparent les caractéristiques d'écriture et d'effacement pour trois épaisseurs de nitrure. Il apparaît que les vitesses de programmation augmentent avec la diminution de l'épaisseur du nitrure. Cette observation est directement liée à la réduction de l'EOT de l'empilement de grille augmentant le champ d'injection.

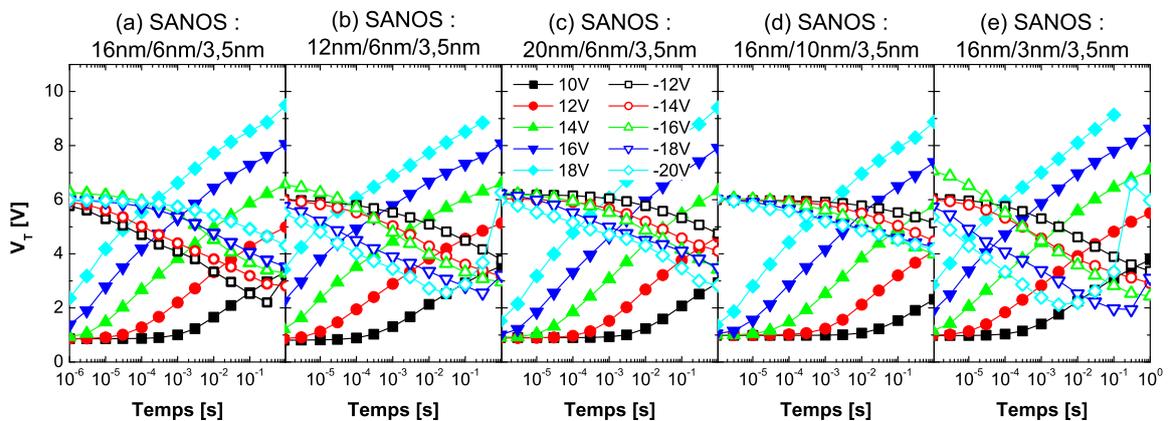


FIG. 4.23: Réseaux des caractéristiques d'écriture et d'effacement, $V_T - T_{\text{Temp}_{\text{Stress}}}$, d'empilements SANOS pour différentes épaisseurs d'alumine et de nitrure. $t_{\text{SiO}_2} = 3,5\text{nm}$.

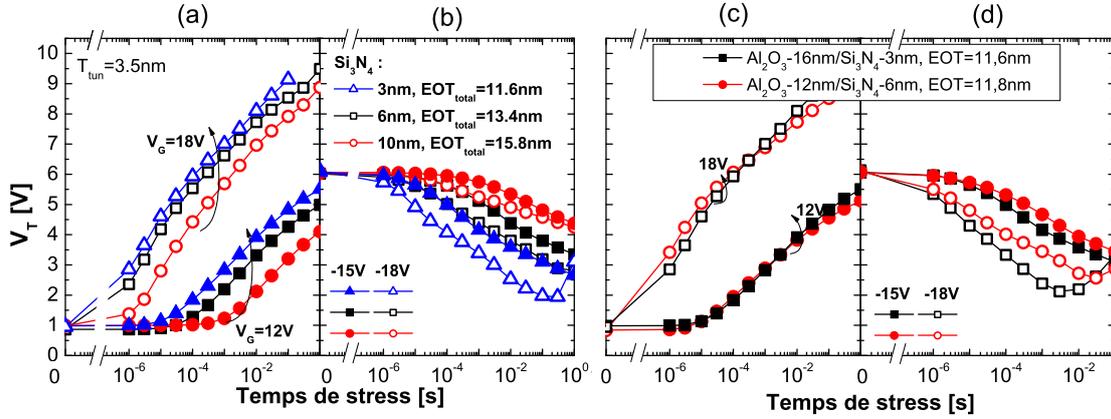


FIG. 4.24: Caractéristiques (a) de programmation et (b) d'effacement $V_T - Temps_{Stress}$ pour différentes épaisseurs de nitrure dans un empilement SANOS ($t_{Al_2O_3} = 16nm$ et $t_{SiO_2} = 3,5nm$). Caractéristiques (c) de programmation et (d) d'effacement $V_T - Temps_{Stress}$ pour différentes épaisseurs de nitrure et d'alumine dans un empilement SANOS ayant la même EOT [Boc09].

Aussi, pour décorrélérer l'impact de l'EOT et de l'épaisseur du nitrure, nous avons comparé deux dispositifs ayant la même EOT, mais des épaisseurs de nitrure et d'alumine différents. On observe, sur la figure 4.24c, la même caractéristique d'écriture. Cela indiquerait que l'écriture est contrôlée par le champ électrique dans l'oxyde *tunnel*. Néanmoins, il est possible que les effets dus aux variations des épaisseurs du nitrure et de l'alumine se compensent : la diminution du Si_3N_4 augmentant le courant de fuite à travers l'oxyde bloquant (cf équation 2.11) et l'augmentation de l' Al_2O_3 réduisant celui-ci.

Sur la figure 4.24d, on observe que l'effacement ne se normalise pas avec l'EOT : le nitrure le plus fin s'effaçant le plus vite. Le calcul des champs dans la structure, avec des positions et des valeurs de charges différentes, n'a pas pu expliquer ce comportement¹¹. Cela indique des mécanismes d'effacement différents en fonction de l'épaisseur du nitrure.

Comme nous l'avons vu dans le chapitre 2 (partie 2.2.4), l'effacement peut être exprimé par :

- Une émission d'électrons provenant du nitrure, équation 4.1.
- Une capture de trous provenant du substrat, équation 4.2.

$$J_{en} = q \cdot e_n \cdot f_T \cdot N_T \quad (4.1)$$

$$J_{cp} = q \cdot c_p \cdot f_T \cdot N_T \quad (4.2)$$

- f_T [-] : Fonction d'occupation des pièges.
- N_T [m^{-2}] : Densité de pièges.
- e_n [1/s] : Coefficients d'émission.
- c_p [1/s] : Coefficients de capture.

¹¹Les champs étaient soit très proches, soit ne correspondaient pas à la tendance observée. On notera aussi que l'injection à travers l'alumine depuis la grille est de type Fowler-Nordheim, nous indiquant que l'épaisseur de l' Al_2O_3 ne doit pas perturber nos conclusions.

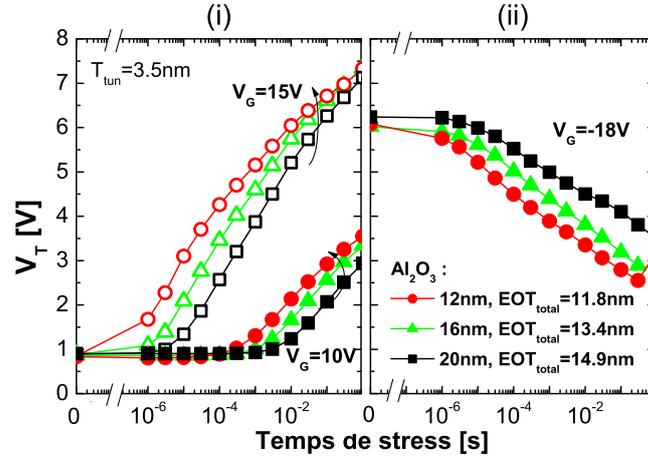


FIG. 4.25: Caractéristiques $V_T - Temps_{Stress}$ (i) d'écriture et (ii) d'effacement pour différentes épaisseurs d'alumine V_G ($t_{Si_3N_4} = 6nm$ et $t_{SiO_2} = 3,5nm$).

Pour un ΔV_T donné, il faut une densité volumique de charges piégées plus importante pour une couche de piégeage fine, aussi le taux d'occupation (f_T) est plus grand¹². Dans ces conditions, le piégeage d'électrons et la capture de trous sont plus importants pour le nitrure fin d'où une cinétique d'effacement plus rapide.

La figure 4.25 regroupe les caractéristiques d'écriture et d'effacement de SANOS ayant différentes épaisseurs d'alumine. La réduction de l'EOT, due à la diminution de l' Al_2O_3 , permet une cinétique de programmation plus rapide grâce à l'augmentation du champ d'injection. Il est intéressant d'observer le pincement des caractéristiques $V_T - Temps_{Stress}$ pour les forts V_G et ΔV_T . Cela est le signe d'une moins bonne efficacité de piégeage pour les alumines plus fines.

Pour aller plus loin dans l'analyse de l'impact des épaisseurs du nitrure et de l'alumine, nous avons tracé le V_T en fonction du champ appliqué : V_G/EOT , après $1ms$ pour l'écriture et $10ms$ pour l'effacement, cf figure 4.26. En écriture on observe que le nitrure de $10nm$ et l'alumine de $20nm$ permettent d'obtenir des V_T plus importants à même V_G/EOT . Cela soulignerait que l'efficacité de piégeage¹³ a diminué avec la réduction de l'épaisseur du nitrure ou de l'alumine. En effet, d'après l'équation du courant de fuite à travers l'oxyde bloquant (cf équation 2.11), l'augmentation de l'épaisseur du nitrure diminuerait la densité de charge sur la bande de conduction du nitrure et réduirait ce courant de fuite. De même l'augmentation de l'épaisseur de l'alumine réduit la transparence de l'oxyde bloquant et diminue la quantité de charges piégées pour avoir le même ΔV_T réduisant ainsi le courant de fuite.

¹²Si on considère la même densité de pièges.

¹³C'est-à-dire : sur l'ensemble des charges qui sont injectées dans l'empilement, combien sont réellement piégées.

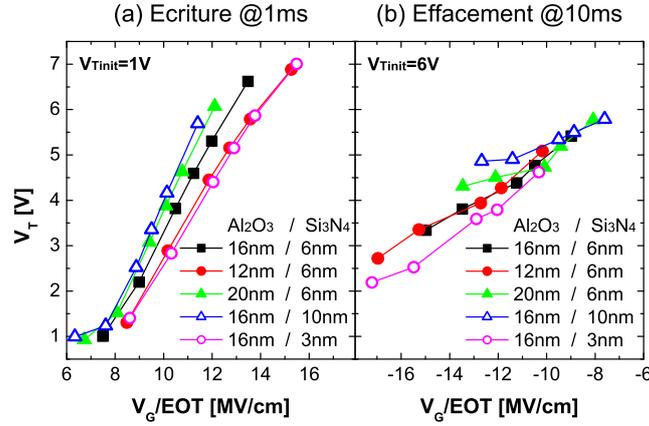


FIG. 4.26: Caractéristiques $V_T - V_G/EOT$ (a) de programmation après $1ms$ et (b) d'effacement après $10ms$ pour des empilements SANOS, intégrant plusieurs épaisseurs de nitrure et d'alumine ($t_{SiO_2} = 3, 5nm$).

De même, en effacement la diminution de l'épaisseur du nitrure favorise l'émission des électrons piégés. Il s'agit du phénomène expliqué plus haut (équation 4.1). Il est intéressant d'observer que l'alumine de $12nm$ ne semble pas dégrader l'effacement.

Ces observations peuvent être mises en regard des résultats présentés dans [Via09a]. Il s'agit d'une confrontation entre les résultats électriques que nous avons obtenus sur nos dispositifs et des simulations d'écriture¹⁴. Ces auteurs s'intéressent à l'efficacité de piégeage de nos mémoires. Pour quantifier cela, ils la définissent comme :

$$\frac{J_{in}^{e-} - J_{out}^{e-}}{J_{in}^{e-}}$$

J_{in}^{e-} : Courant simulé d'électrons entrant dans la couche de piégeage.

J_{out}^{e-} : Courant simulé d'électrons sortant dans la couche de piégeage.

Ainsi, la figure 4.27 regroupe les efficacités de piégeage pour différents V_G , pour différentes épaisseurs de nitrure et pour différentes épaisseurs d'alumine, à même champ V_G/EOT . On observe les mêmes tendances que sur la figure 4.26a :

- La diminution de l'épaisseur du nitrure dégrade grandement l'efficacité de piégeage. Les auteurs de [Via09a] l'expliquent par un trajet plus long des électrons dans le nitrure, leur donnant plus de temps pour se piéger et diminuant leur densité à l'interface SiO_2/Al_2O_3 ; d'où un courant de fuite plus faible.
- La diminution de l'épaisseur de l'alumine ne modifie pas l'efficacité de piégeage à $\Delta V_T = 0V$. Néanmoins, la présence de charges dans l'empilement réduit l'efficacité de piégeage pour l' Al_2O_3 fin. Cela provient de la répartition des champs différents dans la structure et la nécessité d'avoir une densité de charges plus importante pour avoir le même ΔV_T avec la diminution de l'épaisseur de l'oxyde bloquant.

¹⁴Modèle présenté dans la partie 2.3.3 et les paramètres du modèle regroupés dans le tableau 4.2.

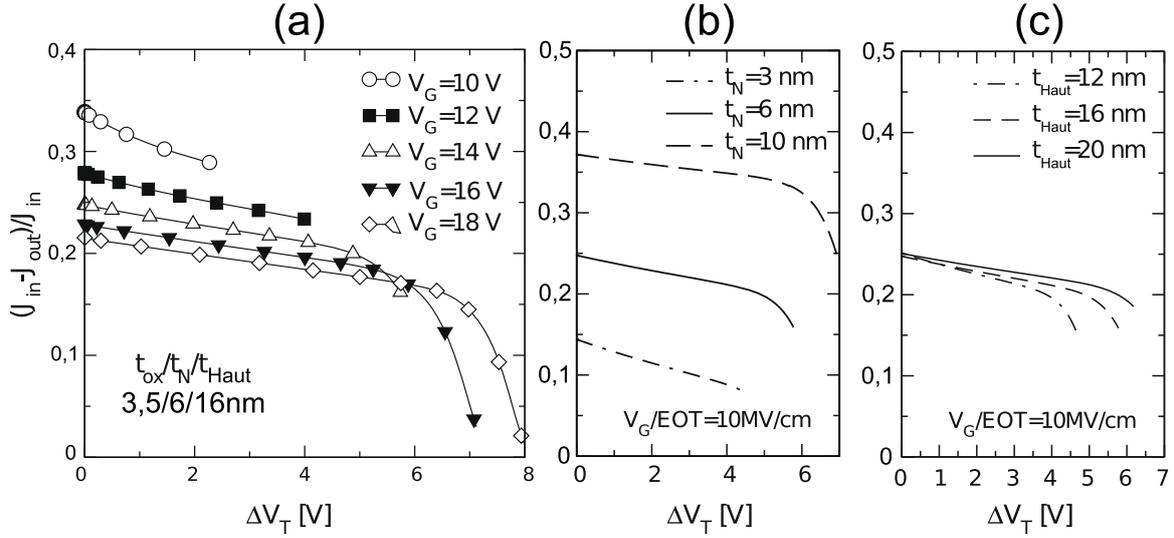


FIG. 4.27: Efficacité du piégeage ($\frac{J_{in}^e - J_{out}^e}{J_{in}^e}$) simulée d'un empilement $Al_2O_3/Si_3N_4/SiO_2$ (t_{tun} 3, 5nm) a) pour différentes tensions, b) pour différentes épaisseurs de Si_3N_4 (3 – 6 – 10nm) et c) de Al_2O_3 (12 – 16 – 20nm) à $V_G/EOT = 10MV/cm$ [Via09a].



En conclusion, le mécanisme dominant pour la programmation et l'effacement reste le champ¹⁵ dans l'oxyde *tunnel*. Mais contrairement aux mémoires à grille flottante, le courant de fuite à travers l'oxyde bloquant modifie grandement la cinétique d'écriture des dispositifs SANOS. De même, la diminution de l'épaisseur du nitrure favorise l'émission des électrons piégés et ainsi l'effacement de la cellule. Mais cette réduction diminue l'efficacité de piégeage ce qui peut être néfaste pour l'endurance¹⁶ et pour la rétention¹⁷.

4.2.3.2 Caractéristiques d'endurance

Comme nous venons de voir, que la diminution des épaisseurs de l'alumine et du nitrure réduisait l'efficacité de piégeage des mémoires SANOS, elle devrait impacter les propriétés d'endurance des mémoires. En effet, dans le cas d'une efficacité réduite, il faut injecter une grande quantité de charges dans la couche de piégeage pour n'en garder qu'une partie. L'augmentation du courant d'injection pour avoir le même ΔV_T va dégrader plus rapidement l'oxyde *tunnel* et l'interface Si/SiO_2 causant une mauvaise endurance [Van09].

Ainsi, pour mettre en évidence cet effet, nous avons choisi d'appliquer les mêmes conditions d'injection pour nos différents dispositifs :

- En écriture : $V_G/EOT = 12MV/cm$ durant 10ms.
- En effacement : $V_G/EOT = -12MV/cm$ durant 30ms.

¹⁵Donc le courant.

¹⁶La programmation d'un ΔV_T demandant un courant injectant plus important.

¹⁷En favorisant l'émission des électrons pour l'effacement, on pénalise la rétention.

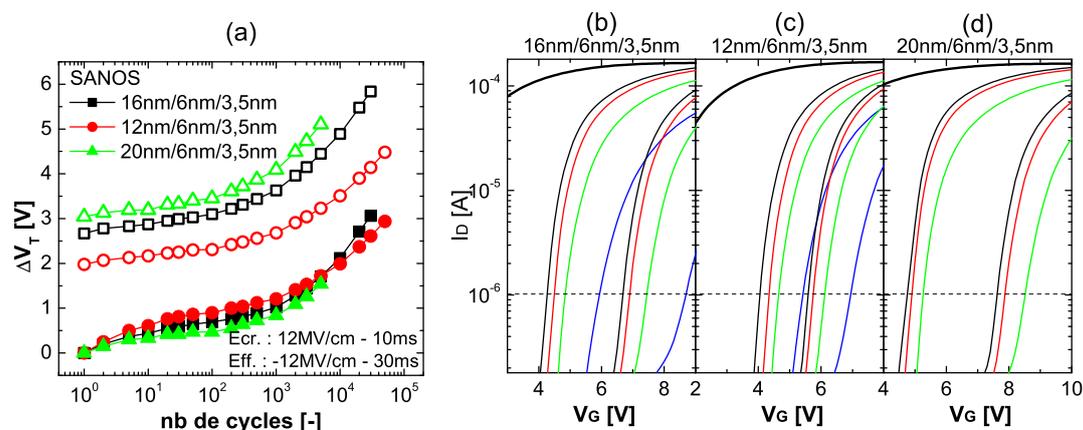


FIG. 4.28: a) Caractéristiques d'endurance $\Delta V_T - nb$ cycles pour des empilements SANOS intégrant différentes épaisseurs d' Al_2O_3 : 12–16–20nm. Réseaux de caractéristiques $I_D - V_G$ lors de l'endurance pour b) $t_{Al_2O_3} = 16nm$, c) $t_{Al_2O_3} = 12nm$ et d) $t_{Al_2O_3} = 20nm$. V_T effacé au 1^{er} cycle $\sim 3,5V$. $t_{Si_3N_4} = 6nm$, $t_{SiO_2} = 3,5nm$.

La figure 4.28a regroupe les caractéristiques d'endurance pour différentes épaisseurs d'alumine. Comme les conditions sont relativement agressives, on observe une importante dégradation de la pente des $I_D - V_G$, signe d'un dommage important de l'interface Si/SiO_2 . Mais comme les conditions d'injection sont similaires, cette dégradation est la même pour l'ensemble des dispositifs. Par contre le ΔV_T obtenu correspond aux efficacités de piégeage des différents empilements. Ainsi l'alumine la plus fine présente, pour la même dégradation, une fenêtre mémoire plus faible que l'alumine épaisse.

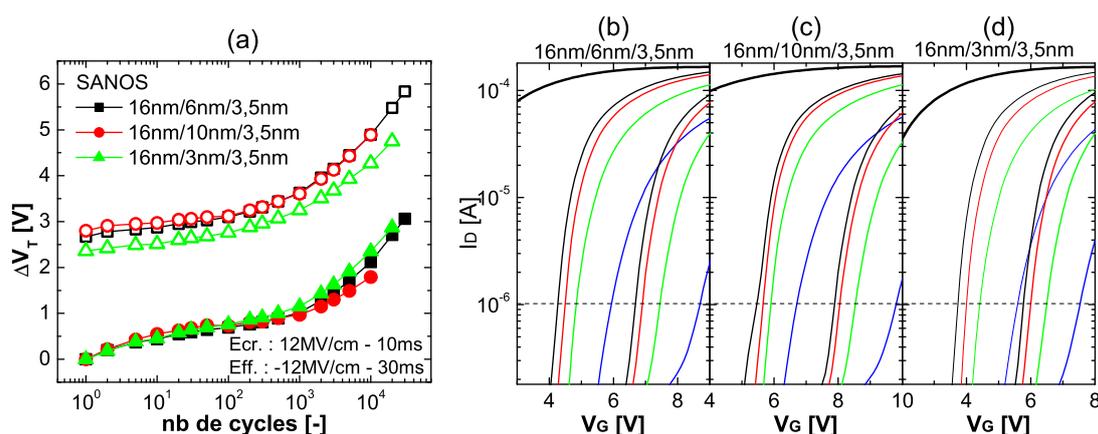


FIG. 4.29: a) Caractéristiques d'endurance $\Delta V_T - nb$ cycles pour des empilements SANOS intégrant différentes épaisseurs de Si_3N_4 : 3–6–10nm. Réseaux de caractéristiques $I_D - V_G$ lors de l'endurance pour b) $t_{Si_3N_4} = 6nm$, c) $t_{Si_3N_4} = 10nm$ et d) $t_{Si_3N_4} = 3nm$. V_T effacé au 1^{er} cycle $\sim 3,5V$. $t_{Al_2O_3} = 16nm$, $t_{SiO_2} = 3,5nm$.

La même étude a été réalisée pour différentes épaisseurs de nitrure, cf figure 4.29. Ainsi, on observe une fenêtre légèrement plus petite pour le nitrure de $3nm$, confirmant une dégradation des performances en endurance avec des dispositifs intégrant des nitrures trop fins.



En conclusion, la dégradation de l'efficacité de piégeage causée par la diminution de l'épaisseur du nitrure ou de l'alumine impacte les caractéristiques d'endurance. Mais le gain notable sur les tensions ou les vitesses d'écriture/effacement dû à la réduction du Si_3N_4 ou de l' Al_2O_3 reste très attractif. L'utilisation d'un nitrure de $4nm$ (au lieu de $6nm$) peut grandement améliorer les caractéristiques de programmation sans trop pénaliser l'endurance. Néanmoins, une trop grande amélioration des vitesses d'effacement se traduit très souvent par une dégradation de la rétention. Il est donc essentiel d'évaluer l'impact de ces épaisseurs sur le maintien de la charge.

4.2.3.3 Caractéristiques de rétention

Nous nous intéressons maintenant à l'impact de l'épaisseur du nitrure sur la rétention [Mel08]. L'objectif est de mieux comprendre les mécanismes de perte de charges. Aussi, nous nous appuyerons sur le modèle de piégeage présenté dans la partie 2.3.3. Il est important de rappeler que, pour ce modèle, l'émission des électrons du piège vers la bande de conduction du nitrure est une activation en température par un mécanisme Poole-Frenkel (cf équation 2.48); et que les électrons sur la bande de conduction du nitrure ont un transport de type drift-diffusion.

De plus, il est important de signaler que les mesures de rétention peuvent être modifiées par la présence ou l'absence du contact de la pointe de mesure sur le plot de grille. En effet :

- L'absence du contact de la pointe donne à la grille un potentiel flottant, non contrôlé, donnant généralement de meilleures rétentions car cela peut réduire le champ dans l'oxyde bloquant ou l'oxyde *tunnel*. Cette condition est similaire à une ligne de mot en haute impédance pour une matrice NAND.
- Le contact de la pointe sur la grille permet de fixer V_G durant la rétention. L'intérêt est de connaître ce potentiel. Néanmoins, les performances sont légèrement moins bonnes que pour un V_G flottant.

Ainsi pour cette étude, il est important de définir le potentiel de grille pour pouvoir simuler correctement la rétention. Aussi, nous avons effectué l'ensemble de ces rétentions en maintenant le contact des pointes sur les plots.¹⁸

¹⁸On notera que les rétentions présentées en dehors de cette partie, si rien n'est précisé, ont été effectuées sans contact des pointes.

