



Etude et validation de boucles d'asservissement permettant le contrôle avancé des procédés en microélectronique : Application à l'étape d'isolation par tranchées peu profondes en technologie CMOS.

Djaffar Belharet

► To cite this version:

Djaffar Belharet. Etude et validation de boucles d'asservissement permettant le contrôle avancé des procédés en microélectronique : Application à l'étape d'isolation par tranchées peu profondes en technologie CMOS.. Micro et nanotechnologies/Microélectronique. Ecole Nationale Supérieure des Mines de Saint-Etienne, 2009. Français. NNT: . tel-00467446

HAL Id: tel-00467446

<https://theses.hal.science/tel-00467446>

Submitted on 26 Mar 2010

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

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



N° d'ordre : **515M**

THESE
présentée par

Djaffar BELHARET

Pour obtenir le grade de Docteur
de l'Ecole Nationale Supérieure des Mines de Saint-Etienne

Spécialité : MICROELECTRONIQUE

Etude et validation de boucles d'asservissement permettant le contrôle avancé des procédés en microélectronique : Application à l'étape d'isolation par tranchées peu profondes en technologie CMOS.

Soutenue à Gardanne le 26 Février 2009

Membres du jury

Jean Pierre LANDESMAN
Bernard LEROY
Jacques PINATON
Agnès ROUSSY
Philippe COLLOT

Professeur/ Institut des matériaux Jean Rouxel, Nantes
Docteur d'Etat
Ingénieur/ ST Microelectronics
Maître-Assistante/ Ecole des Mines de Saint- Etienne
Professeur / Ecole des Mines de Saint- Etienne

A mes parents

Remerciements

Les travaux de cette thèse sont l'œuvre d'une collaboration entre le CMP-GC (EMSE) et ST Microelectronics de Rousset. Dans un premier temps, je remercie Philippe COLLOT, professeur à l'école des Mines de Saint-Etienne et directeur du CMP-GC pour m'avoir accueilli et encadré au cours de ces années de thèse. J'exprime aussi toute ma reconnaissance à Monsieur Stéphane DAUZERE-PEREZ, professeur à l'école des Mines de Saint-Etienne de m'avoir intégré dans le groupe SFL (Sciences de fabrication et logistique). J'adresse également mes remerciements à Mme Agnès Roussy, Maître assistante à l'école des mines de Saint-Etienne pour m'avoir conseillé et co-encadré au cours de ces années de thèse.

Toute ma gratitude va également à mes deux rapporteurs de thèse. Merci à M. Jean Pierre LANDESMAN, Professeur à l'institut des matériaux Jean Rouxel (Nantes) d'avoir accepté de présider mon jury de thèse et de juger mes travaux de recherches. Je veux également remercier M. Bernard LEROY pour avoir accepté les charges de rapporteur et pour ses innombrables remarques constructives.

Je n'oublie pas le personnel de ST Microelectronics, et je tiens à souligner que ces résultats ne seraient pas dans ce manuscrit sans leur précieuse assistance. Je remercie M. Jacques PINATON, ingénieur à ST Microelectronics pour ses conseils, ses encouragements, son aide et toute la confiance qu'il m'a accordée. J'adresse aussi mes remerciements à M. Pascal FORNARA pour avoir co-encadré ma thèse et pour avoir partagé son savoir-faire technologique. Merci à toutes les personnes qui m'ont aidé à réaliser mes travaux expérimentaux et en particulier à Jérôme CIRES (gravure humide), Lydie le PREVOST (PECVD) et toute l'équipe du process CMP. Je n'oublie pas l'équipe de métrologie et en particulier Geslaine et Emilienne et Sébastien. Je remercie également Damien PALMIER pour ses conseils et son aide. Merci à toute l'équipe de process control et en particulier à Franck (extraction des données par SAS) et Laurent (FDC).

Je ne peux pas oublier de remercier M.me Isabelle CONSTANT, qui m'a co-encadré en début de thèse et qui travaille maintenant dans le milieu industriel. De la même façon, je remercie M. Christian RIVERO, ingénieur à ST Microelectronics pour toute son aide et ses conseils. Je veux également remercier tous les anciens doctorants du CMP-GC (Stéphanie, Moustafa, Cathy, Said...) ainsi que tout le personnel actuel de l'équipe SFL (Bernard, Pierre, Claude, Nabil...) et du CMP-GC en général et en particulier M.me Véronique VILLAREAL.