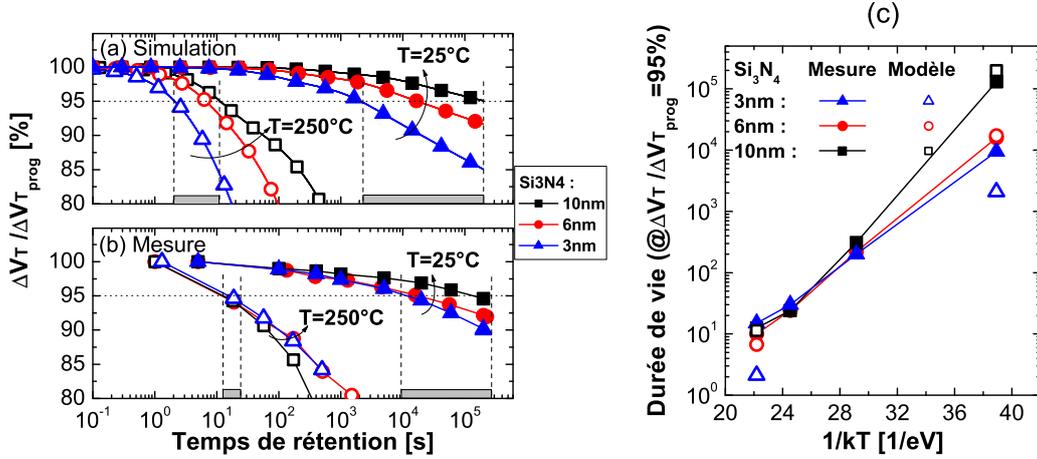


FIG. 4.30: (a) Simulations et (b) mesures expérimentales des caractéristiques de rétention normalisées $\Delta V_T / \Delta V_{T_{prog}}$ pour différentes épaisseurs de nitrure (3nm, 6nm et 10nm) à 25 °C et à 250 °C. (c) Tracés Arrhenius pour différentes épaisseurs de nitrure. $V_{T_{prog}} \sim 6V$ et $\Delta V_{T_{prog}} \sim 5V$ [Boc09].

Les mesures de rétention pour les trois épaisseurs sont présentées sur les figures 4.30b. Le nitrure de 10nm nous donne la meilleure rétention à 25 °C. La figure 4.30c représente les tracés d'Arrhenius de la durée de vie de la charge à 95%¹⁹ en fonction de la température de rétention. Cette figure permet de mettre en évidence l'activation en température de la rétention, souvent appelée énergie d'activation. Elle correspond à la pente dans un tracé d'Arrhenius, soit $\exp\left(\frac{Ea}{kT}\right)$ ²⁰. Elle représente l'énergie qu'il faudrait fournir à un électron piégé pour qu'il soit émis (cf équation 2.47 ou 2.48). Dans ce cas, on observe que l'énergie d'activation croît avec l'augmentation de l'épaisseur du Si_3N_4 , allant de 0,57eV à 0,39eV. Il est intéressant de constater qu'un nitrure plus épais améliore la rétention à température ambiante mais que ce gain est réduit à hautes températures.

	SiO_2	Si_3N_4	Al_2O_3
ΔE_C [eV]	2,9	1,9	2,8
Permittivité, ϵ_r	3,9	7,8	9,3
Masse effective des électrons, m_e	0,5	0,5	0,15
Densité de pièges, ρ [$1/cm^{-3}$]	$1 \cdot 10^{20}$		
Profondeur des pièges, ΔE_T [eV]	$\sim 1,6eV$ cf éq.18 [Via09b]		
Coefficient de piégeage [cm^3/s]	$8 \cdot 10^{-9}$		
Mobilité dans le Si_3N_4 , μ [$cm^2V^{-1}s^{-1}$]	1		

TAB. 4.2: Paramètres utilisés dans les simulations présentées sur les figures 4.30 et 4.31.

¹⁹La durée de vie à 95% est le temps pour lequel la mémoire a perdu 5% de la charge initialement piégée.

²⁰ Ea l'énergie d'activation en [J].

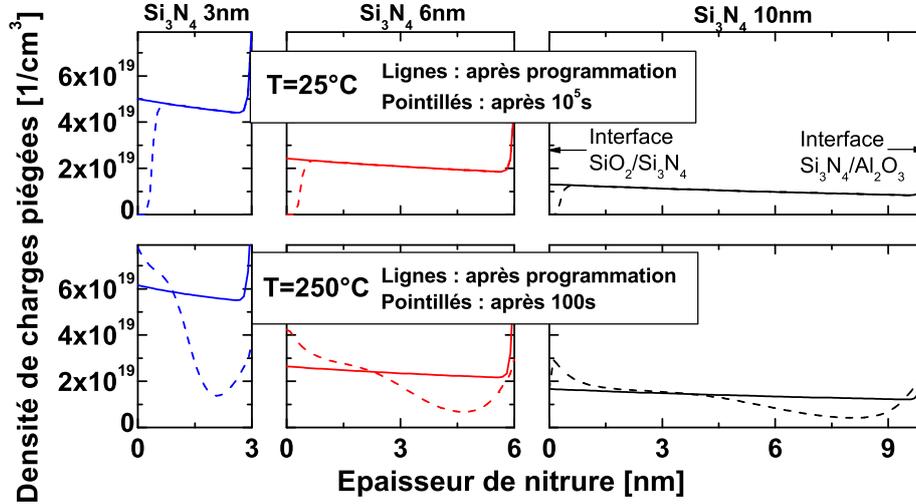


FIG. 4.31: Densité de charges piégées après programmation et durant la rétention, à température ambiante et à 250 °C, pour différentes épaisseurs de nitrure. ($t_{SiO_2} = 3,5nm$ et $t_{Al_2O_3} = 16nm$) [Boc09].

Pour comprendre cet effet, nous avons simulé ces trois structures en rétention (cf figure 4.30a). Les paramètres utilisés sont regroupés dans le tableau 4.2²¹. La répartition initiale de charge impactant grandement la rétention [Arr07], nous avons, dans un premier temps, simulé la programmation de la mémoire. Puis, en partant de la distribution de charge de la programmation, nous réalisons les simulations de rétention (exemple figure 4.32).

Elles nous ont permis d'extraire sur la figure 4.31, la distribution de la charge piégée dans le nitrure après la programmation et durant la rétention. Il apparaît que la densité d'électrons piégés est constante dans tout le nitrure après l'écriture. La valeur de la densité volumique de charge est proportionnelle à l'épaisseur du nitrure : le Si_3N_4 le plus fin donnant la plus forte densité. En effet pour avoir le même ΔV_T , il faut la même quantité de charges surfaciques²² soit approximativement 3,3 fois plus de charges volumiques pour un Si_3N_4 de 3nm que pour un Si_3N_4 de 10nm.

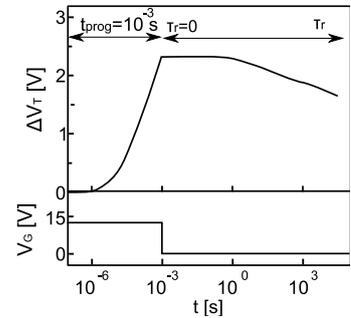


FIG. 4.32: Séquence de simulation pour la programmation et l'effacement [Via09b].

²¹On peut noter que la masse effective utilisée dans ces simulations pour l'alumine est faible ($m_{Al_2O_3} = 0,15m_0$) par rapport aux valeurs de la littérature ($0,28m_0 \sim 0,35m_0$) [Hou03]. Nous pensons que cette faible valeur est le signe de conduction assistées par piège dans l'oxyde bloquant.

²²En négligeant les différentes positions du centroïde de charge $\chi = \frac{\int_0^{t_{Si_3N_4}} x \cdot \rho(x) dx}{\int_0^{t_{Si_3N_4}} \rho(x) dx}$ avec $\rho(x)$ la densité volumique de charges piégées [C/m^3] et $t_{Si_3N_4}$ l'épaisseur de nitrure [m] d'où $\Delta V_T = -Q_T \cdot \left(\frac{\chi}{\epsilon_{Si_3N_4}} + \frac{t_{Al_2O_3}}{\epsilon_{Al_2O_3}} \right)$ [Arr08b].

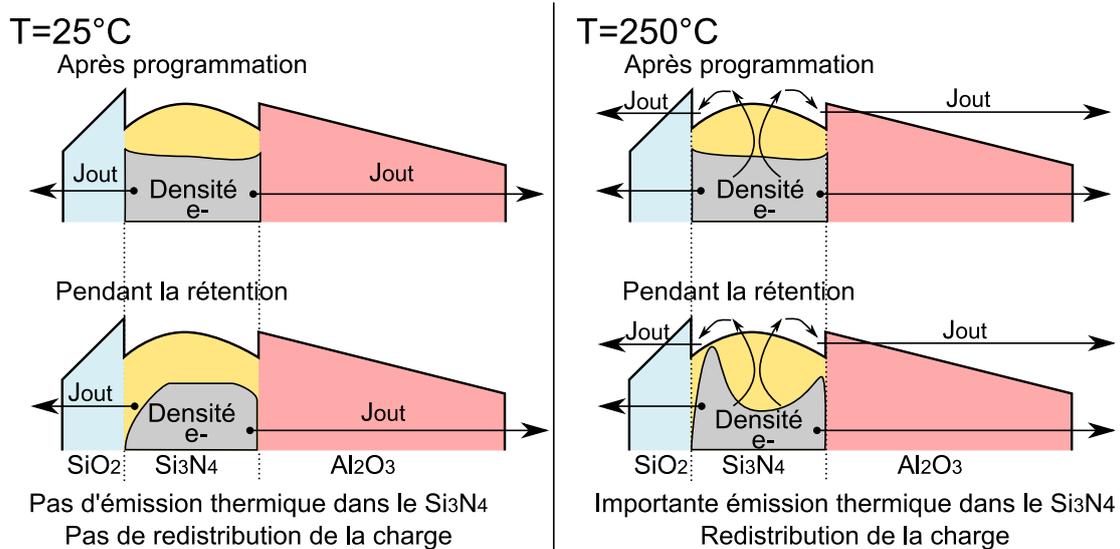


FIG. 4.33: Schémas de bandes et densité de charges piégées lors de la rétention en température. Impact de l'activation en température et la redistribution de la charge. ($t_{\text{SiO}_2} = 3,5\text{nm}$ et $t_{\text{Al}_2\text{O}_3} = 16\text{nm}$) [Boc09].

Nous pouvons comparer l'évolution de la densité de charge durant la rétention à température ambiante et à hautes températures :

- À 25°C , seules les charges aux interfaces $\text{Si}_3\text{N}_4/\text{SiO}_2$ et $\text{Si}_3\text{N}_4/\text{Al}_2\text{O}_3$ sont émises, la quantité de charges piégée dans le volume du nitrure restant constante.
- À 250°C , on observe une importante redistribution de la charge dans le nitrure. En effet, les charges piégées dans le volume migrent vers les interfaces.

Pour expliquer les phénomènes physiques qui interviennent en rétention, nous avons représenté les diagrammes de bandes ainsi que les principaux mécanismes d'émission sur la figure 4.33.

- À 25°C , les électrons piégés dans le volume du nitrure ne sont pas émis thermiquement. Cela explique l'absence de redistribution de charge. De plus, ces électrons ne peuvent pas être injectés directement par effet *tunnel* vers le canal ou la grille. En effet, la transparence vue par ces électrons est trop faible pour permettre ce type de mécanisme. Aussi, seuls les électrons proches des interfaces peuvent être évacués par mécanisme *tunnel*. Or, les mécanismes *tunnel* sont directement proportionnels à la densité de charge disponible. Comme la SANOS avec un Si_3N_4 de 3nm possède ~ 5 fois plus de charges ($5 \cdot 10^{19}\text{cm}^{-3}$) que celle de 10nm (10^{19}cm^{-3}), le courant d'émission est plus important pour les nitrures fins (cf figure 4.34), d'où une moins bonne rétention.
- À 250°C , après l'écriture, il y a redistribution de la charge piégée. Les électrons du volume sont émis sur la bande de conduction du nitrure et migrent rapidement vers les interfaces. Ainsi la densité de charge dans le volume diminue rapidement. Ensuite les électrons situés près des interfaces peuvent être évacués vers le substrat

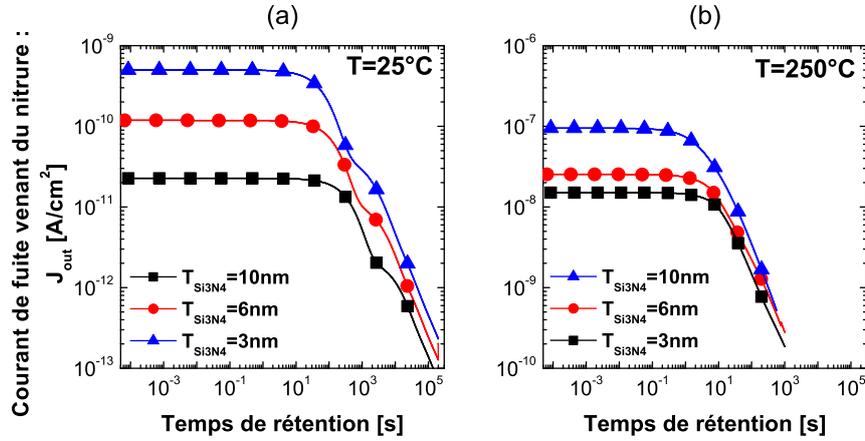


FIG. 4.34: Densité de courant simulée venant du nitrure allant vers le substrat et la grille pour les différentes épaisseurs de nitrure. ($t_{SiO_2} = 3, 5\text{nm}$ et $t_{Al_2O_3} = 16\text{nm}$) [Boc09].

ou la grille. Ainsi, la différence entre les valeurs des courants d'émission des trois épaisseurs de nitrure diminue (cf figure 4.34).

Il est intéressant de remarquer une augmentation de la charge piégée à l'interface Si_3N_4/Al_2O_3 sur la figure 4.31. Elle est causée par une accumulation à cette interface, des électrons sur la bande de conduction, dû au transport dans le nitrure.

En conclusion, à température ambiante, la rétention est contrôlée par l'émission des charges initialement piégées à l'interface Si_3N_4/SiO_2 . Les charges du volume sont conservées. Ainsi, l'augmentation de l'épaisseur du nitrure améliore grandement la rétention à 25°C . À haute température, les charges piégées dans le volume migrent très rapidement aux interfaces et sont émises, on est alors limité par la fuite aux interfaces. Ainsi le bénéfice des nitrures épais²³ diminue.

La figure 4.35 représente l'impact de l'épaisseur de l' Al_2O_3 sur la rétention. La figure 4.35a regroupe les caractéristiques de rétention des empilements SANOS intégrant un oxyde bloquant en alumine de 12nm , 16nm et 20nm pour deux températures : 25°C et 200°C . On remarque une meilleure rétention dans le cas de l'alumine épaisse. La figure 4.35b représente les tracés d'Arrhenius pour une durée de vie de la charge de 95%. On observe une augmentation de l'énergie d'activation avec l'augmentation de l'épaisseur de l'alumine, donnant à haute température, des rétentions relativement proches pour les épaisseurs de 20nm et 16nm .

²³Moins de charges aux interfaces, plus de charges dans le volume.

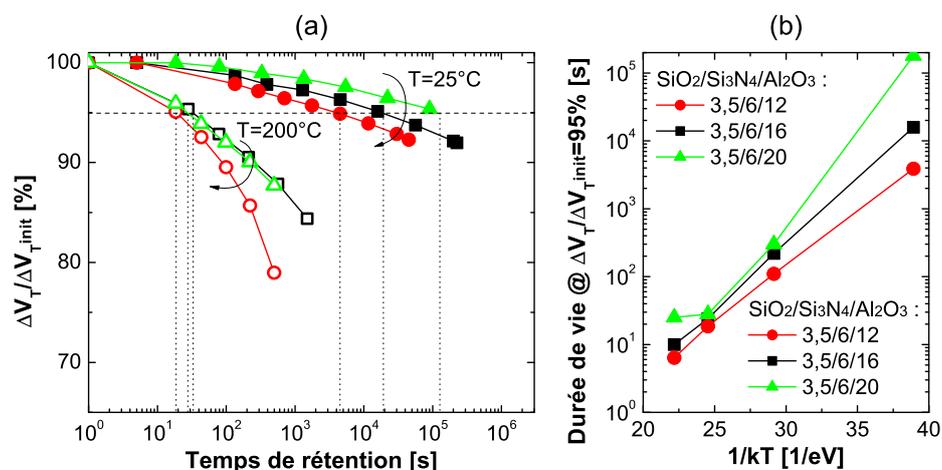


FIG. 4.35: (a) Caractéristiques de rétention normalisées $\Delta V_T / \Delta V_{T_{init}}$ pour différentes épaisseurs d'alumine (12nm, 16nm et 20nm) à 25°C et à 200°C . (b) Tracés Arrhenius pour différentes épaisseurs d'alumine. $V_{T_{prog}} \sim 6\text{V}$ et $\Delta V_{T_{prog}} \sim 5\text{V}$.

La réduction du courant de fuite avec l'augmentation de l'épaisseur de l'alumine est due à plusieurs effets :

- Tout d'abord, l'augmentation de l'épaisseur de l'oxyde bloquant réduit la transparence que voient les électrons qui la traversent, car le trajet dans l'alumine est plus grand.
- Ensuite, une EOT plus grande réduit le champ dans l'alumine.
- Enfin, plus on éloigne la charge piégée de la grille de contrôle, moins il faut piéger d'électrons pour avoir le même ΔV_T (cf équation 1.1). Ainsi, en réduisant la densité de charge piégée, on réduit la probabilité d'émission (cf équation 2.40 et 4.1).

Ainsi l'augmentation de l'épaisseur d'alumine permet d'améliorer la rétention mais reste grandement pénalisante pour les tensions de programmation.



Pour conclure cette partie sur les épaisseurs du nitrure et de l'alumine, nous avons pu mettre en évidence l'augmentation des vitesses d'écriture et d'effacement grâce à la diminution de leur épaisseur à V_G constant. Cette observation est directement liée à la réduction de l'EOT de l'empilement de grille. On retiendra que l'écriture est fortement contrôlée par le champ électrique dans l'oxyde *tunnel* mais que le courant de fuite à travers l'oxyde bloquant impacte lourdement l'ensemble des caractéristiques mémoires. De plus, les nitrures fins présentent une moins bonne efficacité de piégeage, pouvant impacter l'endurance, mais améliorent grandement l'effacement. Ce gain se fait au détriment de la rétention. À température ambiante, le principal mécanisme est l'émission des charges initialement piégées aux interfaces, donnant au nitrure épais une meilleure rétention. En revanche, à haute température, les mesures et les simulations ont montré que les électrons piégés dans le volume sont émis dans la bande de conduction du nitrure puis migrent aux interfaces pour ensuite être évacué par l'oxyde *tunnel* et l'oxyde bloquant. Cela diminue l'impact de l'épaisseur du nitrure. Ainsi, il est capital d'ajuster l'épaisseur du nitrure et de l'alumine pour assurer un bon effacement et une bonne rétention.

4.2.4 Impact des traitements thermiques sur l'alumine

Dans cette partie, nous allons nous intéresser à l'impact des recuits, et de la cristallinité, de la couche d'alumine sur les propriétés électriques d'un empilement mémoire SANOS. Les résultats matériaux sur la couche d'alumine seule sont présentés dans la partie 3.4. Ici, nous allons étudier des empilements capacitifs avec 3,5nm d'oxyde *tunnel* obtenu par croissance thermique, 6nm de Si_3N_4 déposé par LPCVD à 625 ° C, 15nm d' Al_2O_3 déposé par ALD à 350 ° C et une grille en Poly-Si N+. Plusieurs recuits ont été appliqués sur ces dispositifs :

- Recuits après dépôt :
 - Recuit 15min 700 ° C sous O_2 ou N_2 .
 - Recuit 1min ou 30min 900 ° C sous O_2 ou N_2 .
 - Recuit flash, 1min ou 2min 1050 ° C sous O_2 ou N_2 .
- Recuit de simulation de l'activation des dopants :
 - Recuit flash 1050 ° C sous N_2 .

La figure 4.36 présente la photographie MET de nos empilements pour différents recuits. Il est intéressant d'observer la densification de l'alumine et la réduction de son épaisseur d'environ 2,5nm pour les recuits à haute température, ce qui est en accord avec les mesures d'épaisseurs et de densités de la figure 3.39. De plus, pour les recuits haute température, on observe des plans cristallins dans l'alumine alors qu'elle reste amorphe après un recuit à 700 ° C. Cela confirme les mesures FTIR-ATR ainsi que les mesures par EELS (cf figure 3.40) sur la nature polycristalline de l'alumine recuite à plus de 850 ° C.

La figure 4.37 regroupe les EOT de l'ensemble des empilements $Al_2O_3/Si_3N_4/SiO_2$ sur lesquels nous avons appliqué plusieurs types de recuits. Tout d'abord, on observe une importante diminution de l'EOT avec la cristallisation de l'alumine. Cela est en accord avec les mesures élipsométriques montrant une diminution $\sim 2,5nm$, cf figure 3.39.

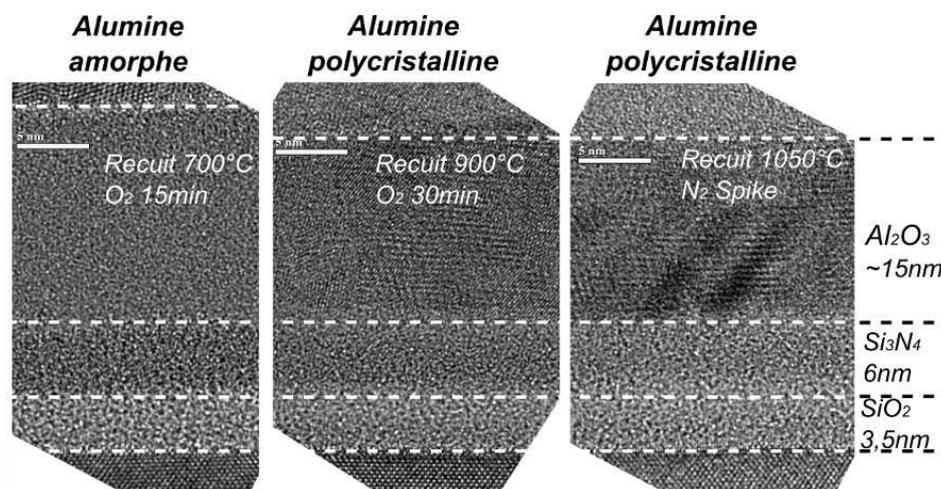


FIG. 4.36: Photographies MET d'empilements Al_2O_3 -15nm/ Si_3N_4 -6nm/ SiO_2 -3,5nm/Si pour différents recuits post-déposition effectués sur Al_2O_3

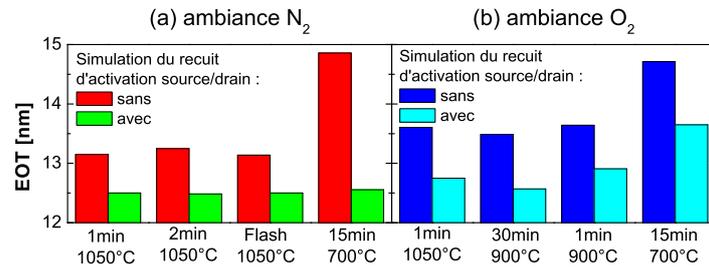


FIG. 4.37: EOT pour l'ensemble de nos empilements Al_2O_3 -15nm/ Si_3N_4 -6nm/ SiO_2 -3,5nm/Si pour différents recuits post-dépôt effectués sur Al_2O_3 .

Néanmoins, il apparaît que les dispositifs qui ont subi un recuit d'activation, effectué après dépôt de la grille et gravure de l'empilement, présentent une EOT plus faible. Il est intéressant de constater que si l'on considère :

- 3,5nm de SiO_2 avec une permittivité de $3,9\epsilon_0$.
- 6nm de SiO_2 avec une permittivité de $7,5\epsilon_0$.
- 15nm d' Al_2O_3 pour l'alumine amorphe (cf figure 3.39).
- 12,5nm d' Al_2O_3 pour l'alumine cristalline (cf figure 3.39).

on obtient une constante diélectrique pour l'alumine de :

- $\sim 8\epsilon_0$ pour les dispositifs ayant subi le recuit d'activation source/drain.
- $\sim 7\epsilon_0$ pour les dispositifs qui n'ont pas subi ce recuit d'activation.

Ainsi, la différence viendrait de la permittivité électrique : les recuits réalisés sur l'alumine juste après dépôt modifieraient l'épaisseur alors que les recuits effectués après dépôt de la grille et gravure de l'empilement joueraient sur la permittivité. L'origine de cette différence n'a pas pu être clairement identifiée. Néanmoins, on peut penser que le fait que l'alumine soit confinée entre deux autres matériaux n'autorise pas les mêmes modifications cristallines qu'une alumine venant d'être déposée.

4.2.4.1 Caractéristiques d'écriture et d'effacement

Les figures 4.38a&b regroupent les caractéristiques d'écriture de ces capacités SANOS pour différents types de recuit. Sur les caractéristiques $\Delta V_{FB} - Temps_{Stress}$, on observe un meilleur piégeage pour les dispositifs intégrant une alumine cristalline. En normalisant en champ, sur les caractéristiques $\Delta V_{FB} - V_G/EOT$, on observe un piégeage similaire. Ainsi, l'augmentation du piégeage précédemment observé vient de la diminution des EOT de l'alumine cristalline. Les éventuelles différences de fuites entre alumine amorphe ou cristalline ne semblent pas perturber la programmation. De même, l'atmosphère du recuit (O_2 ou N_2) de l'alumine n'impacte pas les propriétés de piégeage du dispositif mémoire complet. Seuls les recuits simulant l'activation des sources/drains semblent impacter ces caractéristiques. Ces recuits ont été faits alors que les flancs des empilements étaient à nus, il est possible que cela ait perturbé légèrement l'écriture.

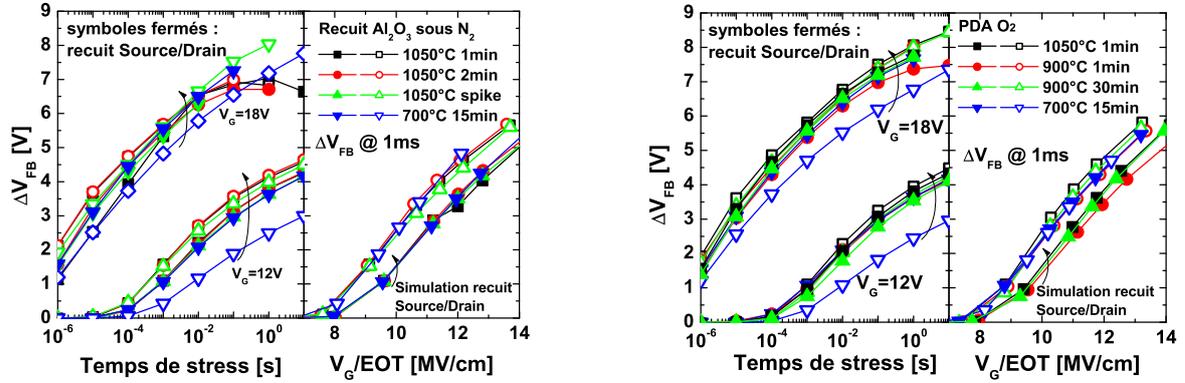
(a) Recuits post-déposition sous N_2 (b) Recuits post-déposition sous O_2

FIG. 4.38: Caractéristiques de programmation $\Delta V_{FB} - Temps_{Stress}$ et $\Delta V_{FB} - V_{G_{Stress}}/EOT @ Temps_{Stress} = 1ms$ d'empilements $Al_2O_3-15nm / Si_3N_4-6nm / SiO_2-3,5nm$ ayant subi différents types de recuits.

La figure 4.39 présente les caractéristiques d'effacement en fonction des différents recuits effectués sur l'alumine. On observe une impossibilité d'effacer les dispositifs possédant une alumine amorphe. Les empilements ayant vu un fort recuit, soit après dépôt de l'alumine, soit en fin de procédé (comme les recuits simulant l'activation des dopants), s'effacent. On observe aussi un meilleur effacement lorsque le dispositif a vu plusieurs forts recuits (cf figure 4.39c). Cette impossibilité d'effacer les empilements avec de l'alumine amorphe ne peut pas provenir de l'EOT. En revanche, elle indique la présence d'un fort

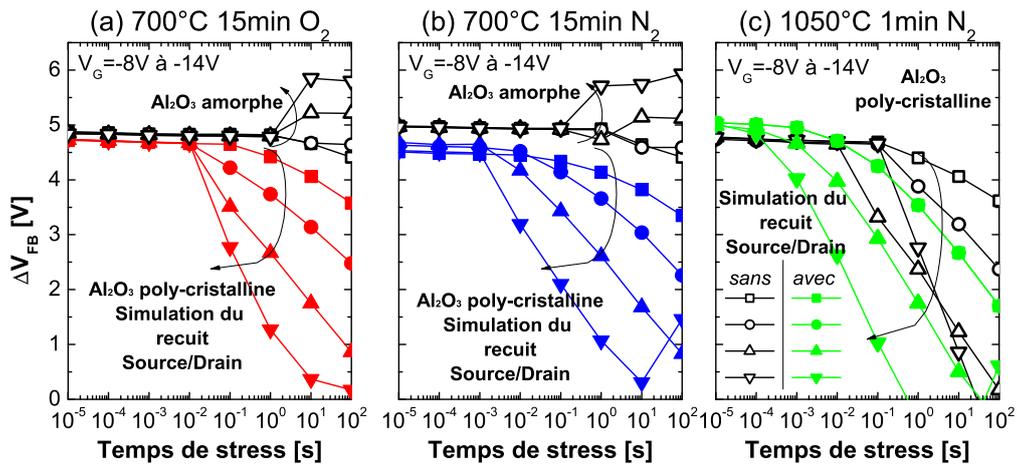


FIG. 4.39: Caractéristiques d'effacement $\Delta V_{FB} - Temps_{Stress}$ d'empilements $Al_2O_3-15nm / Si_3N_4-6nm / SiO_2-3,5nm$ ayant subi différents types de recuits : (a) 700 ° C O_2 15min, (b) 700 ° C N_2 15min et (c) 1050 ° C N_2 1min.

courant de fuite entre la grille de contrôle et le nitrure à travers l'alumine amorphe. En effet, le fort courant d'électrons venant de la grille de contrôle à travers l'oxyde bloquant vient compenser l'effacement. Une hypothèse possible serait une mauvaise compatibilité entre l' Al_2O_3 et la grille polysilicium N+. Les recuits haute température permettent ainsi de protéger l'alumine cristalline grâce à une plus grande densité²⁴ (cf figure 3.39) [Lee00], et/ou de la guérir en recristallisant la couche.

4.2.4.2 Caractéristiques de rétention

La figure 4.40a et la figure 4.40b représentent les caractéristiques de rétention de ces capacités à 25 °C et à 250 °C. On observe une rétention très proche entre les différents empilements à température ambiante. Or à 250 °C, les capacités ayant subi un faible budget thermique ont une meilleure rétention. Les tracés d'Arrhenius de la figure 4.40c, mettent en évidence ce phénomène et l'augmentation de l'énergie d'activation avec la cristallinité de la couche d'alumine.

De plus, pour une grille flottante continue, [Wel06] observent, eux aussi, une meilleure rétention de l'alumine recuit à 700 °C et un meilleur effacement pour leur recuit à 1000 °C. Néanmoins, il est d'élucider de décorrélérer l'effet du nitrure et de l'alumine sur la rétention. Nous pensons malgré tout, en vue de l'ensemble des analyses effectuées sur le nitrure déposé au CEA-Léti, qu'il ne devrait pas être impacté par ces recuits.

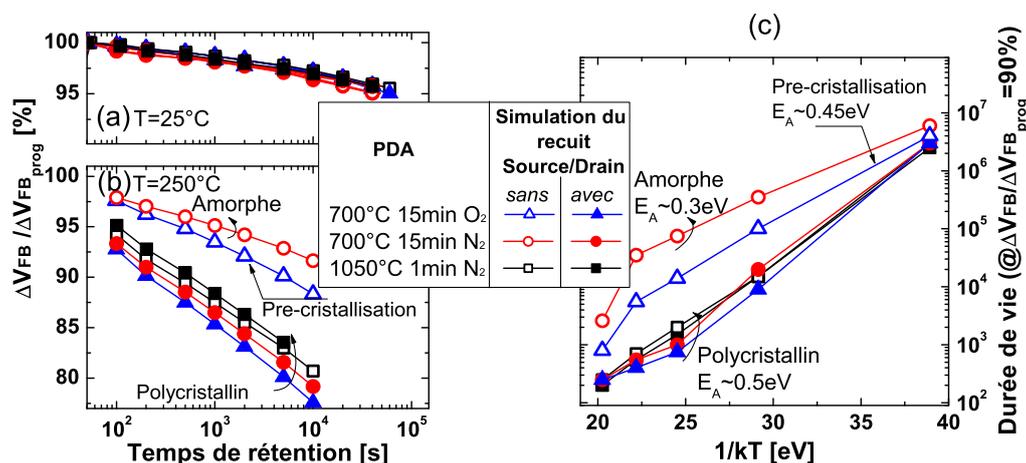


FIG. 4.40: Caractéristiques de rétention $\Delta V_{FB} / \Delta V_{FB-Init}$ a) à 25 °C et b) à 250 °C des capacités SANOS (Al_2O_3 -15nm / Si_3N_4 -6nm / SiO_2 -3,5nm) ayant subi différents types de recuits. c) Tracé d'Arrhenius pour différentes épaisseurs de nitrure. $\Delta V_{FB-Init} \sim 5V$.

²⁴Une densité plus importante donnerait une plus forte cohésion à la matière, la rendant plus robuste aux attaques chimiques d'autres espèces.



En conclusion, l'alumine cristalline semble être plus intéressante qu'une Al_2O_3 amorphe. En effet, sa plus grande densité lui confère une meilleure robustesse. Comme nous l'avons vu pour l'effacement et comme nous allons le voir pour les grilles métalliques (partie 4.2.5.2), l'importance d'une alumine cristalline est essentielle pour de bonnes performances de programmation.

4.2.5 Impact de la grille de contrôle

Comme nous l'avons vu dans la partie 4.2.2.1, l'injection d'électrons de la grille de contrôle vers la couche de piégeage durant l'effacement, appelée *Electron Back Tunneling* (*EBT*), pose de sérieux problèmes de saturation de l'état effacé. Or pour les futures générations de mémoires NAND, le stockage de plusieurs bits par cellule est nécessaire [ITR07]. Ainsi, il est impératif d'obtenir de grandes fenêtres de programmation parfaitement contrôlées. L'une des solutions les plus envisagées est l'utilisation de grilles à fort travail de sortie, comme les grilles en polysilicium P+ ou les grilles métalliques. En effet, elles permettent une plus forte hauteur de barrière entre leurs bandes de conduction et celle de l'oxyde bloquant, réduisant ainsi très fortement l'effet d'*EBT*.

Aussi, dans cette partie, nous évaluerons l'effet d'une grille en polysilicium P+ sur l'effacement, ainsi que l'impact de grilles métalliques en TiN , TaN ou $TaAlN$. Ces matériaux ont été choisis pour leur facilité d'intégration (contrairement au W ou au Pt) : les grilles en TaN sont l'une des briques élémentaires de la TANOS [Lee03] et les grilles TiN sont déjà très utilisées pour les technologies MOS. Nous étudierons, aussi, la compatibilité de ces grilles métalliques avec l'alumine (utilisée en oxyde bloquant) par des mesures de fiabilité des dispositifs mémoires.

4.2.5.1 Étude des grilles en polysilicium

Avant de commencer l'analyse de l'impact d'une grille en polysilicium P+ sur les caractéristiques d'effacement des mémoires SONOS et SAONOS, il faut noter que ce type de grille va augmenter la tension de bande plate (V_{FB}) de l'empilement. Par rapport à une grille polysilicium N+, on peut observer une augmentation de 1V du V_{FB} . Aussi, pour être dans les mêmes conditions d'injection, et ainsi avoir le même ΔV_T , il faudrait une tension de programmation plus forte de 1V pour les grilles en polysilicium P+.

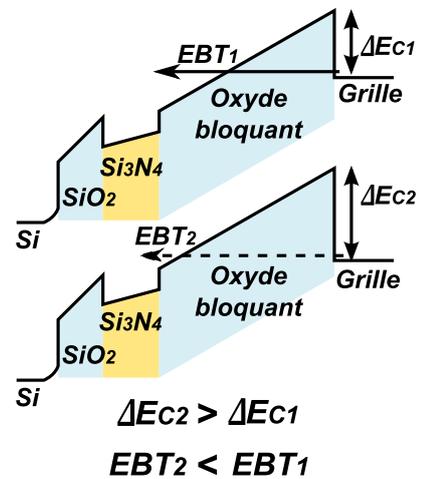


FIG. 4.41: Diagrammes de bandes de mémoires SONOS avec différentes grilles de contrôle.

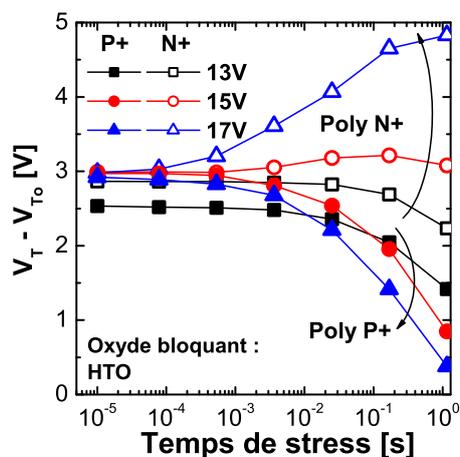


FIG. 4.42: Caractéristiques d'effacement pour des mémoires de type SONOS (SiO_2 -2,5nm/ Si_3N_4 -6nm/ HTO -12nm) intégrant une grille en Si-poly N+ et une grille en Si-poly P+ [Boc08].

La figure 4.42 présente les caractéristiques d'effacement d'une mémoire SONOS (SiO_2 -2,5nm/ Si_3N_4 -6nm/ HTO -12nm). Comme nous l'avons déjà observé dans la partie 4.2.2.1 les empilements SONOS présentent des difficultés à être effacés avec une grille en polysilicium N+. On observe même une écriture pour les tensions fortement négatives. Les électrons injectés depuis la grille (*EBT*) écrivent la mémoire. Cela est le signe :

- D'un *HTO* de moins bonne qualité que le SiO_2 thermique utilisé comme oxyde *tunnel*.
- D'une difficulté à émettre les électrons piégés ou d'injecter des trous dans le nitrure depuis le canal.

Aussi pour les mémoires SONOS, avec un oxyde bloquant en *HTO*, les grilles en polysilicium N+ rendent l'effacement très difficile alors que les grilles en polysilicium P+ le permettent, même à des tensions fortement négatives, grâce à leur fort travail de sortie qui diminue les effets d'*EBT*.

Les figures 4.43a&b présentent les caractéristiques d'effacement de mémoires SAONOS (SiO_2 -2,5nm/ Si_3N_4 -6nm/ HTO -5nm/ Al_2O_3 -8nm) intégrant une grille en polysilicium N+ ou en polysilicium P+. Pour les mémoires SAONOS, il est possible d'effacer même avec la grille en polysilicium N+, contrairement aux dispositifs SONOS. Cela est dû au meilleur couplage que confère l'oxyde bloquant en HTO/Al_2O_3 , permettant d'augmenter le champ électrique dans l'oxyde *tunnel* et diminuer celui dans l'oxyde bloquant. Malgré cette amélioration apportée par l'intégration de l'alumine, on observe une saturation de l'état effacé à $V_G = -18V$ qui n'apparaît pas pour la grille en polysilicium P+. Cette saturation est due au plus faible travail de sortie des grilles en polysilicium N+ entraînant des effets d'*EBT*.

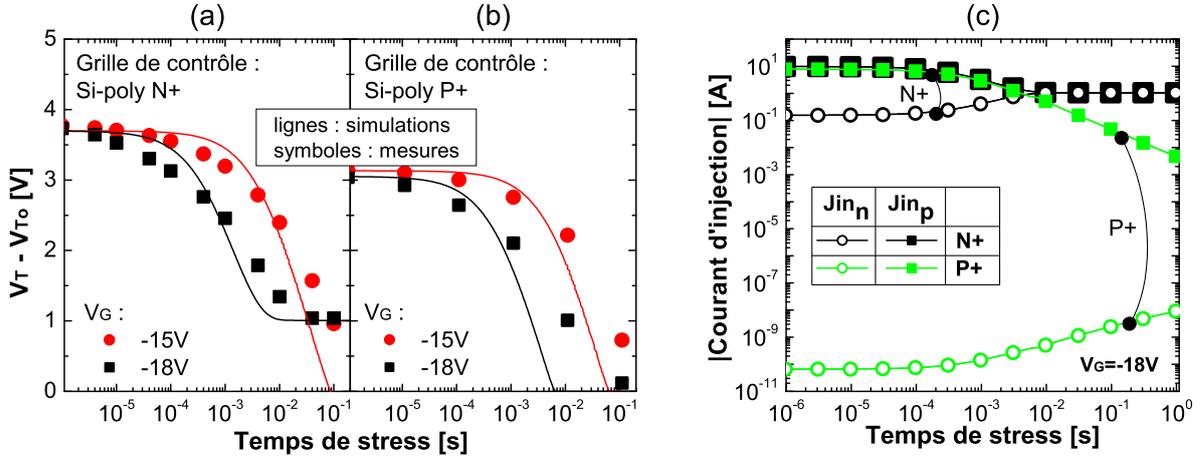


FIG. 4.43: Caractéristiques expérimentales et simulées d'effacement pour des mémoires de type SAONOS (SiO_2 -2, 5nm/ Si_3N_4 -6nm/ HTO -5nm/ Al_2O_3 -8nm) intégrant (a) une grille en Si-poly N+ et (b) une grille en Si-poly P+. (c) Courant de trous (courant d'effacement) et d'électrons (courant parasite d'*EBT*) injectés dans la structure simulée [Boc08].

Pour bien mettre en évidence les mécanismes jouant un rôle important sur l'effacement, nous avons simulé les caractéristiques d'effacement en s'appuyant sur le modèle présenté dans la partie 2.3.2. Les paramètres utilisés sont les mêmes que ceux de la partie 4.2.2.1 et sont regroupés dans le tableau 4.1. Les simulations ont montré que pour ce type de structure, les mécanismes dominants sont l'injection de trous venant du substrat et l'injection d'électrons venant de la grille. L'émission d'électrons venant du nitrure pouvant être négligée.

La figure 4.43c représente l'évolution des courants *tunnel* durant l'effacement extraits de la simulation. On observe que la valeur du courant de trous venant du substrat (J_{in_p}) est similaire pour les deux types de grilles jusqu'à $10^{-3}s$. Cela explique les cinétiques d'effacement très proches entre les deux dispositifs. Néanmoins, on observe que la valeur du courant d'électrons venant de la grille (J_{in_n}) est très élevée pour la grille en polysilicium N+. Le courant d'*EBT* devient égal à J_{in_p} à $10^{-2}s$, faisant saturer l'effacement de la mémoire. La grille en polysilicium P+ a un courant J_{in_n} très faible qui empêche la saturation de l'effacement pour les temps et les tensions utilisés.

4.2.5.2 Étude des grilles métalliques en TiN, TaN ou TaAlN

L'une des difficultés des grilles en polysilicium P+ est leur compatibilité avec l'alumine [Lee00]. Aussi, l'utilisation de grilles métalliques est fortement étudiée. Les grilles en *TaN* ont montré de très bonnes performances avec l' Al_2O_3 grâce à leur travail de sortie, leur excellente stabilité thermique et leur bonne résistance à l'oxydation [Lee03]. Le *TaAlN*,

Grille	Travail de sortie
<i>PolySi - N+</i>	4, 1eV
<i>TaN</i>	4, 45eV
<i>Ta_{1-x}Al_xN_y</i>	4, 5 ~ 4, 8eV
<i>TiN</i>	4, 6eV
<i>PolySi - P+</i>	5, 2eV

TAB. 4.3: Travail de sortie des grilles métalliques étudiées [Als06, Gil09].

grâce à sa teneur en aluminium peut prétendre à une meilleure compatibilité avec l'alumine. Enfin, le *TiN* est très utilisé dans les nouvelles technologies MOS, intégré avec des matériaux *High-κ* à base d'hafnium. Aussi, nous avons voulu évaluer la compatibilité du *TiN* avec les matériaux utilisés dans les technologies TANOS. Le tableau 4.3 regroupe le travail de sortie de ces grilles. Elles ont globalement un travail de sortie de type *mid-gap*, c'est-à-dire que leur bandes de conduction arrivent environ au milieu du gap du silicium.

Dans un premier temps nous avons étudié les courants de fuite à travers une alumine de 16nm ayant subi différents types de recuits après dépôt :

- 700 ° C sous O_2 durant 15min : alumine amorphe.
- 900 ° C sous O_2 durant 30min : alumine cristalline.

et différentes grilles métalliques :

- 10nm de *TiN* déposé par PVD à 680 ° C.
- 10nm de *TaN* déposé par AVD²⁵.
- 5nm de *TaAlN* déposé par AVD.

La figure 4.44 montre les caractéristiques $I_G - V_G$ pour les différents empilements. On observe un claquage prématuré des empilements à base de *TiN* sur alumine amorphe, alors que sur alumine cristalline, le *TiN* donne les mêmes performances que les grilles *TaN*. En effet, le *TiN* est avide d'oxygène. Il dégrade l'interface *TiN/Al₂O₃* et donc la robustesse de l'empilement. Or l'alumine cristalline est plus dense et plus robuste aux dégradations d'interface que l'alumine amorphe (cf figure 3.39). Donc l'alumine cristalline peut être considérée comme compatible avec le *TiN*. On notera aussi que les grilles *TaN* et *TaAlN* présentent une très bonne compatibilité quelle que soit la cristallinité de l'alumine.

Ensuite nous avons voulu étudier la compatibilité des grilles une fois intégrées dans une cellule mémoire complète. Nous avons donc étudié des cellules mémoires *TANOS* (*SiO₂-3nm/Si₃N₄-6nm/Al₂O₃-15nm*) dans le cas d'une alumine cristalline.

²⁵ Dans l'AVD (Aixtron), l'un des précurseurs est dissout dans un solvant. Il est ensuite vaporisé dans la chambre de dépôt grâce à des injecteurs, la quantité injectée par pulses est de l'ordre de quelques microlitres. Ce précurseur va réagir avec un gaz actif uniformément distribué sur le substrat. L'atmosphère de la chambre est contrôlée par un gaz inerte. Plus de détails sur l'AVD sont donnés dans [Web05]. AVD (un précurseur est présent en permanence, un autre est introduit par cycles) peut être vu comme un compromis entre la CVD (les précurseurs sont présents durant tout le dépôt) et l'ALD (les précurseurs sont introduits les uns après les autres, par cycles). La principale différence entre AVD et la CVD ou l'ALD est sa grande vitesse de croissance contrôlée (quelques *nm/min*)[Kar07].

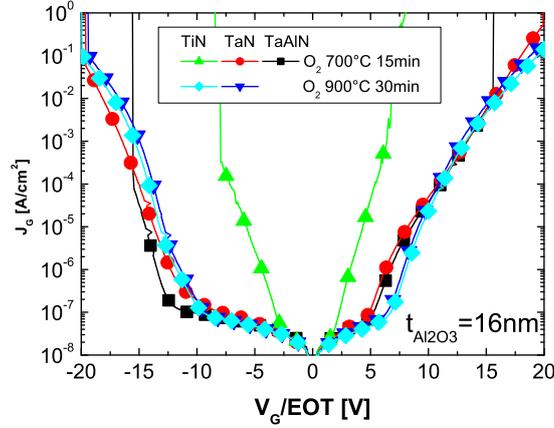


FIG. 4.44: Caractéristiques $I_G - V_G$ d' Al_2O_3 -16nm ayant subi différents recuits et différents types de grilles.

Comme le montre la figure 4.45a, à même V_G/EOT ²⁶, les grilles TiN, TaN et TaAlN présentent une saturation de l'effacement plus faible qu'une grille en polysilicium N+. De plus, sur le réseau des caractéristiques d'écriture/effacement, figure 4.45b, on observe que ces grilles métalliques permettent un sureffacement jusqu'à $-2V$ avec les tensions et les épaisseurs considérées. On peut donc en conclure que ces grilles présentent un fort potentiel pour l'augmentation de la fenêtre mémoire. De plus, le *TiN* ne semble pas poser de problèmes de compatibilité lorsque qu'il est déposé sur de l'alumine cristalline pour l'écriture et l'effacement.

Les figures 4.46a-c comparent les caractéristiques de rétention et d'endurance de *TANOS* avec des grilles en *TaN* ou en *TiN*. Les mêmes caractéristiques sont observées pour les deux grilles. Ainsi, la grille en *TiN* semble présenter la même compatibilité avec l'alumine cristalline que la grille en *TaN* pour l'ensemble des caractéristiques mémoires lorsque le dépôt de grille est effectué sur de l'alumine cristalline.

Pour aller plus loin dans l'analyse de ces résultats, on peut observer que la rétention n'est pas dégradée après une endurance de 10^4 cycles (cf figure 4.46b). En effet, si l'on observe la caractéristique de l'endurance (cf figure 4.46c), on remarque que les deux états (écrit et effacé) remontent parallèlement, gardant la fenêtre de programmation parfaitement ouverte et constante. De plus, les $I_D - V_G$ montrent une diminution de la pente sous le seuil, caractéristique d'une dégradation de l'interface *Si/SiO₂*. Aussi, l'absence de dégradation de la rétention après endurance et l'absence de décalage des $I_D - V_G$ après plusieurs cycles de programmation, nous indiquent que les 10^4 cycles n'ont pas dégradé les différentes couches de l'empilement. Seule l'interface *Si/SiO₂* a été dégradée en raison

²⁶Nous avons choisi de comparer ces dispositifs à V_G/EOT constant, car les plaques intégrant des grilles en polysilicium N+ présentaient des EOT légèrement différentes : $EOT_{Poly-N+} \approx 13,4nm$ et $EOT_{Metal} \approx 12nm$.

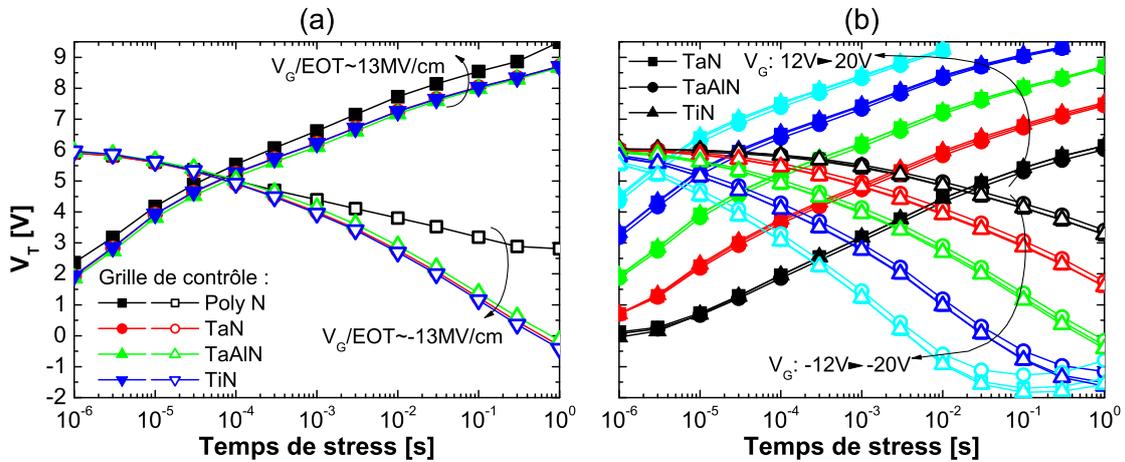


FIG. 4.45: Caractéristiques d'écriture et d'effacement pour des mémoires de type SANOS ($SiO_2-3nm/Si_3N_4-6nm/Al_2O_3-15nm$) intégrant différents types de grilles de contrôle. (a) Comparaison entre une grille en Si-poly N+ et des grilles en TiN , TaN et $TaAlN$ à $V_G/EOT = 13MV/cm$. (b) Réseaux des caractéristiques de programmation des grilles métalliques. L'alumine a été recuite à $900^\circ C$ sous O_2 pendant 30min.

d'un champ d'injection très fort ($> 15MV/cm$) [Van09]. Cela met en évidence que les différentes grilles présentent la même compatibilité avec l'alumine, sans la dégrader ni l'améliorer, et qu'il faut réaliser le dépôt après cristallisation de la couche d'alumine.

Nous avons voulu soumettre ces empilements à des conditions d'endurance plus agressives en terme de tension de programmation, pour déceler un éventuel claquage de l'alumine en fonction de la grille de contrôle. La figure 4.47 nous montre, encore une fois, des

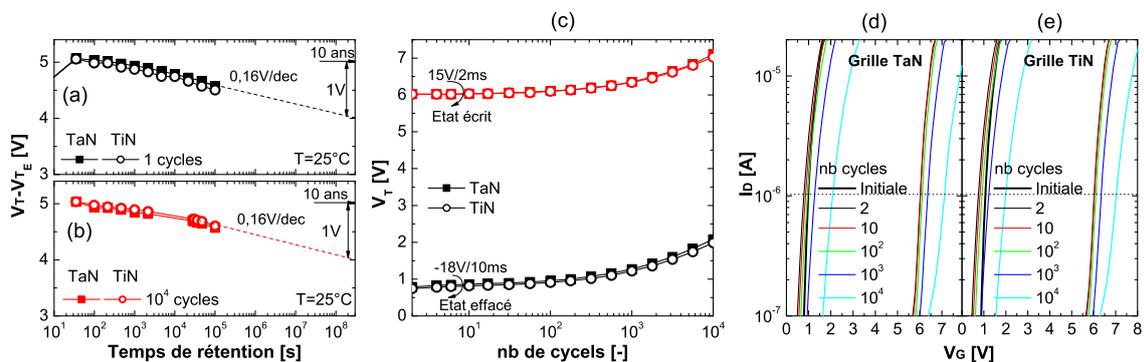


FIG. 4.46: Caractéristiques de rétention de mémoires de type TANOS ($SiO_2-3nm/Si_3N_4-6nm/Al_2O_3-15nm$) intégrant une grille TaN ou TiN (a) avant endurance et (b) après 10^4 cycles. (c) Caractéristiques d'endurance. Caractéristiques $I_D - V_G$ durant l'endurance pour (d) la grille TaN et (e) pour la grille TiN . L'alumine a été recuite à $900^\circ C$ sous O_2 pendant 30min.

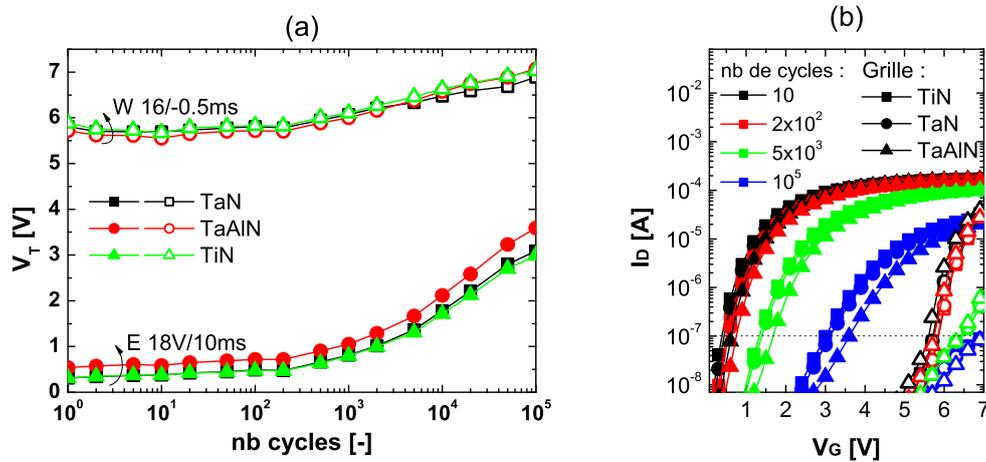


FIG. 4.47: (a) Caractéristiques d'endurance pour des mémoires de type TANOS (SiO_2 -3nm/ Si_3N_4 -6nm/ Al_2O_3 -15nm) intégrant différents types de grilles de contrôle. (b) Réseaux des caractéristiques $I_D - V_G$ lors de l'endurance. L'alumine a été recuite à 900 ° C sous O_2 pendant 30min.

caractéristiques similaires entre *TaN* et *TiN* déposés sur alumine cristalline même après 10^5 cycles et confirment les résultats précédents.



En conclusion, il est important de souligner l'importance de l'intégration de grilles à fort travail de sortie, notamment *TaN*, pour réduire la saturation de l'état effacé et permettre le sureffacement. De plus, nous avons mis en évidence la très mauvaise compatibilité entre la grille *TiN* et l'alumine amorphe, alors que sur alumine cristalline, une grille en *TiN* présente les mêmes caractéristiques d'écriture, d'effacement, d'endurance et de rétention qu'une grille *TaN*. Il est donc essentiel de cristalliner l'alumine avant le dépôt de grille.

4.3 Mémoire à base de nitrure avec oxyde *tunnel* en $SiO_2/HfSiON$

Les fortes tensions de programmation des mémoires Flash-NAND demandent une importante périphérie²⁷ pour rendre ces tensions compatibles avec celles de la microélectronique. C'est aussi l'un des principaux problèmes pour leur utilisation dans les applications embarquées. De plus, nous avons vu (cf figures 4.8, 4.19, 4.46 et 4.47) que l'utilisation de forts champs d'injection impacte fortement les caractéristiques d'endurance en dégradant l'interface Si/SiO_2 . Ainsi la réduction des tensions de programmation est un point clé de l'intégration des futures mémoires Flash-NAND. L'une des solutions les plus efficaces est l'ingénierie de bande.

Le principe de l'ingénierie de bande consiste à choisir des matériaux en fonction de leur hauteur de barrière et de leur permittivité pour optimiser leurs propriétés d'injection ou d'isolation. Dans le cas d'un oxyde *tunnel*, il est nécessaire de réduire les fuites à faible champ, pour améliorer la rétention, et d'augmenter l'injection à fort champ, pour accélérer la vitesse de programmation. Comme nous l'avons vu dans les parties 1.4.2 et 3.5, plusieurs types de barrière sont envisagés en tant qu'oxyde *tunnel* :

- Les monocouches : $HfSiON$ [Van05, Van06b].
- Les bicouches : SiO_2 -2, 9nm/ Si_3N_4 -1, 6nm [Wan07, Lin08].
- Les tricouches : par exemple SiO_2 -2nm/ Al_2O_3 -8nm/ HTO -2nm [Blo09a], $SiO_2/Si_3N_4/HTO$ [Lai07] ou SiO_2 -1, 9nm/ $HfSiON$ -3, 6nm/ HTO -4, 1nm [Ver09].

L'utilisation de tricouches à base de Si_3N_4 est actuellement l'option la plus étudiée [Lue08a, Lue08b], elles sont appelées les BE-SONOS²⁸. Néanmoins, un piégeage parasite dans le nitrure de l'oxyde *tunnel* peut être problématique. Ainsi d'autres matériaux sont à envisager.

Dans cette partie nous nous intéresserons à l'intégration d'un empilement bicouche $SiO_2/HfSiON$ en tant qu'oxyde *tunnel*. L'étude que nous avons réalisée dans la partie 3.5, a mis en évidence son intérêt par rapport à un oxyde de SiO_2 . Nous allons donc étudier ce bicouche dans des mémoires à couche de piégeage nitrure.

Les détails de l'analyse matériau sont regroupés dans la partie 3.5. On y retrouve le procédé de fabrication du $HfSiON$. On rappellera que le $HfSiO$ a été nitruré pour le maintenir amorphe et assurer une bonne fiabilité en terme de piégeage et de courant de fuite à bas champ.

²⁷Coûteuse en surface sur le silicium.

²⁸BE-SONOS : Bandgap Engineering SONOS

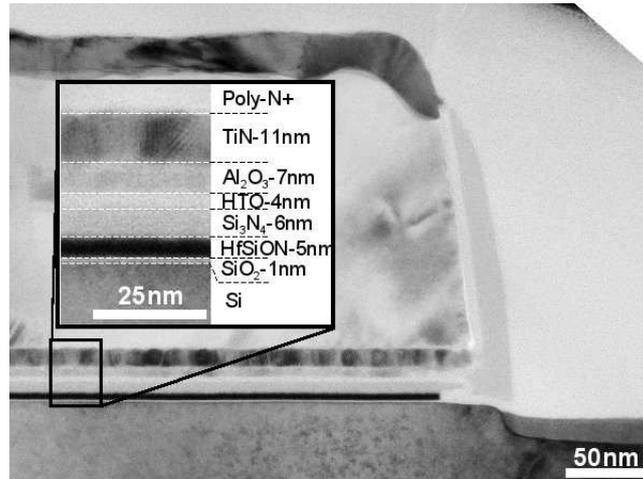


FIG. 4.48: Photographies MET de l'empilement mémoire complet THONHOS : $TiN-11nm/Al_2O_3$:
4-7nm/ $HTO-4nm/Si_3N_4-6/HfSiON-5/SiO_2-1nm$.

Il est important de souligner que les empilements étudiés dans cette partie peuvent être considérés comme la synthèse des résultats présentés dans ce chapitre. En effet, nous avons intégré :

- Un oxyde bloquant bicouche HTO/Al_2O_3 , pour améliorer les propriétés de rétention.
- Des diélectrique d'interpoly intégrant du $HfAlO-1:4$ ont été réalisés, pour optimiser le compromis programmation/rétention.
- Des grilles en TiN , pour réduire l'injection parasite de la grille de contrôle durant l'effacement par rapport à une grille en polysilicium dopé N^+ .

La figure 4.48 est une photographie HR-TEM d'un des empilements étudiés. Il a la particularité d'intégrer une grille en TiN . Comme le $HfSiON$ est déposé en deux étapes, il est important de remarquer que l'on n'observe aucune discontinuité dans la couche.

4.3.1 Caractéristiques d'écriture et d'effacement

La figure 4.49 présente une comparaison des performances d'écriture et d'effacement entre une SAONOS, avec un oxyde *tunnel* en SiO_2 de 2,5nm, et une SAONHOS²⁹, intégrant un bicouche $SiO_2-1nm/HfSiON-5nm$. On observe une très nette augmentation des vitesses d'écriture et d'effacement pour la SAONHOS, malgré une EOT similaire. Cela met en évidence les propriétés d'injection supérieures du bicouche $SiO_2/HfSiON$ sur un SiO_2 "classique" (cf figure 3.51). On obtient ainsi un ΔV_T de 4V pour des conditions de programmation de 16V – 100μs / -16V – 1ms pour la SAONHOS alors qu'il faut 16V – 1ms / -16V – 100ms pour la SAONOS.

²⁹SAONHOS : $Si/Al_2O_3/HTO/Si_3N_4/HfSiON/SiO_2/Si$

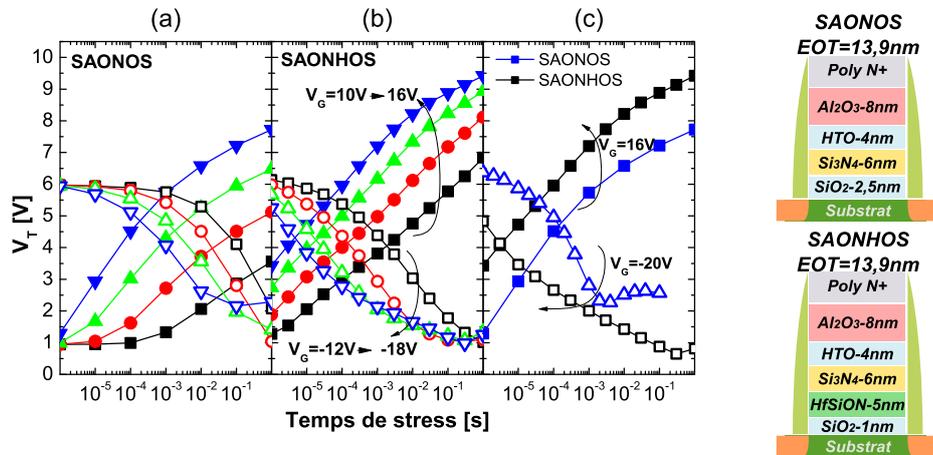


FIG. 4.49: Caractéristiques de programmation et d'effacement $V_T - Temps_{Stress}$ (a) pour un *tunnel* standard 2,5nm et (b) pour un bicouche $SiO_2-1nm/HfSiON-5nm$. (c) Comparaison de ces deux dispositifs.

Les figures 4.50a&b comparent l'endurance de la SAONOS et de la SAONHOS. On observe une caractéristique d'endurance constante dans le cas du bicouche $SiO_2-1nm/HfSiON-5nm$ avec un $\Delta V_T \sim 4V$. Les courbes $I_D - V_G$ (cf figure 4.50c) sont superposées, alors que pour la SAONOS, on observe une remontée des états écrits et effacés due à la dégradation de l'interface Si/SiO_2 (cf figure 4.50d). En effet, la plus grande sensibilité au champ du bicouche permet une bonne injection avec des champs faibles ($< 10MV/cm$) limitant ainsi la dégradation de l'interface Si/SiO_2 et permettant une grande amélioration de l'endurance. On peut aussi noter qu'aucun piègeage parasite irréversible dans le $HfSiON$ ne vient perturber l'endurance.

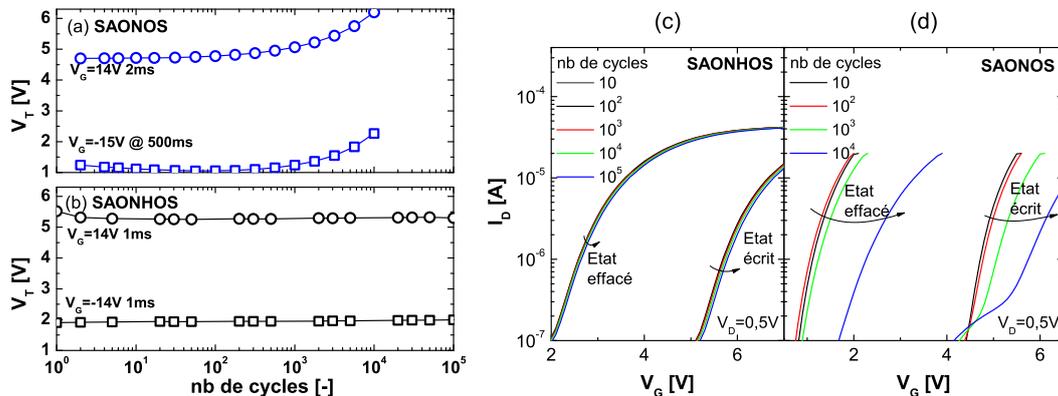


FIG. 4.50: Caractéristiques d'endurance a) d'un empilement SAONOS (8nm/4nm/6nm/2,5nm) b) d'un empilement SAONHOS (8nm/4nm/6nm/5nm/1nm). $I_D - V_G$ durant l'endurance pour c) la SAONHOS et d) la SAONOS.

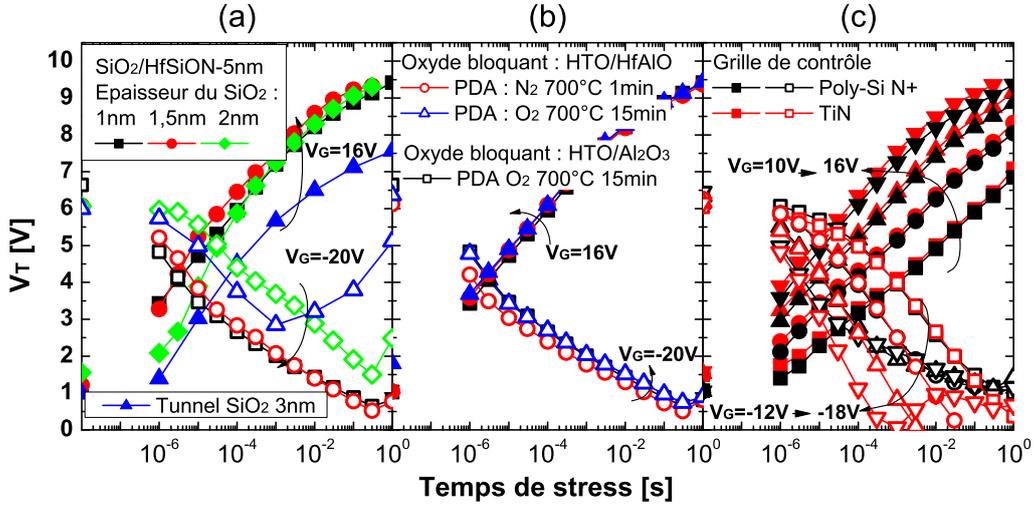


FIG. 4.51: Caractéristiques $V_T - \text{Temps}_{\text{Stress}}$ d'écriture et d'effacement d'empilements (a) *Tunnel/Si₃N₄-6nm/HTO-4nm/Al₂O₃-8nm* pour différents types d'oxydes tunnel, (b) *SiO₂-1nm/HfSiON-5nm/Si₃N₄-6nm/HTO-4nm/High-κ-8nm* pour différents *High-κ* ayant subi différents recuits et (c) *SiO₂-1nm/HfSiON-5nm/Si₃N₄-6nm/HTO-4nm/Al₂O₃-8nm* pour différentes grilles de contrôle.

La figure 4.51a représente les caractéristiques de programmation pour différentes épaisseurs SiO_2 (1nm, 1,5nm et 2nm) dans le bicouche tunnel ($\text{SiO}_2/\text{HfSiON}$). Il apparaît que l'augmentation du SiO_2 perturbe peu l'écriture mais ralentit l'effacement. Dans tous les cas, les propriétés d'injection sont supérieures au SiO_2 classique de 3nm.

Sur la figure 4.51b, des *High-κ* différents ont été intégrés dans le bicouche de l'oxyde bloquant ($\text{HTO}/\text{High-κ}$) :

- Soit de l'alumine.
- Soit du HfAlO -1:4 ayant subi plusieurs types de recuit juste après dépôt.

On peut observer que le *High-κ* dans l'oxyde bloquant impacte peu les caractéristiques de programmation. En effet, le HTO de l'oxyde bloquant joue le rôle de barrière pour les électrons injectés et masque d'éventuelles fuites de ce *High-κ*. De plus, la différence d'EOT entre le dispositif intégrant de l'alumine ou de l'aluminate hafnium n'est que de $\sim 1\text{nm}$ (soit $\sim 10\%$ sur l'EOT total). Ainsi, cette différence impacte peu le champ dans l'oxyde tunnel³⁰.

Enfin, la figure 4.51c présente l'effet de la grille TiN . Les dispositifs à grille polysilicium N+ présentent une saturation de leurs états effacés pour les fortes tensions et les temps longs. La grille en TiN réduisant fortement l'*Electron Back Tunneling* (cf figure 4.45), permet d'améliorer nettement l'effacement. Il ne faut pas, malgré tout, dégrader la cellule avec des tensions trop fortes et des temps trop longs (exemple à -18V après 1ms).

³⁰À $V_G = 16\text{V}$ et $\Delta V_T = 0\text{V}$, le champ dans le SiO_2 est $\sim 11\text{MV}/\text{cm}$ pour l'oxyde bloquant à base d' Al_2O_3 et de $\sim 11,5\text{MV}/\text{cm}$ pour celui à base de HfAlO .

4.3.2 Caractéristiques de rétention

L'étude réalisée sur les bicouches, dans la partie 3.5.2, a montré que ces structures pouvaient présenter des courants de fuite, à faible champ, plus faibles que l'oxyde *tunnel* classique de $3nm$ de SiO_2 . Dans cette partie, nous nous intéresserons au bicouche $SiO_2/HfSiON$ sur la rétention d'une mémoire à couche de piégeage en nitrure.

La figure 4.52a présente la caractéristique de rétention d'une SAONOS et d'une SAONHOS. On observe, sur la caractéristique de la SAONHOS, une perte de charge rapide pour les temps courts. Plusieurs mécanismes peuvent expliquer ce phénomène :

- Une redistribution de la charge dans le nitrure.
- Une émission des électrons piégés dans le $HfSiON$ par *tunnel* direct dans le substrat (Trap to Band) (cf partie 2.2.4.2).
- Une émission des électrons piégés à l'interface $HfSiON/SiO_2$ par *tunnel* direct (Trap to Band).
- Une évacuation des électrons présents sur la bande de conduction du nitrure.

Néanmoins après 10^4s , la perte de charge devient similaire à la SAONOS. Ainsi, il est possible d'obtenir la même rétention qu'une SAONOS en écrivant un peu plus la mémoire SAONHOS (environ $0,6V$ de plus). Il est important de souligner que, malgré l'augmentation du temps ou de la tension de programmation, la SAONHOS reste plus rapide que la SAONOS.

Les figures 4.52b-c représentent les caractéristiques de rétention à $25^\circ C$, $125^\circ C$ et $200^\circ C$ pour les différents empilements mémoires étudiés : la SANOS, la SAONOS et la SAONHOS. On observe l'intérêt d'intégrer une couche de *HTO* entre l'alumine et le

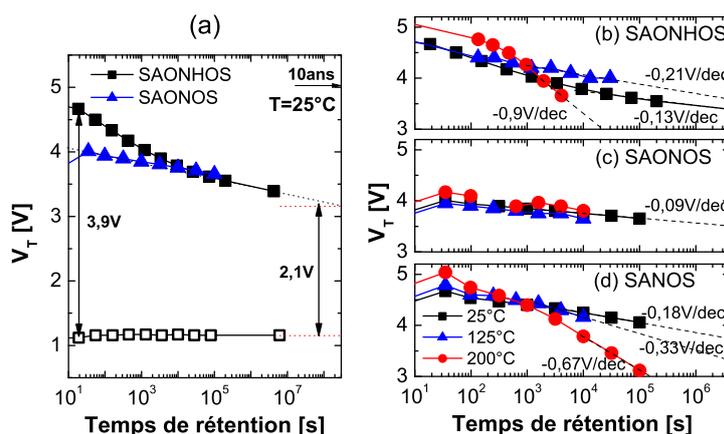


FIG. 4.52: (a) Comparaison des caractéristiques de rétention $V_T - Temps_{Ret}$ d'un empilement SAONHOS ($8nm/4nm/6nm/5nm/1nm$) et SAONOS ($8nm/4nm/6nm/2,5nm$) à $25^\circ C$. Caractéristiques de rétention à $25^\circ C$, $125^\circ C$ et à $200^\circ C$ pour (b) la SAONHOS, (c) la SAONOS et (d) la SANOS ($16nm/6nm/2,5nm$).

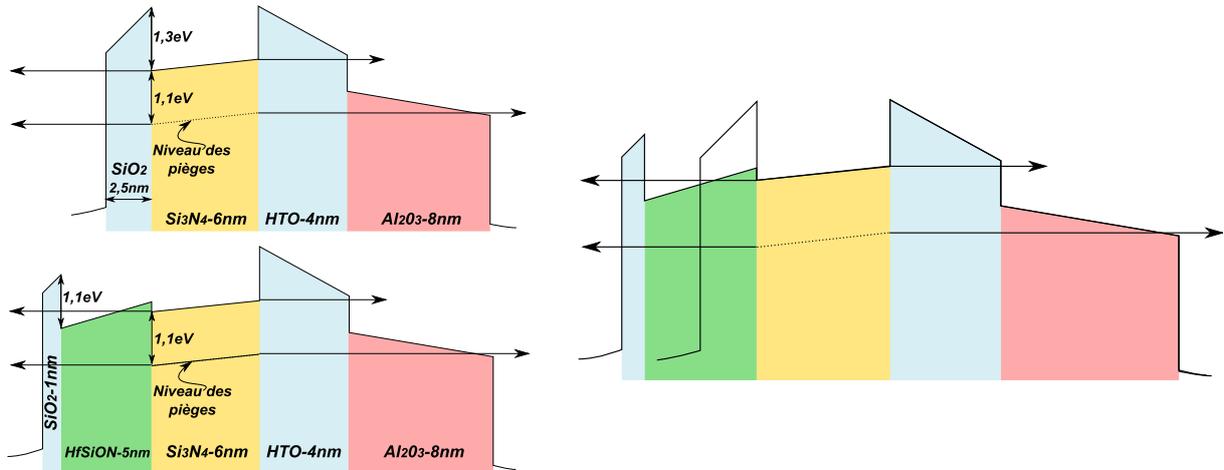


FIG. 4.53: Diagrammes de bandes des deux structures, SAONHOS (8nm/4nm/6nm/5nm/1nm) et une SAONOS (8nm/4nm/6nm/2,5nm), calculés à $V_G = 0V$ et pour un ΔV_T de 5V.

nitride pour améliorer la rétention et diminuer l'activation en température. On observe aussi la dégradation de la rétention avec l'intégration d'un oxyde *tunnel* bicouche.

Néanmoins, à 125 ° C, nous n'observons plus cette perte rapide de charges. Il est possible qu'elle intervienne à des temps plus courts. De plus comme la température favorise l'injection, on observe le même ΔV_T qu'à 25 ° C après quelques secondes de rétention.

La figure 4.53 représente les diagrammes de bandes de la SAONOS et de la SAONHOS. On observe que :

- La transparence de l'oxyde bloquant est identique pour les deux dispositifs.
- La transparence vue par les électrons piégés à 1,1eV est plus faible pour le bicouche $SiO_2/HfSiON$, ce qui va à l'encontre de la perte rapide de charges.
- Des électrons sur la bande de conduction du nitride voient une très faible hauteur de barrière du $HfSiON$ favorisant la perte de charges dans le cas du bicouche.

Il semble donc que la perte rapide de charges du bicouche est due, soit :

- À l'émission des électrons piégés dans le $HfSiON$.
- Aux éventuels électrons localisés sur la bande de conduction du nitride.

Dans tous les cas, il apparaît que le faible écart entre la bande de conduction de l' $HfSiON$ et du Si_3N_4 est un problème pour la rétention. L'une des solutions possibles est l'utilisation d'une zone de stockage de charges à faible gap comme les nanocristaux ou AlN [Lai05] (cf tableau 3.1).



En conclusion, les bicouches $SiO_2/HfSiON$ présentent d'excellentes performances en écriture, en effacement et en endurance grâce à leurs propriétés d'injection supérieures à un oxyde *tunnel* classique. En rétention, nous avons observé une perte rapide de charges pour les temps courts ($< 10^3s$). Il est possible de la compenser avec un V_{Tprog} plus élevé. Même avec cette programmation plus longue, le bicouche reste plus rapide qu'un oxyde *tunnel* classique. Ainsi, l'utilisation de *High-κ* dans l'oxyde *tunnel* est une approche très intéressante pour les applications demandant une grande endurance.

4.4 Conclusion

Nous avons étudié dans ce chapitre l'intérêt des matériaux *High- κ* intégrés dans des mémoires non-volatiles pour les applications NAND. Tout d'abord nous avons vu la nécessité d'utiliser des diélectriques d'interpoly à base de *High- κ* pour permettre la programmation des mémoires à nanocristaux en Fowler-Nordheim. Les tricouches, *HTO/HfAlO/HTO*, permettent de bons compromis entre tensions de programmation et rétention. Pour les applications à couche de piégeage en nitrure, nous avons regroupé l'ensemble des performances des différents dispositifs étudiés dans le tableau 4.4.

Pour les mémoires à couche de piégeage nitrure, nous avons mis en évidence que l'utilisation d'oxyde bloquant à base d' Al_2O_3 est indispensable, pour assurer une bonne écriture mais surtout un bon effacement. L'utilisation d'un bicouche *HTO/Al₂O₃* permet une importante amélioration des caractéristiques de rétention par rapport à une mémoire SANOS sans dégrader l'EOT. En effet, le *HTO* présente une plus grande barrière aux électrons piégés empêchant leur fuite à travers l'oxyde bloquant.

Nous avons aussi mis en évidence que l'augmentation des vitesses d'écriture et d'effacement due à la diminution de l'épaisseur de l'alumine et/ou du nitrure à $V_G = cst$ était directement liée à la réduction de l'EOT de l'empilement de grille et à l'augmentation du champ d'injection. On retiendra que l'écriture est fortement contrôlée par le champ électrique de l'oxyde *tunnel*, mais que l'épaisseur du nitrure et le courant de fuite à travers l'oxyde bloquant jouent un rôle important sur l'effacement, la rétention et l'endurance. En effet, il apparaît que les nitrures fins présentent une moins bonne efficacité de piégeage pouvant dégrader les performances d'endurance. Pour la rétention à température ambiante, on retiendra que le principal mécanisme est l'émission des charges initialement piégées aux interfaces, donnant au nitrure épais une meilleure rétention. À hautes températures, les émissions thermo-assistées diminuent les bénéfices des nitrures épais.

Concernant l'étude de la grille de contrôle, il est très important de souligner la nécessité d'intégrer des grilles à fort travail de sortie. En effet, nous avons vu que l'un des principaux problèmes des mémoires SANOS est d'allier rétention et effacement. Il est donc impératif qu'aucun mécanisme ne vienne ralentir l'effacement, comme le courant d'électrons venant de la grille de contrôle (EBT). Les grilles *TaN* sont annoncées comme prometteuses. Nous avons mis en évidence qu'il est préférable de déposer les grilles sur une alumine cristallisée pour réduire la défektivité en particulier avec le *TiN*. La grille *TiN* présente les mêmes caractéristiques d'écriture, d'effacement, d'endurance et de rétention qu'une grille *TaN*.

Enfin, l'intégration de bicouches *SiO₂/HfSiON* comme oxyde *tunnel* constitue une approche intéressante pour réduire les tensions de programmation. En effet, leurs fortes sensibilités aux champs leur permettent une injection importante d'électrons à des champs plus faibles qu'un oxyde *tunnel* classique. Ainsi, ils permettent une très nette amélioration de l'endurance.

	EOT [nm]	Écrit./Effac. $\Delta V_T \sim 3V$	Charges après 10 ans à 25 ° C	Endurance $W/E_{1cyc.} \rightarrow W/E_{10^5 cyc.}$
SONOS & SAONOS : <i>SiO₂-2,5nm/Si₃N₄-6nm/Oxyde bloquant</i>			$\Delta V_{Tinit} \sim 3V$	
<i>HTO-12nm</i>	17,9	Effacement trop long	~ 80%	—
<i>HTO/Al₂O₃ 4nm/8nm</i>	13,9	16V-1ms -16V-30ms	~ 27%	—
SANOS : <i>SiO₂-3,5nm/Si₃N₄-t_{SiN}/Al₂O₃-t_{Al₂O₃}</i>			$V_{Tinit} \sim 6V$	$W/E_{1cyc.} \rightarrow W/E_{10^4 cyc.}$
<i>t_{SiN}/t_{Al₂O₃}</i>				
<i>6nm/16nm</i>	13,4	18V-3ms -18V-3ms	~ 84%	16V-10ms/-16V-30ms 6,5V/3,8V → 8,7V/6,9V
<i>6nm/12nm</i>	11,8	16V-10ms -16V-20ms	~ 83%	14V-10ms/-14V-30ms 5,5V/3,5V → 6,9V/5,4V
<i>6nm/20nm</i>	14,9	18V-10ms -20V-20ms	~ 89%	17,8V-10ms/-17,8V-30ms 7,5V/4,4V → <i>claquer</i>
<i>10nm/16nm</i>	15,8	19V-3ms -19V-10ms	~ 87%	18,9V-10ms/-18,9V-30ms 7,7V/4,9V → 6,6V/9,8V
<i>3nm/16nm</i>	11,6	16V-5ms -16V-10ms	~ 78%	13,9V-10ms/-13,9V-30ms 5,6V/3,2V → 7,5V/5,6V
TANOS : <i>SiO₂-3nm/Si₃N₄-6nm/Al₂O₃-15nm</i>			$\Delta V_{Tinit} \sim 5V$	
<i>TaN</i>	~ 12	16V-1ms -16V-4ms	~ 80%	16V-500μs/-18V-10ms 5,8V/0,3V → 6,9V/3,1V
<i>TiN</i>	~ 12	16V-1ms -16V-4ms	~ 80%	16V-500μs/-18V-10ms 5,9V/0,3V → 7,0V/3,0V
SAONHOS : <i>SiO₂/HfSiON/Si₃N₄/HTO/Al₂O₃</i>			$\Delta V_{Tinit} \sim 4V$	
<i>1/5/6/4/8nm</i>	13,9	14V-1ms -14V-1ms	~ 54%	14V-1ms/-14V-1ms 5,3V/1,9V → 5,3V/2V

TAB. 4.4: Résumé des propriétés des mémoires de type SONOS étudiées

Conclusion générale

Le travail que nous venons de présenter a porté sur l'étude de l'intégration des diélectriques *High- κ* dans les mémoires Flash non-volatiles.

Conclusions

Dans le **Chapitre 1**, nous avons présenté le contexte actuel et les perspectives du marché des mémoires Flash. Nous avons vu que les mémoires non-volatiles EEPROM-Flash sont confrontées à de nombreuses difficultés qui nécessitent une rupture technologique et l'utilisation de nouvelles approches. Dans ce contexte, les mémoires à couche de piégeage discret sont de plus en plus envisagées de même que l'intégration de matériaux *High- κ* dans l'empilement interpoly. L'un des matériaux *High- κ* les plus envisagés est l'alumine et son optimisation fait l'objet de nombreuses études. De même, les mémoires à couche de piégeage discret, notamment les empilements de type TANOS ($TaN/Al_2O_3/Si_3N_4/SiO_2/Si$), font l'objet d'une attention particulière de la part des industriels. L'augmentation des propriétés d'injection de l'oxyde tunnel est également une piste très étudiée. Dans ce contexte, nous nous sommes intéressés à l'intégration des matériaux *High- κ* dans les différentes couches de l'empilement, pour la compréhension des phénomènes physiques, en vue de l'optimisation des performances mémoires.

Dans le **Chapitre 2**, nous avons introduit l'ensemble des éléments théoriques permettant l'analyse des résultats expérimentaux qui ont été exposés dans les chapitres suivants. Aussi, nous avons tout d'abord présenté des rappels sur les modes de conduction électronique dans les diélectriques telle que la conduction *tunnel*, Fowler-Nordheim ou les conductions assistées par pièges (conduction Poole-Frenkel, TAT, ..). Puis différentes approches pour modéliser le piégeage de charges dans ces couches ont été présentées. Enfin, en s'appuyant sur ces notions, nous avons présenté deux modèles de piégeage pour les mémoires de type SONOS, à injection Fowler-Nordheim. Ces modèles étant complémentaires, ils permettent de couvrir un grand spectre de phénomènes physiques. Ainsi, ils nous ont permis une meilleure compréhension des mécanismes d'écriture, d'effacement et de rétention, dans les chapitres suivants.

Dans le **Chapitre 3**, nous avons étudié les propriétés des matériaux *High- κ* en vue de leur intégration dans des mémoires non-volatiles. Nous avons mis en évidence l'intérêt des aluminates d'hafnium grâce au contrôle de leurs propriétés physiques par l'ajustement de la concentration en Hf. Pour ces matériaux, nous avons aussi mis en évidence des charges fixes négatives et un piégeage aux interfaces *HfAlO/SiO₂*.

De plus, dans le cas d'un empilement tricouche, nous avons observé que l'épaisseur du *High- κ* ne modifiait pas les courants à V_G/EOT constant. Nous avons aussi montré que pour réduire l'*EOT* d'un diélectrique d'interpoly tricouche, il est possible de réduire fortement, voir de supprimer, le *HTO* côté anode sans dégradation notable des courants de fuite.

Nous avons aussi mis en évidence l'importance de cristalliser l'alumine après son dépôt pour assurer une bonne compatibilité avec une grille en *TiN*.

Les silicates d'hafnium sont apparus comme de bons candidats pour être intégrés en tant qu'oxyde *tunnel*, en particulier les structures *SiO₂/HfSiON* ont présentées une très bonne sensibilité au champ.

Dans le **Chapitre 4**, nous avons étudié l'intérêt des matériaux *High- κ* intégrés dans des cellules mémoires à nanocristaux et à couche de piégeage nitrure pour les applications en programmation Fowler-Nordheim. Tout d'abord nous avons démontré la nécessité d'utiliser des diélectriques d'interpoly à base de *High- κ* pour permettre la programmation des mémoires à nanocristaux en Fowler-Nordheim. Les tricouches, *HTO/HfAlO/HTO*, permettent de bons compromis entre tensions de programmation et rétention.

Pour les mémoires à couche de piégeage nitrure, nous avons mis en évidence que l'utilisation d'oxyde bloquant à base de *High- κ* comme l'*Al₂O₃* est indispensable, pour assurer une bonne écriture mais surtout un bon effacement. De plus, l'augmentation des vitesses d'écriture et d'effacement est fortement impactée par l'épaisseur de l'alumine et/ou du nitrure. En effet, il apparaît que les couches fines présentent une moins bonne efficacité de piégeage pouvant dégrader les performances d'endurance. En rétention, on retiendra que le principal mécanisme à température ambiante est l'émission des charges initialement piégées aux interfaces, donnant au nitrure épais une meilleure rétention. A hautes températures, les émissions thermo-assistées diminuent les bénéfices des nitrures épais. On notera aussi que l'utilisation d'un bicouche *SiO₂/Al₂O₃* permet une importante amélioration des caractéristiques de rétention par rapport à une monocouche d'alumine sans dégrader l'*EOT*.

De plus, nous avons vu que l'un des principaux problèmes des mémoires SANOS est le compromis entre rétention et effacement. Il est donc impératif qu'aucun mécanisme ne vienne ralentir l'effacement. Il est donc nécessaire d'intégrer des grilles à fort travail de sortie (*TiN*, *TaN*, *TaAlN*). Enfin, les bicouches *SiO₂/HfSiON* intégrés en oxyde *tunnel*, ont montré une forte sensibilité aux champs, leur permettant une réduction des tensions de programmation et l'amélioration de l'endurance.

Les matériaux *High- κ* sont donc inévitables dans la poursuite de la miniaturisation des mémoires Flash. L'alumine est le candidat le plus présenté en tant que diélectrique d'interpoly pour les prochaines générations. De même, l'intégration de grille métallique de type *TiN/TaN* (mid gap) est suffisant pour limiter la saturation de l'effacement. Pour les générations plus avancées, l'utilisation de matériaux à constante diélectrique plus élevée est à envisager. Dans ce cadre les aluminates d'hafnium sont un bon compromis entre couplage et rétention. Enfin, l'ingénierie de l'oxyde *tunnel* est l'une des solutions les plus prometteuses pour la réduction des tensions de programmation.

Perspectives

Cette étude a montré que les problèmes de fiabilité (essentiellement rétention à haute température) peuvent être en grande partie résolus grâce à l'optimisation de l'empilement de grille. Par conséquent, les mémoires de type TANOS constituent une solution très prometteuse pour les applications de stockage de masse. En revanche, les tensions mises en jeu sont trop élevées ($\sim 18V$) pour permettre une utilisation de ces mémoires en système embarqué. Une voie pour ouvrir ce domaine d'application a concerné l'utilisation de matériaux *High- κ* dans l'empilement *tunnel*, permettant de réduire de plusieurs volts les tensions de programmation. Cependant les performances en rétention sont encore peu satisfaisantes. Ainsi l'une des premières perspectives serait la poursuite de l'optimisation de l'empilement *tunnel*.

Un travail sur l'amélioration de la rétention des approches BE-SONOS peut s'orienter autour de deux axes :

- Ingénierie de l'empilement *tunnel*, en travaillant sur les épaisseurs, les matériaux *High- κ* utilisés et la structure pour des dispositifs complets.
- Ingénierie de la couche de piégeage, en utilisant des matériaux possédant des écarts plus importants entre leurs bandes de conduction, comme par exemple les *HfAlO* ou l'*AlN*. Un piégeage à des niveaux énergétiques plus profonds peut aussi être une solution intéressante.

La poursuite de l'ingénierie commencée sur les grilles de contrôle et leurs impacts sur la fiabilité est donc à continuer. Il est impérative d'éduiter en profondeur la compatibilité entre les couches de l'empilement de grille, notamment avec l'intégration de nouveaux métaux. De même, l'impact des traitements thermiques sur l'alumine reste à approfondir.

Enfin, d'un point de vue modélisation, il serait intéressant d'allier les avantages des deux modèles utilisés pour avoir un modèle plus général prenant en compte les électrons, les trous et leurs transports dans le nitrure. De même, il serait pertinent de prendre en compte les mécanismes assistés par pièges et le piégeage parasite dans l'oxyde de contrôle qui jouent un rôle clé dans la fiabilité de ces mémoires.

On peut malgré tout s'interroger sur les limites intrinsèques des mémoires à couche de piégeage intégrant des *High- κ* . Avec l'avancée des générations technologiques, il est de plus en plus difficile d'allier tensions de programmation faibles et fiabilité (rétention, endurance...). Aussi, l'investigation d'autres approches est essentielle pour préparer l'avenir de l'industrie des mémoires, comme les PCRAMs (mémoires résistives), semblent très prometteuses. Un effort de développement est donc à réaliser pour rendre leur technologie industrialisable.

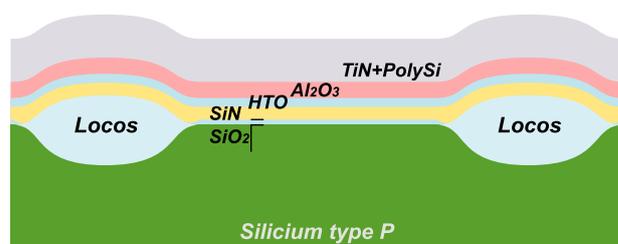
Annexe A

Procédé de fabrication d'une cellule mémoire réalisée au Létì



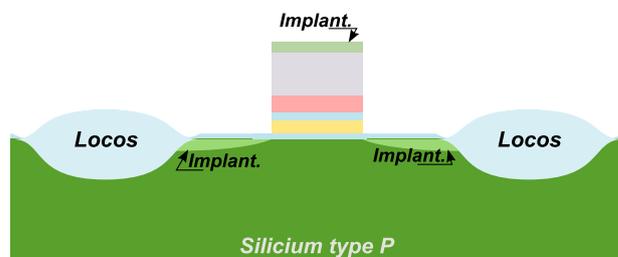
1- Préparation du substrat

- Utilisation substrat type p
- Isolation Locos
- Implantation caisson



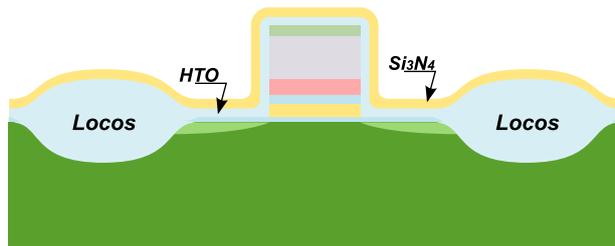
2- Réalisation de l'empilement mémoire (1/2)

- Dépôt des diélectriques
exemple : SiO₂~3,5nm/SiN~6nm/HTO~4nm/Al₂O₃~8nm
- Dépôt de grille
exemple : TiN~10nm/Poly-Si~200nm



3- Réalisation de l'empilement mémoire (2/2)

- Photo-lithographie grille
- Gravure grille : arrêt sur SiO₂ tunnel
- Stripping résine
- Implantation (si nécessaire)



4- Réalisation des espaceurs (1/2)

→ Dépôt HTO~15nm LPCVD

L'épaisseur du dépôt doit être suffisante pour remplir les volumes générés par les gravures isotropes de l'empilement de grille

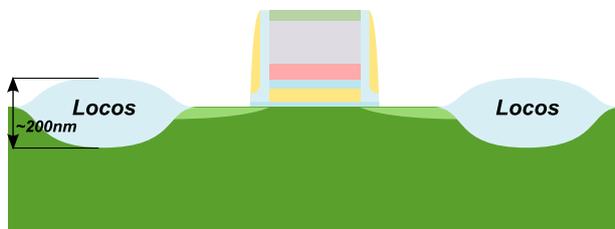
→ Dépôt Si₃N₄~30nm LPCVD

→ Photo-Lithographie : protection face avant

→ CMP face arrière : arrêt sur le substrat

Retrait des couches Nitrure/Oxyde/Poly-Si/empilement de grille déposées aussi en face arrière. On effectue ce retrait à cette étape car la face avant est protégé par un nitrure et qu'il n'y a plus de dépôt qui ne soit pas enlevé par chimie.

→ Stripping résine face avant

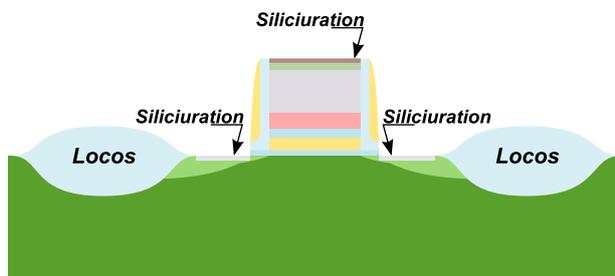


5- Réalisation des espaceurs (2/2)

→ Gravure espaceur Si₃N₄/HTO

→ Dépôt HTO~3nm LPCVD

Oxyde d'implantation



6- Implantation DSN/HFF & Siliciuration

→ Implantation

→ Recuit (950°C 15s)

avec le retrait de l'oxyde d'implantation ~3nm

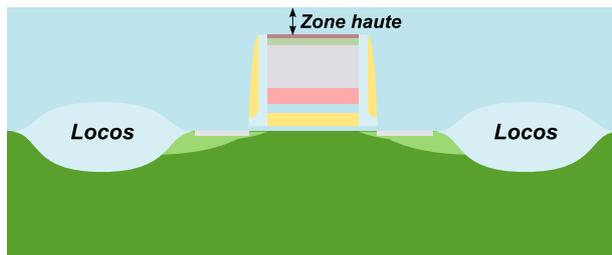
→ Siliciuration

Nettoyage, dépôt Ni~9nm et TiN~10nm, retrait sélectif du Ni et TiN qui n'ont pas réagi, et recuit de siliciuration.



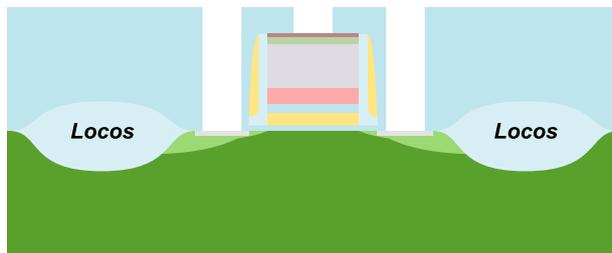
7- Niveau contact (1/4)

→ Dépôt oxyde de passivation ~1,1µm

**8- Niveau contact (2/4)**

→ **CMP de passivation**

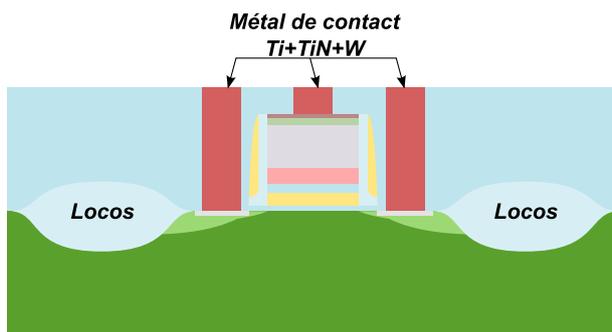
Il reste environ 450nm~500nm d'oxyde en zone haute

**9- Niveau contact (3/4)**

→ **Photo-lithographie contact**

→ **Gravure : arrêt sur NiSi**

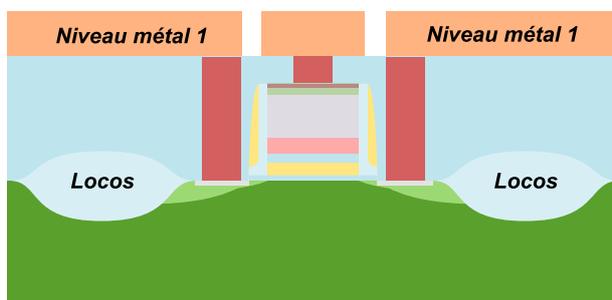
→ **Stripping**

**10- Niveau contact (4/4)**

→ **Remplissage**

Ti~40nm+TiN~60nm+W~500nm

→ **CMP du W**

**11- Niveau métal**

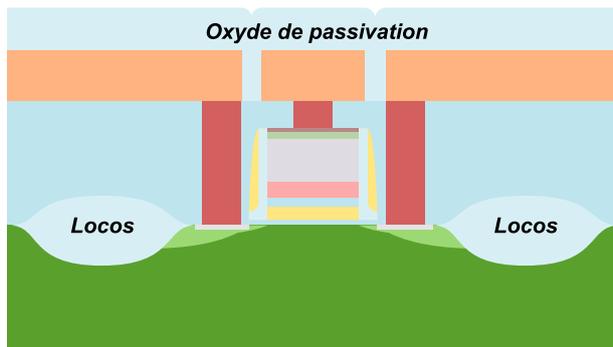
→ **Dépôt métal**

nettoyage, dépôt Ti~10nm/AICu~440nm/TiN~40nm

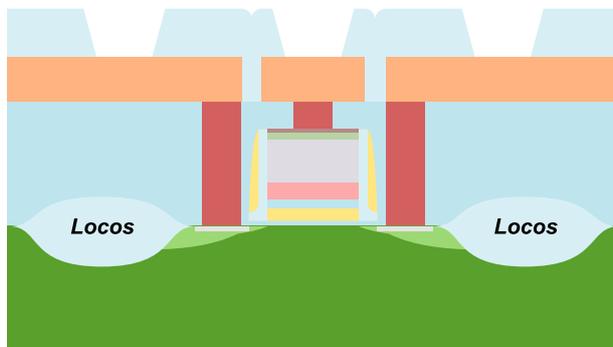
→ **Photo-lithographie métal**

→ **Gravure**

→ **Stripping**



12- Niveau plots (1/2)
→ Dépôt de passivation



13- Niveau plots (2/2)
→ Photo-lithographie plots
→ Gravure
→ Stripping
→ Recuit

Bibliographie

- [Akb04] Akbar M.S., Cho H.-J., Choi R., Kang C.S., Kang C.Y., Choi C.H., Rhee S.J., Kim Y.H. et Lee J.C., *Optimized NH₃ Annealing Process for High-Quality HfSiON Gate Oxide*, IEEE Electron Device Letters **25** (2004), 465–467.
- [Als06] Alshareef H.N., Choi K., Wen H.C., Luan H., Harris H., Senzaki Y., Majhi P., Lee B.H. et Jammy R., *Composition dependance of the work function of Ta_{1-x}Al_xN_y metal gates*, Applied Physics Letter **88** (2006), 072108.
- [Arn75] Arnett P.C, *Transient conduction in insulators at high fields*, Journal of Applied Physics **46** (1975), 5236–5243.
- [Arr07] Arreghini A., Akil N., Driussi F., Esseni D., Selmi L. et van Duuren M.J., *Characterization and Modeling of long term retention in SONOS Non Volatile Memories*, Proceedings of ESSDERC, 2007, pp. 406–409.
- [Arr08a] ———, *Long term charge retention dynamics of SONOS cells*, Solid-State Electronics **52** (2008), 1460–1466.
- [Arr08b] Arreghini A., Driussi F., Vianello E., Esseni D., van Duuren M.J., Golubović D.S., Akil N. et van Schaijk R., *Experimental Characterization of the Vertical Position of the Trapped Charge in Si Nitride-Based Nonvolatile Memory Cells*, IEEE Trans. on Electron Devices **255** (2008), 1211–1219.
- [Atw97] Atwood G., Fazio A., Mills D. et Reaves B., *Intel StrataFlash™ Memory Technology Overview*, Intel Technology Journal Q4 (1997).
- [Bac01] Bachhofer H., Reisinger H., Bertagnolli E. et von Philipsborn H., *Transient conduction in multielectric siliconoxidenitrideoxide semiconductor structures*, Journal of Applied Physics **89** (2001), 2791–2800.
- [Beu09] Beug M.F., Melde T., Paul J., Bewersdorff-Sarlette U., Czernohorsky M., Beyer V., Hoffmann R., Seidel K., Löhr D.A., Bach L., Knoefler R. et Tilke A.T., *Improvement of 48nm TANOS NAND Cell Performance by Introduction of a Removable Encapsulation Liner*, International Memory Workshop, 2009, pp. 88–89.
- [Blo09a] Blomme P., De Vos J. et van Houdt J., *Optimization of Al₂O₃ based VARIOT engineered tunnel dielectric for floating gate Flash scaling*, International Memory Workshop, 2009.
- [Blo09b] Blomme P. et van Houdt J., *Scalability of fully planar NAND Flach memory arrays below 45nm*, International Memory Workshop, 2009.
-

- [Boc08] Bocquet M., Molas G., Perniola L., Garros X., Buckley J., Gély M., Colonna J.P., Grampeix H., Martin F., Vidal V., Toffoli A., De Salvo B., Ghibaudo G., Pananakakis G. et Deleonibus S., *On the Role of a HTO/Al₂O₃ Bi-Layer Blocking Oxide in Nitride-Trap Non-Volatile Memories*, Proceedings of ESSDERC, 2008.
- [Boc09] Bocquet M., Vianello E., Molas G., Perniola L., Grampeix H., Martin F., Colonna J.P., Papon A.M., Brianceau P., Gély M., Ghibaudo G., Selmi L., Pananakakis G. et De Salvo B., *An in-depth investigation of physical mechanisms governing SANOS memories characteristics*, International Memory Workshop, 2009.
- [Boh04] Boher P., Evrard P., Condat O., Dos Reis C., Defranoux C., Piel J.P. et Bellandi E., *Automated metrology system including VUV spectroscopic ellipsometry and X-ray reflectometry for 300 mm silicon microelectronics*, Thin solid films **450** (2004), 114–119.
- [Bre08] Breuil L., Furnémont A., Rothschild A., Van den Bosch G., Cacciato A. et Van Houdt J., *Improvement of TANOS NAND Flash performance by the optimization of a sealing layer*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 126–127.
- [Bri97] Brière O., Barla K., Halimaoui A. et Ghibaudo G., *Oscillatory behavior of the tunneling current in ultra thin gate dielectrics : influence of various physical and technological parameters*, Solid-State Electronics **41** (1997), 987–990.
- [Bro97] Brown W.D. et Brewer J.E. (ed.), *Nonvolatile Semiconductor Memory Technology*, IEEE PRESS, 1997.
- [Buc05] Buckley J., De Salvo B., Molas G., Gély M. et Deleonibus S., *Experimental and theoretical study of layered tunnel barriers for nonvolatile memories*, Proceedings of ESSDERC (Grenoble - France), 2005.
- [Buc06] Buckley J., Bocquet M., Molas G., Gély M., Brianceau P., Rochat N., Martinez E., Martin F., Grampeix H., Colonna J.P., Toffoli A., Vidal V., Leroux C., Ghibaudo G., Pananakakis G., Bongiorno C., Corso D., Lombardo S., DeSalvo B. et Deleonibus S., *In-depth Investigation of Hf-based High-k Dielectrics as Storage Layer of Charge-Trap NVMs*, IEDM Tech. Dig., 2006.
- [Cac07] Cacciato A., Furnémont A., Breuil L., De Vos J., Haspelagh L., van Houdt J., *Effect of Al₂O₃ morphology on the erase saturation performance in SANOS-type memory cells*, International Conference on Memory Technology and Design, 2007, pp. 217–220.
- [Cha84] Chang C., *Tunneling in thin gate oxide MOS structures*, Ph.D. thesis, Université de Californie, Berkeley, 1984.
- [Cha08] Chang S.-I., Lee C.-H., ang C., Jeon S., Kim J., Choi B.-I., Park Y., Park J., Jeong W., You J., Choi B., Sel J., Sim J.S., Y. Shin, Choi J. et Lee W.-S., *Reliability Characteristics of TANOS (TaN/AlO/SiN/Oxide/Si) NAND Flash Memory with Rounded Corner (RC) Structure*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 117–118.
-

-
- [Chi05] Chin A., Laio C.C., Chen C., Chiang K.C., Yu D.S., Yoo W.J., Samudra G.S., Wang T., Hsieh I.J., McAlister S.P. et Chi C.C., *Low Voltage High Speed SiO₂/AlGaN/AlLaO₃/TaN Memory with Good Retention*, IEDM Tech. Dig., 2005.
- [Cho97] Chou A.I., Lai K., Kumar K., Chowdhury P. et J.C. Lee, *Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism*, Applied Physics Letter **70** (1997), 3407–3409.
- [Cho05] Choi S., Choi H., Kim T.-W., Yang H., Lee T., Jeon S., Kim C. et Hwang H., *High density silicon nanocrystal embedded in SiN prepared by low energy (<500eV) SiH₄ plasma immersion ion implantation for non-volatile memory applications*, IEDM Tech. Dig., 2005, pp. 173–176.
- [Cho08] Choi S., Baik S.J. et Moon J.-T., *Band Engineered Charge Trap NAND Flash with sub-40nm Process Technologies*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 925–928.
- [Cho09] ———, *Band Engineered Charge Trap NAND Flash with sub-40nm Process Technologies*, IEDM Tech. Dig., 2009, pp. 925–928.
- [Cle01] Clerc R., *Étude des effets quantiques dans les composants CMOS oxydes de grille ultra-minces – modélisation et caractérisation*, Ph.D. thesis, Institut National Polytechnique de Grenoble, 2001.
- [Coi09] Coignus J., Clerc R., Ghibaudo G. et Boulanger F., *Analytical modeling of tunnel current through SiO₂ – HfO₂ stacks in metal oxide semiconductor structures*, Journal of Vacuum Science Technology B **27** (2009), 338–345.
- [Com03] Compagnoni C. M., Ielmini D., Spinelli A. S., Lacaïta A. L., Previtali C. et Gerardi C., *Study of data retention for nanocrystal Flash memories*, IEEE Annual International Reliability Physics Symposium, 2003, pp. 506–512.
- [De 99] De Salvo B., *Etude du transport électrique et de la fiabilité des isolants des mémoires non volatiles à grille flottante*, Ph.D. thesis, École doctorale EEATS-INPG, 29 Janvier 1999.
- [De 01] De Salvo B., Ghibaudo G., Pananakakis G., Masson P., Baron T., Buffet N., Fernandes A. et Guillaumot B., *Experimental and Theoretical Investigation of Nano-Crystal and Nitride-Trap Memory Devices*, IEEE Trans. on Electron Devices **48** (2001), 1789–1799.
- [DeS07] DeSalvo B., 2007.
- [Doh03] Doh S.J., Jung H.-S., Kim Y.-S., Lim H.-J., Kim J.P., Lee J.H., Lee J.-H., Lee N.-I., Kang H.-K., Suh K.-P., Park S.G., Kang S.B., Choi G.H., Chung Y.-S., Baik H.-S., Chaug H.-S., Cho M.-H., Moon D.-W., Park H. B., Cho M. et Hwang C.S., *Improvement of NBTI and Electrical Characteristics by Ozone Pre-treatment and PBTI issues in HfAlO(N) High-κ Gate Dielectrics*, IEDM Tech. Dig., 2003, pp. 943–946.
- [Eng07] Engstrom O., Raeissi B., Hall S., Bui O., Lemme M.C., Gottlob H.D.B., Hurley P.K. et Cherkaoui K., *Navigation aids in the search for future high-k dielectrics : Physical and electrical trends*, Solid-State Electronics **51** (2007), 622–626.
-

-
- [Eri97] Ericsson P., Bengtsson S. et Skarp J., *Properties of Al₂O₃-films deposited on silicon by atomic layer epitaxy*, MicroElectronic Engineering **36** (1997), 91–94.
- [Fab08] <http://www.fabtech.org/content/view/full/3046/132/>, 2008.
- [Fis07] Fischer D. et Kersch A., *Ab initio study of high permittivity phase stabilization in HfSiO*, MicroElectronic Engineering **84** (2007), 2039–2042.
- [Fow28] Fowler R.H. et Nordheim L., *Electron emission in intense electric fields*, Containing Papers of a Mathematical and Physical Character **119** (1928), 173–181.
- [Fre38] Frenkel J., *On pre-breakdown phenomena in insulators and electronic semiconductors*, Phys. Rev **54** (1938), 647–648.
- [Fuk07] Fukuzumi Y., Matsuoka Y., Kito M., Kido M., Sato M., Tanaka H., Nagata Y., Iwata Y., Aochi H. et Nitayama A., *Optimal Integration and Characteristics of Vertical Array Devices for Ultra-High Density, Bit-Cost Scalable Flash Memory*, IEDM Tech. Dig., 2007, pp. 449–452.
- [Fur06] Furnémont A., Rosmeulen M., Van Houdt J., De Meyer K. et Maes H., *Model for electron redistribution in silicon nitride*, Proceedings of ESSDERC, 2006.
- [Fur07] Furnémont A., Rosmeulen M., Cacciato A., Breuil L., De Meyer K., Maes H. et Van Houdt J., *A consistent model for the SANOS programming operation*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2007, pp. 96–97.
- [Fur08] Furnémont A., Cacciato A., Breuil L., Rosmeulen M., Maes H., De Meyer K. et Van Houdt J., *Physical understanding and modeling of SANOS retention in programmed state*, Solid-State Electronics **52** (2008), 577–583.
- [Gar04] Garros X., *Caractérisation et modélisation de l'oxyde d'hafnium comme alternative la silice pour les futures technologies CMOS submicroniques*, Ph.D. thesis, Université de Provence-Aix-Marseille, 2004.
- [Gar07] Garros X., Mitard J., Leroux C., Reibold G. et Boulanger F., *In depth analysis of V_T instabilities in HfO₂ technologies by charge pumping measurements and electrical modeling*, IEEE Annual International Reliability Physics Symposium, 2007.
- [Gav05] Gavartin J.L., Shluger A.L., Foster A.S. et Bersuker G.I., *The role of nitrogen-related defects in High- κ dielectric oxides : Density-functional studies*, Journal of Applied Physics **97** (2005), 053704.
- [Gav06] Gavartin J.L., Muñoz Ramo D., Shluger L., Bersuker G. et Lee B.H., *Negative oxygen vacancies in HfO₂ as charge traps in High- κ stacks*, Applied Physics Letter **89** (2006), 082908.
- [Geh03] Gehring A., Jiménez-Molinos F., Kosina H., Palma A., Gámiz F. et Selberherr S., *Modeling of retention time degradation due to inelastic trap-assisted tunneling in EEPROM devices*, Microelectronic Reliability **43** (2003), 1495–1500.
- [Geh04] Gehring A. et Selberherr S., *Modeling of Tunneling Current and Gate Dielectric Reliability for Nonvolatile Memory Devices*, IEEE Trans. on Device and Materials Reliability **4** (2004), 306–319.
-

-
- [Ger08] Gerardi C., Molas G., Albin G., Tripiciano E., Gely M., Emmi A., Fiore O., Nowak E., Mello D., Vecchio M., Masarotto L., Portoghese R., De Salvo B., Deleonibus S.,Maurelli A., *Performance and reliability of a 4Mb Si nanocrystal NOR Flash memory with optimized 1T memory cells*, IEDM Tech. Dig., 2008.
- [Gil09] Gilmer D.C., Goel N., Park H., Park C., Barnett J., Kirsch P.D. et R. Jammy, *High Work-function Oxygen-Bearing Electrodes for Improved Performance in MANOS Charge-Trap NVM and MIM-DRAM Type Devices*, International Memory Workshop, 2009, pp. 90–91.
- [Gou03] Gougousi T., Niu D.N., Ashcraft R.W. et Parsons G.N., *Carbonate formation during post-deposition ambient exposure of High- κ dielectrics*, Applied Physics Letter **83** (2003), 3543–3545.
- [Gov03] Govoreanu B., Blomme P., Rosmeulen M., Van Houdt J. et De Meyer K., *VARIOT : A Novel Multilayer Tunnel Barrier Concept for Low-Voltage Non-volatile Memory Devices*, IEEE Electron Device Letters **24** (2003), 99–101.
- [Gov06] Govoreanu B., Wellekens D., Haspelslagh L., De Vos J. et Van Houldt J., *Investigation of the low-field leakage through High- κ interpoly dielectric stacks and its impact on nonvolatile memory data retention*, IEDM Tech. Dig., IMEC, 2006, pp. 479–482.
- [Gra07] Grampeix H., J.-P. Colonna, G. Molas, M. Bocquet, F. Martin, N. Rochat, E. Martinez, C. Licitra, T. Veyron, A.-M. Papon, M. Gely et K. Yckache, *Effect of Nitridation for High- k layers by ALCVD in order to decrease the Trapping in Non Volatile Memories*, Proc. Of the ECS Meeting, vol. 11, 2007, pp. 213–225.
- [Gri03] Gritsenko V.A., Nasyrov K.A., Novikov Y.N., Aseev A.L., Yoon S.Y., Lee J.-W., Lee E.-H., Kim C.W., *A new low voltage fast SONOS memory with High textit- κ dielectric*, Solid-State Electronics **47** (2003), 1651–1656.
- [Gro02] Groner M.D., Elam J.W., Fabreguette F.H. et George S.M., *Electrical characterization of thin Al_2O_3 films grown by atomic layer deposition on silicon and various metal substrates*, Thin solid films **413** (2002), 186–197.
- [Gu 07] Gu S.-H., Hsu C.-W., Wang T., W.-P. Lu, Ku Y.-H. J., Lu C.-Y., *Numerical simultation of bottom oxide thickness effect on charge retention in SONOS flash memory cells*, IEEE Trans. on Electron Devices **54** (2007), 90–97.
- [Ha 02] Ha W., Choo M.H., et Im S., *Poole-Frenkel conduction in amorphous solids*, Journal of Non-Crystalline Solids **303** (2002), 78–82.
- [Hau00] Haukka S., Tuomienen M. et Granneman E., *Atomic Layer Chemical Vapor Deposition of High- κ gate dielectrics*, Semicon Europa / Semieducation, 2000.
- [Her95] Herrmann M. et Schenk A., *Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories : Measurement and modeling*, Journal of Applied Physics **77** (1995), 4522–4540.
- [Hes86] Hesto P., *Instabilities in silicon devices*, Natural or electronic conduction, 1986.
- [Hil71] Hill R.M., *Poole-Frenkel conduction in amorphous solids*, The Philosophical Magazine **23** (1971), 59–86.
-

-
- [Ho 02] Ho M.-Y., H. Gong, G.D. Wilk, B.W. Busch, M.L. Green, W.H. Lin, A. See, S. K. Lahiri, M.E. Loomans, Petri I. Risnen, and T. Gustafsson, *Morphology and crystallization kinetics in HfO₂ thin films grown by atomic layer deposition*, Applied Physics Letter **93** (2002), 1477–1481.
- [Hou03] Hou Y.T., et al, , IEEE Electron Device Letters **24** (2003), 96.
- [Huo07] Huo Z.L., Yang J.K., Lim S.H., Baik S.J., Lee J., Han J.H., Yeo I.-S., Chung U.-I., Moon J.T. et Ryu B.I., *Band Engineered Charge Trap Layer for highly Reliable MLC Flash Memory*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2007, pp. 138–139.
- [Huo09] ———, *A study of stored charge interference and fringing field effects in sub-30nm charge-trapping NAND Flash*, International Memory Workshop, 2009, pp. 34–35.
- [Ice08] <http://www.ice-corp.com/blogs.aspx?id=5418&blogid=86>, 2008.
- [ICI07] *The McClean Report 2007 edition : an in-depth analysis and forecast of the integrated circuit industry*, Tech. report, IC Insights, 2007.
- [ICI09] *The McClean Report 2009 edition*, Tech. report, IC Insights, 2009.
- [Iel00] Ielmini D., Spinelli A., Rigamonti M.A. et Lacaita A.L., *Modeling of SILC Based on Electron and Hole Tunneling - Part I : Transient Effects*, IEEE Trans. on Electron Devices **47** (2000), no. 6, 1258–1265.
- [Ito04] Ito T., Kato H., Nango T. et Ohki Y., *Energy Band Profile of Hafnium Silicates Estimated by X-Ray Photoelectron Spectroscopy*, Japanes Journal of Applied Physics **43** (2004), 8199–8202.
- [ITR07] ITRS, 2007.
- [Jac07] Jacob S., De Salvo B., Perniola L., Festes G., Bodnar S., Coppard R., Thierry J.F., Pate-Cazal T., Bongiorno C., Lombardo S., Dufourcq J., Jalaguier E., Pedron T., Boulanger F. et Deleonibus S., *Integration of CVD Silicon Nanocrystals in 32Mb NOR Flash Memory*, Solid-State Electronics (2007).
- [Jac08] Jacob Stéphanie, *Intégration, caractérisation et modélisation des mémoires non-volatiles avancées à sites de stockage discrets*, Ph.D. thesis, Université de Provence – Aix-Marseille I, 2008.
- [Jah05] Jahan C., Faynot O., Cassé M., Ritzenthaler R., Brévard L., Tosti L., Garros X., Vizioz C., Allain F., Papon A.M., Dansas H., Martin F., Vinet M., Guillaumot B., Toffoli A., Giffard B. et Deleonibus S., *ΩFETs transistors with TiN metal gate and HfO₂ down to 10nm*, 2005, pp. 112–113.
- [Jah08] Jahan C., Nowak E., Perniola L., Gély M., Molas G., Lombardo S., De Salvo B. et Deleonibus S., *Performance enhancement in ultra-scaled SONOS FinFlash by inclusion of high-k dielectric in the gate stack*, 2008, pp. 106–108.
- [Jam07] James D., *Nano-Scale Flash in the Mid-Decade*, 2007, pp. 371–376.
- [Jeo06] Jeon S. et Kim C., *The effect of fixed charge in Al₂O₃ blocking dielectric on memory properties of charge trap flash memory devices*, vol. 9, 2006, pp. 265–267.
-

- [Joh02] Johson R.S., Hong J.G., Hinkle C., Lucovsky G., *Electron trapping in noncrystalline remote plasma deposited Hf-aluminate alloys for gate dielectric applications*, Journal of Vacuum Science Technology B **20** (2002), 1126–1131.
- [Jun02] Jung H.-S., Kim Y.-S., Kim J.P., Lee J.H., Lee J.-H., Lee N.-I., Kang H.-K., Suh K.-P., Ryu H.J., Oh C.-B., Kim Y.-W., Cho K.-H., Baik H.-S., Chung Y.S., Chang H.S. et Moon D.W., *Improved Current Performance of CMOSFETs with Nitrogen Incorporated HfO₂-Al₂O₃ Laminate Gate Dielectrique*, 2002, pp. 853–856.
- [Jun06] Jung S.-M., Jang J., Cho W., Cho H., Jeong J., Chang Y., Kim J., Rah Y., Son Y., Park J., Song M.-S., Kim K.-H., Lim J.-S., et Kim K., *Three Dimensionally Stacked NAND Flash Memory Technology Using Stacking Single Crystal Si Layers on ILD and TANOS Structure for Beyond 30nm Node*, 2006.
- [Kah67] Kahng D. et Sze S.M., *A floating gate and its application to memory devices*, Bell Systems Technology Journal **46** (1967), 1288–1295.
- [Kam05] Kamimuta Y., Koike M., Ino T., Suzuki M., Koyama M., Tsunashima Y. et Nishiyama A., *Determination of Band Alignment of Hafnium Silicon Oxynitride/Silicon (HfSiON/Si) Structures using Electron Spectroscopy*, Japanes Journal of Applied Physics **44** (2005), 1301–1305.
- [Kan07] Kang C., Choi J., Sim J., Lee C., Shin Y., Park J., Sel J., Jeon S., Park Y. et Kim K., *Effets of Lateral Charge Spreading on the Reliability of TANOS (TaN/AlO/SiN/Oxide/Si) NAND Flash Memory*, 2007.
- [Kar07] Karim Z., Barbar G., Boissière O., Lehnen P., Christoph L., Seidel T, Adelman, Conard T., O’Sullivan B., Ragnarsson L.-A., Schram T., van Elshocht S. et De Gendt S., *AVD and MOCVD TaCN-based Films for Gate Metal Applications on High- κ Gate Dielectrics*, 2007.
- [Kat02] Kato H., Nango T., Miyagawa T., Katagiri T., Seol K.S. et Ohki Y., *Plasma-enhanced chemical vapor deposition and characterization of high-permittivity hafnium and zirconium silicate films*, Journal of Applied Physics **92** (2002), 1106.
- [Kim02] Kim J.B., Kwon D.R., Chakrabarti K., Lee C., Oh K.Y. et Lee J.H., *Improvement in Al₂O₃ dielectric behavior by using ozone as oxidant for the atomic layer deposition technique*, Journal of Applied Physics **92** (2002), 6739–6742.
- [Kim03] Kim Y., Lim C., Young C.D., Matthews K., Barnett J., Fornan B., Agarwal A., Brown G.A., Bersuker G., Zeitzoff P., Gardner M., Murto R.W., Lason L., Metzner C., Kher S. et Huff H.R., *Conventional poly-Si gate MOS-transistors with a novel, ultra-thin Hf-oxide layer*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2003, pp. 167–168.
- [Kim05] Kim K., *Technology for sub-50nm DRAM and NAND Flash Manufacturing*, IEDM Tech. Dig., 2005, pp. 333–336.
- [Kim07] Kim S., Cho W., Kim J., Lee B. et Park S., *Air-Gap Application and Simulation Results for Low Capacitance in 60nm NAND Flash Memory*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2007, pp. 54–55.
-

- [Kim08] Kim T.-K., Chang S.N., Hong S.W., Chae D.H., Lee K. et Choi J.-H., *16-Gigabit, 8-level NAND Flash Memory with 51nm 44-Cell String Technology*, Proceedings of ESSDERC, 2008, pp. 111–113.
- [Kit09] Kittl J. A., Opsomer K., Popovici M., Menou N., Kaczer B., P. Wang X., Adelman C., Pawlak M. A., Tomida K., Rothschild A., Govoreanu B., Degraeve R., Schaekers M., Zahid M., Delabie A., Meersschant J., Polspoel W., Clima S., Pourtois G., Knaepen W., Detavernier C., Afanas'ev V.V., Blomberg T., Pierreux D., Swerts J., Fischer P., Maes J.W., Manger D., Vandervorst W., Conard T., Franquet A., Favia P., Bender H., Brijs B., Van Elshocht S., Jurczak M., Van Houdt J., Wouters D.J., *High- κ dielectrics and metal gates for future generation memory devices*, 2009.
- [Kra05] Krautheim G., Hecht T., Jakschik S., Schröder U. et Zahn W., *Mechanical stress in ALD- Al_2O_3 films*, Applied Surface Science **252** (2005), 200–204.
- [Kwa07] Kwak D., Park J., Kim K., Yim Y., Ahn S., Park Y., Kim J., Jeong W., Kim J., Park M., Yoo B., Song S., Kim H., Sim J., Kwon S., Hwang B., Park H., Kim S., Lee Y., Shin H., Yim N., Lee K., Kim M., Lee Y., Park J., Park S., Jung J. et Kim K., *Integration Technology of 30nm Generation Multi-Level NAND Flash for 64Gb NAND Flash Memory*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2007.
- [Lai98] Lai S., *Tunnel oxide and ETOX Flash scaling limitation*, International Non Volatile Memory Conference, 1998, pp. 6–7.
- [Lai05] Lai C.H., Chin A., Chiang K.C., Yoo W.J., Cheng C. F., McAlister S.P., Chi C.C. et Wu P., *Novel $SiO_2/AlN/HfAlO/IrO_2$ Memory with Fast Erase, Large ΔV_{th} and Good Retention*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2005, pp. 210–211.
- [Lai06] Lai C.H., Chin A., Kao H.L., Chen K.M., Hong M., Kwo J. et Chi C.C., *Very Low Voltage $SiO_2/HfON/HfAlO/TaN$ Memory with Fast Speed and Good Retention*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2006, pp. 210–211.
- [Lai07] Lai S.C., Lue H.-T., Yang M.-J., Hsieh J.-Y., Wang S.-Y., Wu T.-B., Luo G.-L., Chien C.-H., Lai E.-K., Hsieh K.-Y., Liu R. et Lu C.-Y., *MA BE-SONOS : A Bandgap Engineered SONOS using Metal Gate and Al_2O_3 Blocking Layer to Overcome Erase Saturation*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2007, pp. 88–89.
- [Lar01] Larcher L., Paccagella A. et Ghidini G., *A Model of the Stress Induced Leakage Current in Gate Oxides*, IEEE Trans. on Electron Devices **48** (2001), 285–288.
- [Lar04] Larcher L., Pavan P. et Eitan B., *On the physical mechanism of the NROM memory erase*, IEEE Trans. on Electron Devices **51** (2004), 1593–1599.
- [Lee99] Lee B.H., Kang L., Qi W.-J., Nieh R., Jeon Y., Onishi K. et Lee J.C., *Ultrathin hafnium oxide low leakage and excellent reliability for alternative gate dielectric application*, IEDM Tech. Dig., 1999, pp. 133–136.
- [Lee00] Lee J. H., Koh K., Lee N.I., Cho M.H., Kim Y.K., Jeon J.S., Cho K.H., Shin H.S., Kim M. H., Fujihara K., Kang H.K. et Moon J.T., *Effect of polysilicon*
-

- gate on the flatband voltage shift and mobility degradation for *ald-al2o3* gate dielectric, IEDM Tech. Dig., 2000, pp. 645–648.
- [Lee02] Lee J.-H., Kim Y.-S., Jung H.-S., Lee J.-H., Lee N.-I., Kang H.-K., Ku J.-H., Kang H.S., Kim Y.-K., Cho K.-H. et Suh K.-P., *Poly-Si gate CMOSFETs with HfO₂-Al₂O₃ laminate gate dielectric for low power application*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2002, pp. 84–85.
- [Lee03] Lee C.H., Choi K.I., Cho M.K., Song Y.H., Park K.C. et Kim K., *A Novel SONOS Structure of SiO₂/SiN/Al₂O₃ with TaN metal gate for multi-giga bit flash memorie*, IEDM Tech. Dig., 2003, pp. 613–616.
- [Lee06] Lee C.-H., Choi J., Kang C., Shin Y., Lee J.-S., Sel J., Sim J., Jeon S., Choe B.-I., Bae1 D., Park K. et Kim K., *Multi-Level NAND Flash Memory with 63 nm-node TANOS (Si-Oxide-SiN-Al₂O₃-TaN) Cell Structure*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2006.
- [Lee08] Lee C.H., Wu C.W., Lin S.W., Yeh T.H., Gu S.H., Chen K.F., Chen Y.J., Hsieh J.Y., Huang I.J., Zous N.K., Han T.T., Chen M.S., Lu W.P., Tahui W. et Lu C.Y., *Numerical Simulation of Programming Transient Behavior in Charge Trapping Storage Memory*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 109–110.
- [Lef07] Lefez B., Lopitiaux J., Hannoyer B., Bacos M.-P. et Beucher E., *Fourier transform infrared reflection absorption spectroscopy and microspectroscopy, a tool to investigate thermally grown oxide scales*, Thin solid films **516** (2007), 349–354.
- [Ler04] Leroux C., Mitard J., Ghibaudo G., Garros X., Reimbold G., Guillaumot B., Martin F., *Characterization and modeling of hysteresis phenomena in High- κ dielectrics*, IEDM Tech. Dig., 2004.
- [Les02] Leskelä M. et Ritala M., *Atomic Layer Deposition (ALD) : from precursors to thin film structures*, Thin solid films **409** (2002), 138–146.
- [Li 06] Li Q., Koo K.M., Lau W.M, Lee P.F., Dai J.Y, Hou Z.F et Gong G., *Effects of Al addition on the native defects in hafnia*, Applied Physics Letter **88** (2006), 182903.
- [Li 09] Li S. et Zhang T., *Approching the information theoretical bound of multi-level NAND Flash memory storage Efficiency*, International Memory Workshop, 2009, pp. 46–48.
- [Lib90] Libsch F.R et White M.H, *Charge transport and storage of low programming voltage SONOS/MONOS mamory devices*, Solid-State Electronics **33** (1990), 105–126.
- [Lin08] Lin S.H., Chin A., Yeh F.S. et McAlister S.P., *Good 150 ° C Retention and Fast Erase Characteristics in Charge-Trap-Engineered Memory having a Scaled Si₃N₄ Layer*, 2008, pp. 843–846.
- [Liu09] Liu D. et Robertson J., *Oxygen vacancy levels and interfaces of Al₂O₃*, MicroElectronic Engineering **86** (2009), 1668.
- [Luc00] Lucovsky G. et Phillips J.C., *Limitations for aggressively scaled CMOS Si devices due to bond coordination constraints and reduced band offset energies at Si-high-k dielectric interfaces*, Applied Surface Science **166** (2000), 497–503.
-

- [Lue05] Lue H.-T., Wang S.-Y., Lai E.-K., Shih Y.-H., Lai S.-C., Yang L.-W., Chen K.-C., Ku J., Hsieh K.-Y., Liu R. et Lu C.-Y., *BE-SONOS : a Bandgap Engineered SONOS with excellent performance and reliability*, IEDM Tech. Dig., 2005.
- [Lue08a] Lue H.-T., Hsu T.-H., Lai S.C., Hsiao Y.H., Peng W.C., Liao C.W., Huang Y.F., Hong S.P., Wu M.T., Hsu F.H., Lien N.Z., Wang S.Y., Yang L.W., Yang T., Chen K.C., Hsieh K.Y., Liu R. et Lu C.-Y., *Scaling Evaluation of Be-SONOS NAND Flash Beyond 20nm*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2008.
- [Lue08b] Lue H.-T., Lai E.-K., Wang S.-Y., Yang L.-W., Yang T., Chen K.-C., Hsieh K.-Y., Liu R. et Lu C.-Y., *A Novel Gate-Injection Program/Erase P-Channel NAND-Type Flash Memory with High (10M Cycle) Endurance*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2008.
- [Lui97] Lui O. K. B. et Migliorato P., *A new generation-recombination model for device simulation including the Poole-Frenkel and phonon-assisted tunnelling*, Solid-State Electronics **41** (1997), 575–583.
- [Man01] Manchanda L., Morris M.D., Green M.L., van Dover R.B., Klemens F., Sorsch T.W., Silverman P.J., Wilk G., Busch B. et Aravamdhan S., *Multi-component High- κ gate dielectrics for the silicon industry*, MicroElectronic Engineering **59** (2001), 351–359.
- [Mas82] Maserjan J. et Zamani N., *Observation of positively charged state generation near the Si/SiO₂ interface during Fowler-Nordheim tunnelling*, 743–146.
- [Mat04] Mathieu H., *Physique des semiconducteurs et des composants électroniques*, Dunod, 2004.
- [McW90] McWhorter P.J., Miller S.L. et Dellin T.A., *Modeling the memory retention characteristics of silicon-nitride-oxide-silicon nonvolatile transistors in varying thermal environment*, Journal of Applied Physics **68** (1990), 1902–1909.
- [Mel08] Melde T., Beug M.F., Bach L., Riedel S., Ludwig C. et Mikolajick T., *Nitride thickness scaling limitation in TANOS charge trapping devices*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 130–132.
- [Men05] Menéndez-Proupin E. et Gutiérrez G., *Electronic properties of bulk $\gamma - Al_2O_3$* , 035116.
- [Mir06] Miranda A.H., van Schaijk R., van Duuren M., Akil N. et Golubović D.S., *Reliability Comparison of Al₂O₃ and HfSiON for use as Interpoly Dielectric in Flash Arrays*, Proceedings of ESSDERC, 2006.
- [Mol04] Molas G., Deleruyelle D., De Salvo B., Ghibaud G., Gély M., Jacob S., Lafond D. et Deleonibu S., *Impact of few electron phenomena on floating-gate memory reliability*, IEDM Tech. Dig., 2004, pp. 877–880.
- [Mol06] Molas G., Grampeix H., Buckley J., Bocquet M., Garros X., Martin F., Colonna J.P., Brianceau P., Vidal V., Gély M., De Salvo B., Bongiorno C., Lombardo S. et Deleonibus S., *In-depth Investigation of HfAlO Layers as Interpoly Dielectrics of Future Flash Memories*, Proceedings of ESSDERC, 2006.
-

- [Mol07] Molas G., Bocquet M., Buckley J., Colonna J.P., Masarotto L., Grampeix H., Martin F., Vidal V., Toffoli A., Brianceau P., Scheiblin P., Gély M., Papon A.M., Auvert G., Perniola L., Licitra C., Veyron T., Rochat N., Bongiorno C., Lombardo S., De Salvo B. et Deleonibus S., *Thorough investigation of Si-nc memories with high-k interpoly for sub-45nm node Flash NAND applications*, IEDM Tech. Dig., 2007.
- [Mor02] Morisaki Y., Aoyama T., Sugita Y., Irino K. Sugii T. et Nakamura T., *Ultra-thin ($t_{eff}^{inv} = 1.7nm$) poly-si-gated sin/hfO₂/sion high- κ strack dielectrics with high thermal stability (1050°C)*, 2002, pp. 861–864.
- [Mot71] Mott N.F. et Davis E.A., *Electronic processes in non-cristalline materials*, Clarendon press (1971).
- [Nar88] Naruke K., Taguchi S. et Wada M., *Stress Induced Leakage Current Limiting to Scale Down EEPROM Tunnel Oxide Thickness*, IEDM Tech. Dig., 1988, pp. 424–427.
- [Nas04] Nasyrov K.A., Gritsenko V.A., Novikov Y.U., Lee E.-H., Yoon S.Y. et Kim C.W., *Two-bands charge transport in silicon nitride due to phonon-assisted trap ionization*, Journal of Applied Physics **96** (2004), 4293–4296.
- [Nea00] Neaton J.B., Muller D.A. et Ashcroft N.W., *Electronic properties of the Si/SiO₂ interface from frist principles*, Physical Review Letters **85** (2000), 1298–1301.
- [Ngu05] Nguyen N.V., Sayan S., Levin I., Ehrstein J.R., Baumvol I.J.R., Driemeier C., Krug C., Wielunski L., Hung P.Y. et Diebold A., *Optical band gaps and composition dependence of hafnium aluminate thin films grown by atomic layer chemical vapor deposition*, Journal of Vacuum Science Technology A **23** (2005), 1706–1713.
- [Now08] Nowak E., Bocquet M., Perniola L., Ghibaudo G., Molas G., Jahan C., Kies R., Reimbold G., De Salvo B. et Boulanger F., *New Physical Model for ultra-scaled 3D Nitride-Trapping Non-Volatile Memories*, IEDM Tech. Dig., 2008.
- [Oh 04] Oh C.W., Suk S.D., Lee Y.L., Sung S.K., Choe J.-D., Lee S.-Y., Choi D.U., Yeo K.H., Kim M.S., Kim S.-M., Li M., Kim S.H., Yoon E.-J., Kim D.-W., Park D., Kim K. et Ryu B., *Damascene Gate FinFET SONOS Memory Implemented on Bulk Silicon Wafer*, IEDM Tech. Dig., 2004, pp. 893–896.
- [Oli88] Olivo P., Nguyen T.N. et Ricco B., *High-field-induced degradation in ultra-thin SiO₂ film*, IEEE Trans. on Electron Devices **35** (1988), 2259–2267.
- [Pan95] Pananakakis G., Ghibaudo G., Kies R. et Papadas C., *Temperature dependence of the Fowler-Nordheim current in metal oxides-degenerate semiconductor structures*, Journal of Applied Physics **78** (1995), 2635–2641.
- [Par04] Park J.-H., Hur S.-H., Lee J.-H., Park J.-T., Sel J.-S., Kim J.-W., Song S.-B., Lee J.-Y., Lee J.-H., Son S.-J., Kim Y.-S., Park M.-C., Chai S.-J., Choi J.-D., Chung U.-I., Moon J.-T., Kim K.-T., Kim K. et Ryu B.-I., *8Gb MLC (Multi-Level Cell) NAND Flash Memory using 63nm Process Technology*, IEDM Tech. Dig., 2004, pp. 873–876.
-

- [Par06] Park Y., Choi J., Kang C., Lee C., Shin Y., Choi B., Kim J., Jeon S., Sel J., Park J., Choi K., Yoo T., Sim J. et Kim K., *Highly Manufacturable 32Gb Multi - Level NAND Flash Memory with 0.0098 μm^2 Cell Size using TANOS(Si - Oxide - Al_2O_3 - TaN) Cell Technology*, IEDM Tech. Dig., 2006.
- [Pau06] Paul A., Sridhar C., Gedam S. and Mahapatra S., *Comprehensive Simulation of Program, Erase and Retention in Charge Trapping Flash Memories*, IEDM Tech. Dig., 2006, pp. 873–876.
- [Pow08] Power J.R., Shum D., Gong Y., Bogacz S., Haeupel J., Estel H., Strenz R., Kakoschke R., van der Zanden K., et Allinger R., *Improved Retention for a Al_2O_3 IPD Embedded Flash Cell without Top-Oxide*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 93–96.
- [Pra06] Prall K., *Scaling Non-Volatile Memory Below 30nm*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2006, pp. 5–10.
- [Pun04] Panchaipetch P., Okamoto T., Nakamura H., Uraoka Y., Fuyuki T. et Horii S., *Effect of nitrogen on electrical and physical properties of polyatomic layer Chemical Vapor Deposition HfSi_xO_y gate dielectrics*, Japanes Journal of Applied Physics **42** (2004), 7815–7820.
- [Ren02] Renault O., Gosset L.G., Rouchon D. et Ermolieff A., *Angle-resolved x-ray photoelectron spectroscopy of ultrathin Al_2O_3 films grown by atomic layer deposition*, Journal of Vacuum Science Technology A **20** (2002), 1867–1876, Etude Al_2O_3 spectroscopie X, intrinsieque defauts Al-OH, peut etre du a l'oxydation de l'interface qui fait diffuse ses Oxygene dans l'alumine pendant le RTA.
- [Rob00] Robertson John, *Band offsets of wide-band-gap oxides and implications for future electronic devices*, Journal of Vacuum Science Technology B **18** (2000), 1785–179.
- [Rob02] ———, *Band offsets of high dielectric constant gate oxides on silicon*, Journal of Non-Crystalline Solids **303** (2002), 94100.
- [Rob04] ———, *High dielectric constant oxides*, The European Physical Journal Applied Physics **28** (2004), 265–291.
- [Roc03] Rochat Névine, K. Dabertrand, V. Cosnier, S. Zoll, P. Besson et U. Weber, *Infrared spectroscopy of High-k thin layer by multiple internal reflection and attenuated total reflection*, Phys. Status Solidi (c) **0** (2003), 2961–2965.
- [San97] Sanchez-Portal D. et al., *Density-functional method for very large systems with LCAO basis sets*, Int. J. Quantum Chem **65** (1997), 453.
- [San09] Sankaran K. , Pourtois G., Rignanese G-M., Van Houdt J. et Kittl J., *First principles modelling of high- ϵ dielectric for flash memory applications*, CECAM, 2009.
- [Ser05] Servalli G., Brazzelli D., Camerlenghi E., Capetti G., Costantini S., Cupeta C., DeSimone D., Ghetti A., Ghilardi T., Gulli P., Mariani M., Pavan A. et Somaschini R., *A 65nm NOR Flash Technology with 0.042 μm^2 Cell Size for High Performance Multilevel Application*, IEDM Tech. Dig., 2005, pp. 849–852.
-

-
- [Shi98] Shih W.-K., Wang E.X., Jallepalli S., Leon F., Maziar C.M. et Tashjr A.F., *Modeling gate leakage current in nMOS structures due to tunneling through an ultra-thin oxide*, Solid-State Electronics **42** (1998), 97–106.
- [Shi05] Shin Y., Choi J., Kang C., Lee C., Park K.-T., Lee J.-S., Sel J., Kim V., Choi B., Sim J., Kim D., Cho H.-J. et Kim K., *A novel NAND-type MONOS memory using 63nm process technology for Multi-Gigabit Flash EEPROMs*, IEDM Tech. Dig., 2005.
- [Sho52] Shockley W. et Read, W.T., *Statistics of the recombinations of holes and electron*, Physical Review **87** (1952), 835–842.
- [Shu09] Shum D., Jaschke G., Canning M., Kakoschke R., Duschl R., Sikorski R., Erler F., Stiftinger M., Duch A., Power J.R., Tempel G., Strenz R. et Allinger R., *ALD-Al₂O₃ as Inter-Poly Dielectric for a Product Demonstrator in a proven eFlash Technology*, International Memory Workshop, 2009, pp. 42–45.
- [Son08a] Song Y.-C., Lui X.-Y., Du G., Kang J.-F. et Han R.-Q., *Carriers recombination processes in charge trapping memory cell by simulation*, Chinese Physics B **17** (2008), 2678–2682.
- [Son08b] Song Y.-C., Lui X.-Y., Wang Z.Y., Zhao K., Du G., Kang J.-F. et Han R.-Q., *Evaluating the effects of physical mechanisms on the program, erase and retention in the charge trapping memory*, International Conference on Simulation of Semiconductor Processes and Devices, 2008, pp. 41–44.
- [Spe04] Specht M., Kommling R., Dreeskornfeld L., Weher W., Hofmann F., Alvarez D., Kretz J., Luyken R.J., Rosner W., Reisinger H., Landgraf E., Schulz T., Hartwich J., Stadele M., Klandievski V., Hartmann E. et Risch L., *Sub-40nm tri-gate charge trapping nonvolatile memory cells for high-density applications*, IEEE Symposium On VLSI Technology Digest of Technical Papers, 2004, pp. 244–245.
- [Suz86] Suzuki E., Schroder D.K. et Y. Hayashi, *Carrier conduction in ultrathin nitrided oxide films*, Journal of Applied Physics **60** (1986), 36163621.
- [Tak99] Takagi S.-I., Yasuda N. et Toriumi Akira, *A new I-V model for stress-induced leakage current including inelastic tunneling*, IEEE Trans. on Electron Devices **46** (1999), 348–354.
- [Tak04] Takeuchi H., Wong H.Y., Ha D. et King T.-J., *Impact of Oxygen Vacancies on High-K Gate Stack Engineering*, IEDM Tech. Dig., 2004, pp. 829–832.
- [Tam84] Tam S., Ko P.-K. et Hu C., *Lucky-electron model of channel hot-electron injection in MOSFETs*, IEEE Trans. on Electron Devices **31** (1984), 1116–1125.
- [Tan98] Tang S., Wallace R.M., Seabaugh A. et King-Smith D., *Evaluating the minimum thickness of gate oxide on silicon using first-principles method*, Applied Surface Science **135** (1998), 137–142.
- [Tan04] Tan Y.N., Chim W.K., Choi W.K., Joo M.S., Ng T.H. et Cho B.J., *High- κ HfAlO Charge Trapping Layer in SONOS-type Nonvolatile Memory Device for High Speed Operation*, IEDM Tech. Dig., 2004.
-

- [Tan06] Tan Y.N., Chim W.K., Choi W.K., Joo M.S. et Cho B.J., *Hafnium Aluminum Oxide as Charge Storage and Blocking-Oxide Layers in SONOS-Type Nonvolatile Memory for High-Speed Operation*, IEEE Trans. on Electron Devices **53** (2006), 654–662.
- [Tom06] Tomida K., Kita K. et Toriumi A., *Dielectric constant enhancement due to Si incorporation into HfO_2* , Applied Physics Letter **89** (2006), 142902.
- [Tos08] http://www.solid-state.com/display_article/316150/5/none/none/TETAK/Toshiba,-SanDisk-ramping-43nm-NAND-flash-with-HK+MG,-3b/cell, 2008.
- [Tsa08] Tsai P.-H., Chang-Liao K.-S., Liu C.-Y., Wang T.-K., Tzeng P. J., Lin C.H., Lee L.S., et Tsai M.-J., *Novel SONOS-Type nonvolatile memory device with optimal Al doping in $HfAlO$ charge-trapping layer*, IEEE Electron Device Letters **29** (2008), 265–268.
- [Tse06] Tse K. et Robertson J., *Defect passivation in HfO_2 gate oxide by fluorine*, Applied Physics Letter **89** (2006), 142914.
- [Tsy90] Tsyganenko A.A., Mirnov K.S.S., Rzhetskij A.M. et Mardilovich P.P., *Infrared spectroscopic evidence for the structural OH groups of spinelle alumina modifications*, Mater. Chem. Phys. **26** (1990), 35–46.
- [Van05] Van Schaijk R., van Duuren M., Neuilly F., Baks W., Miranda A.H., Slotboom M., Akil N. et Tello P.G., *SONOS flash memories with HfO_2 or $HfSiON$* , International Conference on Memory Technology and Design, Philips, 2005, pp. 219–221.
- [Van06a] Van Duuren M., van Schaijk R., Slotboom M., Tello P.G., Akil N., Miranda H.A. et Golubović D.S., *Pushing the scaling limits of embedded non-volatile memories with High- κ materials*, International Conference on IC Design and Technology, 2006.
- [Van06b] Van Schaijk R., van Duuren M., Akil N., Huerta A., Beckx S., Neuilly F., Rittersma Z., Slotboom M., Van Elshocht S. et Wouters J., *A Novel SONOS Memory with $HfSiON/Si_3N_4/HfSiON$ Stack for Improved Retention*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2006.
- [Van08] Van den Bosch G., Furnémont A., Zahid M.B., Degraeve R., Breuil L., Cacciato A., Rothschild A., Olsen C., Ganguly U. et Van Houdt J., *Nitride engineering for improved erase performance and retention of TANOS NAND Flash memory*, Proc. IEEE Non-Volatile Semiconductor Memory Workshop, 2008, pp. 128–129.
- [Van09] Van den Bosch G., Breuil L., Cacciato A., Jurczak M. et Van Houdt J., *Investigation of Window Instability in Program/Erase Cycling of TANOS NAND Flash memory*, International Memory Workshop, 2009, pp. 84–85.
- [Ver09] Verma S., Bersuker G., Gilmer D.C., Padovani A., Park H., Nainani A., Heh D., Huang J., Jiang J., Parat K., Kirsch P.D, Larcher L., Tseng H.-H., Saraswat C. et Jammy R., *A Novel Fluorine Incorporated Band Engineered (BE) Tunnel ($SiO_2/HfSiO/SiO_2$) TANOS with excellent Program/Erase & Endurance to 10^5 cycles*, International Memory Workshop, 2009.
-

-
- [Via06] Vianello E., Driussi F., Esseni D., Selmi L., Van Duuren M.J. et Widdershoven F., *Does multi-trap assisted tunneling explain the oxide thickness dependence of the statistics of SILC in Flash memory arrays*, Proceedings of ESSDERC, 2006, pp. 403–406.
- [Via08] Vianello E., Driussi F., Palertri P., Arregnini A., Esseni D., Selmi L., Akil N., van Duuren M. et Golubovic D., *Impact of the charge transport in conduction band on the retention of Si-Nitride based memories*, Proceedings of ESSDERC, 2008.
- [Via09a] Vianello E., Bocquet M., Driussi F., Perniola L., Molas G. et Selmi L., *Program efficiency and high temperature retention of Si₃N₄/High-κ based memories*, MicroElectronic Engineering, 2009.
- [Via09b] Vianello E., Driussi F., Arregnini A., Palertri P., Esseni D., Selmi L., Akil N., van Duuren M.J. et Golubovic D.S., *Experimental and simulation analysis of program/retention transients in silicon nitride-based NVM Cells*, IEEE Trans. on Electron Devices **56** (2009), 1980–1990.
- [Via09c] Vianello E., Perniola L., Blaise P., Molas G., Colonna J.P., Driussi F., Palestri P., Esseni D., Selmi L., Rochat N., Licitra C., Lafond D., Kies R., Reibold G., De Salvo B. et Boulanger F., *New insight on the charge trapping mechanisms of SiN-based memory by atomistic simulation and electrical modeling*, IEDM Tech. Dig., 2009.
- [Vis02] Visokay M.R., Chambers J.J., Rotondaro A.L.P., Shanware A. et Colombo L., *Application of HfSiON as a gate dielectric material*, Applied Physics Letter **80** (2002), 3183–3185.
- [Wan04] Wang X., Liu J., Bai W., Kwong D.-L., *A Novel MONOS-Type Nonvolatile Memory Using High-k Dielectrics for Improved Data Retention and Programming Speed*, IEEE Trans. on Electron Devices **51** (2004), 597–602.
- [Wan05a] Wang Y. et White M. H., *An analytical retention model for SONOS nonvolatile memory devices in the excess electron state*, Solid-State Electronics **49** (2005), 97–107.
- [Wan05b] Wang Y.Q., Singh P.K., Yoo W.J., Yeo Y.C., Samudra G., Chin A., Hwang W.S., Chen J.H., Wang S.J. et Kwong D.-L., *Long Retention and Low Voltage Operation Using IrO₂/HfAlO/HfSiO/HfAlO Gate Stack for Memory Application*, IEDM Tech. Dig., 2005, pp. 169–172.
- [Wan06] Wang X. et Kwong D.-L., *A Novel High-k SONOS Memory Using TaN/Al₂O₃/Ta₂O₅/HfO₂/Si Structure for Fast Speed and Long Retention Operation*, IEEE Trans. on Electron Devices **53** (2006), 78–82.
- [Wan07] Wang Y.Q., Hwang W.S., Zhang G., Samudra G., Yeo Y.-C. et Yoo W.J., *Electrical characteristics of memory devices with a High-κ HfO₂ trapping layer and dual SiO₂/Si₃N₄ tunneling layer*, IEEE Trans. on Electron Devices **54** (2007), 2699–2705.
- [Web05] Weber U., Boissière O., Lindner J., Schumacher M., Lehnen P., Manke C., van Elshocht S., Caymax M., Cosnier V. et McEntee T., *Improving CMOS per-*
-

- formance by AVD[®] grown High- κ dielectrics and advanced metal electrodes, Proc. Of the ECS Meeting, 2005.
- [Weg67] Wegener H.A.R., Lincoln A.J., Pao H.C., O'Connell M.R. et Oleksiak R.E., *The variable threshold transistor, a new electrically alterable, non-destructive read-only storage device*, IEDM Tech. Dig., 1967.
- [Wei77] Weinberg Z.A., *Tunneling of electron from Si thermally grown SiO₂*, Solid-State Electronics **22** (1977), 11–18.
- [Wel06] Wellekens D., Blomme P., Govoreanu B., De Vos J., Haspeslagh L., van Houdt J., Brunco D.P. et van der Zanden K., *Al₂O₃ Based Flach Interpoly Dielectrics : a Comparative Retention Study*, Proceedings of ESSDERC, 2006.
- [Wie03] Wiemer C., Fanciulli M., Crivelli B., Pavia G. et Alessandri M., *Evolution of crystallographic ordering in Hf_{1-x}Al_xO_y*, Applied Physics Letter **83** (2003), 5271–5273.
- [Wil01] Wilk G.D., Wallace R.M. et Anthony J.M., *High- κ gate dielectrics : Current status and materials properties considerations*, Journal of Applied Physics **89** (2001), 5243–5275.
- [Xio07] Xiong K., Robertson J. et Clark S.J., *Behavior of hydrogen in wide band gap oxides*, Journal of Applied Physics **102** (2007), 083710.
- [Yam05] Yamamoto, *PVD-HfSiON gate dielectrics with Ni-FUSI electrode for 65nm LSTP application*, MicroElectronic Engineering **80** (2005), 198–201.
- [Yan99] Yang N., Henson W.K., Hauser J. et Wortman J.J., *Modeling study of ultra-thin gate using direct tunneling current and capacitance-voltage measurements in MOS devices*, IEEE Trans. on Electron Devices **46** (1999), 1464–1471.
- [Yan00] Yang Y. et White M.H., *Charge retention of scaled SONOS nonvolatile memory devices at elevated temperature*, Solid-State Electronics **44** (2000), 949–958.
- [Yeo04] Yeo C.C., Joo M.S., Cho B.J. et Whang S.J., *Effect of annealing on the composition and structure of HfO₂ and nitrogen-incorporated HfO₂*, Thin solid films **462-463** (2004), 90–95.
- [Yu 02] Yu H.Y., Li M.F., Cho B.J., Yeo C.C., Joo M.S., Kwong D.-L., Pan J.S., Ang C.H., Zheng J.Z. et Ramanathan S., *Energy gap and band alignment for (HfO₂)_x(Al₂O₃)_{1-x} on (100) Si*, Applied Physics Letter **81** (2002), 376–378.
- [Yu 07] Yu W., Yu M. et Zhu C., *The Role of Nitrogen on Charge-Trapping-Induced Vth Instability in HfAlON High- κ Gate Dielectric With Metal and Poly-Si Gate Electrodes*, IEEE Trans. on Electron Devices **54** (2007), 1972–1977.
- [Zhu01] Zhu W., Ma T.P., Tamagawa T., Di Y., Kim J., Carruthers R., Gibson M. et Furukawa T., *HfO₂ and HfAlO for CMOS : Thermal stability and current transport*, IEDM Tech. Dig., 2001, pp. 463–466.
- [Zhu02] Zhu W.J., Tamagawa T., Gibson M., Furukawa T. et Ma T.P., *Effect of Al inclusion in HfO₂ on the physical and electrical properties of the dielectrics*, IEEE Electron Device Letters **23** (2002), 649–651.
-

Liste des publications

Dans cette bibliographie figure la liste de mes publications. La liste des publications est établie par catégorie de publication.

Conférences internationales

- [1] G. Molas, H. Grampeix, J. Buckley, M. Bocquet, X. Garros, F. Martin, J. P. Colonna, P. Brianceau, V. Vidal, M. Gély, B. De Salvo, S. Deleonibus, C. Bongiorno, S. Lombardo, *In-depth Investigation of HfAlO Layers as Interpoly Dielectrics of Future Flash Memories*, Proceeding of ESSDERC, 2006, pp.242-245.
 - [2] J. Buckley, M. Bocquet, G. Molas, M. Gely, P. Brianceau, N. Rochat, E. Martinez, F. Martin, H. Grampeix, JP. Colonna, A. Toffoli, V. Vidal, C. Leroux, G. Ghibaudo, G. Pananakakis, C. Bongiorno, D. Corso, S. Lombardo, B. DeSalvo, S. Deleonibus, *In-depth Investigation of Hf-based High-k Dielectrics as Storage Layer of Charge-Trap NVMs*, IEDM Tech. Dig., 2006, pp.251-254.
 - [3] M. Bocquet, G. Molas, H. Grampeix, J. Buckley, F. Martin, J. P. Colonna, M. Gély, G. Pananakakis, G. Ghibaudo, B. De Salvo, S. Deleonibus, *Intrinsic fixed charge and trapping properties of HfAlO interpoly dielectric layers*, Proc. ICMTD, 2007, pp.239-242.
 - [4] H. Grampeix, J.-P. Colonna, G. Molas, M. Bocquet, F. Martin, N. Rochat, E. Martinez, C. Licitra, T. Veyron, A.-M. Papon, M. Gely, K. Yckache, *Effect of Nitridation for High-k layers by ALCVD in order to decrease the Trapping in Non Volatile Memories*, ECS Transactions, 2007.
 - [5] G. Molas, M. Bocquet, J. Buckley, J. P. Colonna, L. Masarotto, H. Grampeix, F. Martin, V. Vidal, A. Toffoli, P. Brianceau, P. Scheiblin, M. Gély, A. M. Papon, G. Auvert, L. Perniola, C. Licitra, T. Veyron, N. Rochat, C. Bongiorno, S. Lombardo, B. De Salvo, and S. Deleonibus, *Thorough investigation of Si-nc memories with high-k interpoly for sub-45nm node Flash NAND applications*, IEDM Tech. Dig., 2007, pp.455-456.
 - [6] G. Molas, D. S. Golubović, M. Bocquet, J. Buckley, H. Grampeix, J. P. Colonna, L. Masarotto, F. Martin, M. Gély, C. Bongiorno, S. Lombardo, M. J. van Duuren, B. De Salvo, S. Deleonibus, *Integration of silicon nanocrystal memory arrays with HfAlOx based interpoly dielectric*, Proc. IEEE NVSMW, 2008.
 - [7] M. Bocquet, G. Molas, L. Perniola, X. Garros, J. Buckley, M. Gély, J. P. Colonna, H. Grampeix, F. Martin, V. Vidal, A. Toffoli, B. De Salvo, S. Deleonibus, G. Pananakakis, G. Ghibaudo, *On the Role of a HTO/Al₂O₃ Bi-Layer Blocking Oxide in Nitride-Trap Non-Volatile Memories*, Proc. ESSDERC, 2008, pp.119-122.
 - [8] G. Molas, M. Bocquet, H. Grampeix, J. P. Colonna, L. Masarotto, C. Licitra, N. Rochat, E. Martinez, F. Martin, M. Gély, C. Bongiorno, S. Lombardo, G. Pananakakis, G. Ghibaudo, B. De Salvo, *Reliability of charge trapping memories with high-k control dielectrics*, Proc. ISAGST, 2008.
-

-
- [9] E. Nowak, M. Bocquet, L. Perniola, G. Ghibaudo, G. Molas, C. Jahan, R. Kies, G. Reimbold, B. De Salvo, F. Boulanger, *New Physical Model for ultra-scaled 3D Nitride-Trapping Non-Volatile Memories*, IEDM Tech. Dig., 2008.
- [10] L. Perniola, E. Nowak, M. Bocquet, C. Jahan, R. Kies, B. De Salvo, F. Boulanger and G. Reimbold, *Superior Write/Erase performance thanks to Alumina as top dielectric in planar and SOI 3D TANOS FinFlash*, Proc. SISC, 2008.
- [11] M. Bocquet, E. Vianello, G. Molas, L. Perniola, H. Grampeix, F. Martin, J. P. Colonna, A. M. Papon, M. Gély, G. Pananakakis, G. Ghibaudo, L. Selmi, B. De Salvo, *An in-depth investigation of physical mechanisms governing SANOS memories characteristics*, Proc. IMW, 2009, pp.60-63.
- [12] G. Gay, G. Molas, M. Bocquet, E. Jalaguier, M. Gély, L. Masarotto, J. P. Colonna, H. Grampeix, F. Martin, P. Brianceau, V. Vidal, R. Kies, K.Yckache, G. Ghibaudo, T. Baron, C. Bongiorno, S. Lombardo and B. De Salvo, *Performance and reliability of Si-nanocrystal double layer memory devices with high-k control dielectrics*, Proc. IMW, 2009, pp.74-77.
- [13] E. Vianello, M. Bocquet, F. Driussi, L. Perniola, G. Molas and L. Selmi, *Program Efficiency and High Temperature Retention of SiN / High-KBased Memories*, Proc. Infos, 2009.
- [14] G. Molas, M. Bocquet, E. Vianello, L. Perniola, H. Grampeix, J.P. Colonna, L. Masarotto, F. Martin, P. Brianceau, M. Gély, C. Bongiorno, S. Lombardo, G. Pananakakis, G. Ghibaudo and B. De Salvo, *Reliability of charge trapping memories with high-k control dielectrics*, Proc. Infos, 2009.
- [15] Etienne Nowak, Elisa Vianello, Luca Perniola, Marc Bocquet, Gabriel Molas, Rabah Kies, Marc Gely, Gerard Ghibaudo, Barbara De Salvo, Gilles Reimbold, Fabien Boulanger, *Charge Localization During Program and Retention in NROM-like Non-volatile Memory Devices*, Proc. SSDM, 2009.

Articles de revues

- [16] G. Molas, M. Bocquet, J. Buckley, H. Grampeix, M. Gély, J.-P. Colonna, C. Licitra, N. Rochat, T. Veyront, X. Garros, F. Martin, P. Brianceau, V. Vidal, C. Bongiorno, S. Lombardo, B. De Salvo, S. Deleonibus, *Investigation of hafnium-aluminate alloys in view of integration as interpoly dielectrics of future Flash memories*, Solid-State Electronics **51** (2007), pp.1540-1546.
- [17] G. Molas, M. Bocquet, J. Buckley, H. Grampeix, M. Gély, J.P. Colonna, F. Martin, P. Brianceau, V. Vidal, C. Bongiorno, S. Lombardo, G. Pananakakis, G. Ghibaudo, B. De Salvo, S. Deleonibus, *Evaluation of HfAlO high-k materials for control dielectric applications in non-volatile memories*, Microelectronic Engineering **85** (2008), pp.2393-2399.
- [18] M. Bocquet, G. Molas, L. Perniola, X. Garros, J. Buckley, M. Gély, J.P. Colonna, H. Grampeix, F. Martin, V. Vidal, A. Toffoli, S. Deleonibus, G. Ghibaudo, G. Panana-
-

kakis, B. De Salvo, *Impact of a HfO₂/Al₂O₃ bi-layer blocking oxide in nitride-trap non-volatile memories*, Solid-State Electronics **53** (2009), pp.786-791.

Poster – Conférence internationale

- [19] M. Bocquet, G. Molas, E. Martinez, H. Grampeix, F. Martin, J. P. Colonna, J. Buckley, C. Licitra, N. Rochat, T. Veyron, A. M. Papon, F. Aussenac, V. Delaye, M. Gély, G. Pananakakis, B. De Salvo, G. Ghibaudo, S. Deleonibus, *NH₃ treatments of Hf-based layers for application as NVM active dielectrics*, SISC 2007.
-

TITRE : Intégration de matériaux à forte permittivité électrique (*High-κ*) dans les mémoires non-volatiles pour les générations sub-45nm

Résumé :

Les mémoires non-volatiles Flash sont aujourd'hui un élément clé du développement de l'électronique portable demandant de plus en plus de capacité de stockage à bas coût (netbook, téléphones mobiles, PDA, clé USB...). Afin d'assurer son maintien pour les années à venir, il est nécessaire de poursuivre l'amélioration de cette technologie. Ainsi, l'intégration de matériaux à forte permittivité électrique (appelés : *High-κ*) et l'utilisation de mémoires à couche de piégeage discret sont de plus en plus envisagées. Le travail de cette thèse s'inscrit dans ce contexte. Il comprend tout d'abord une étude électrique (charge fixe, piégeage, courants de fuite...) de matériaux *High-κ* (HfO_2 , $HfAlO$, Al_2O_3 , $HfSiON$) en vue de leur intégration dans les mémoires non-volatiles. Les empilements les plus prometteurs ont ensuite été intégrés dans des mémoires à nanocristaux de silicium ou à couche de piégeage nitrure. Les performances électriques ont été reliées aux propriétés matériaux des couches utilisées. L'analyse des résultats électriques ainsi que la compréhension physique des mécanismes mis en jeu a été permise par une étude de modélisation. En particulier, un modèle complet de mémoire à couche de piégeage discret a été développé.

Spécialité :

Micro- et Nano-électronique

Mots-clés :

Microélectronique, mémoire flash, mémoire non-volatile, matériaux à forte permittivité électrique, mémoires à couche de piégeage discret, SONOS, TANOS, alumine, oxyde d'hafnium, piégeage.

TITLE : Integration of *High-κ* materials for the non-volatiles memories in sub-45nm generation

Abstract :

Flash memory is today a major element for the development of the portable electronics which require more and more memory capability at low cost (netbook, cell phones, PDA, USB sticks...). In order to maintain it for the years to come, it is necessary to continue improving this technology. Also, the integration of *High-κ* materials and the use of trap charge memories are strongly envisaged. This PhD focuses on the integration and the electrical study (fixed charge, trapping, leakage currents...) of the most promising *High-κ* materials (HfO_2 , $HfAlO$, Al_2O_3 , $HfSiON$) for non-volatile memory applications. These materials are then integrated in nanocrystal memories and nitride charge trap memories. The analysis of the memory performances was made through a modelling study of the involved physical mechanisms. In particular, a complete SONOS-like model is proposed to explain the experimental results.

Speciality :

Micro- et Nano-electronics

Key words :

Microelectronic, flash memory, non-volatile memory, *High-κ* materials, charge trap memory, SONOS, TANOS, alumina, hafnium oxide, hafnium aluminates, trapping.

Thèse préparée au sein :

- Du Laboratoire d'Électronique et de Technologie de l'Information (LETI), CEA-Grenoble, Minattec, 17 av. des Martyrs, 38054 Grenoble Cedex 9, France.
- De l'Institut de Microélectronique, Electromagnétisme et Photonique (IMEP), MINATEC – INPG – 3, Parvis Louis Néel, 38016 Grenoble, France.