Et enfin, je remercie toute ma famille qui m'a soutenu pendant toute cette période. Ce soutien a toujours été pour moi une source intarissable de motivation. Merci beaucoup papa et maman, malgré la distance qui nous sépare, vous êtes toujours dans mon cœur. Je remercie également mes frère et sœurs pour leurs encouragements et leur soutien permanent. Merci à Nordine, sa femme et Mahdi (l'aixois).

TABLE DES MATIERES

Table des matières	iii
Liste des figures	vi
Liste des tableaux	xii
Liste des abréviations	xii
Introduction générale	1
1 Le contrôle des procédés	6
1.1 Les techniques de contrôle.....	7
1.1.1 Les méthodes de contrôle	7
1.1.2 Les enjeux du contrôle des procédés face aux nouvelles technologies.....	8
1.1.3 Le temps de cycle.....	10
1.1.4 Les sources de variations dans la production des semi-conducteurs.....	11
1.1.5 Le contrôle des procédés et équipements	12
1.2 Etat de l'art sur le contrôle et la modélisation des procédés	16
1.2.1 Travaux sur le contrôle et la modélisation des procédés	16
1.2.2 Le contrôle avancé des procédés dans le milieu industriel.....	17
1.2.3 La conception de contrôleur R2R	20
1.3 Stratégie de l'étude et outils statistiques	21
1.3.1 Evaluation de la variabilité du procédé	22
1.3.2 Identification de l'étape et de ses paramètres critiques	27
1.3.3 Modèle du procédé	32
1.3.4 Validation du modèle	35
1.3.5 Contraintes de production	35
1.3.6 Evaluation du gain sur la réduction de la variabilité	36
1.4 Conclusion	37
2 Isolation par les tranchées peu profondes (STI) dans la technologie CMOS	41
2.1 La technologie CMOS.....	42
2.1.1 La loi de Moore	42
2.1.2 Introduction de la technologie CMOS	44
2.1.3 Les procédés de fabrication d'un circuit CMOS	45
2.2 Techniques d'isolation pour les Circuits Intégrés	45
2.2.1 Introduction sur l'isolation dans les circuits intégrés	46
2.2.2 La technologie LOCOS	48
2.2.3 La technologie LOCOS modifiée ou améliorée	52
2.2.4 Technologies NON LOCOS – Tranchées d'isolation	55
2.3 Impact du STI sur le rendement électrique des circuits.....	62
2.3.1 Impact du STI sur les paramètres électriques.....	62
2.3.2 Impact du STI sur les contraintes mécaniques.....	63
2.4 Conclusion	65

3 Corrélation entre la morphologie de la tranchée d'isolation et les propriétés électriques des transistors	69
3.1 Introduction.....	70
3.1 Corrélations entre la morphologie de la tranchée d'isolation et les propriétés électriques des transistors	72
3.2.1 Mesures de résistances.....	72
3.2.2 Principales caractéristiques des transistors	81
3.2.3 Indicateur électrique de suivi des boucles de régulations R2R du module STI.....	85
3.3 Corrélation entre la morphologie de la tranchée d'isolation et les contraintes mécaniques.....	85
3.3.1 Rappel sur l'effet des contraintes STI	85
3.3.2 Cristal soumis aux contraintes de l'isolement STI.....	86
3.3.3 Effet sur la mobilité des porteurs	87
3.3.4 Influence de la distance entre la grille et le STI	89
3.3.5 Modélisation numérique.....	95
3.4 Conclusion.....	101
4 Boucles de régulation R2R sur le module STI	104
4.1 Analyses de variabilités	105
4.1.1 Introduction.....	105
4.1.2 Rappel du procédé de la réalisation de la brique STI.....	106
4.1.3 Étude de la variabilité.....	107
4.1.4 Boucles de régulation sur le module STI.....	108
4.1.5 Organisation du travail.....	110
4.2 Procédé de dépôt chimique en phase vapeur assisté par plasma haute densité (HDP-CVD)	112
4.2.1 Introduction.....	113
4.2.2 Principe du dépôt HDP-CVD.....	113
4.2.3 Les axes principaux de recherches consacrés à l'implémentation des boucles de régulation R2R pour le procédé CVD	119
4.2.4 Variabilités du procédé HDP-CVD.....	121
4.2.5 Modélisation du Procédé HDP-CVD.....	129
4.2.6 Conclusion.....	139
4.3 Le procédé de polissage mécanico-chimique (CMP)	140
4.3.1 Introduction.....	141
4.3.2 Description du procédé CMP.....	142
4.3.3 La modélisation du polissage mécano-chimique.....	144
4.3.4 Définition de la non uniformité à travers la plaque	150
4.3.5 La distribution de la pression.....	151
4.3.6 Modélisation du procédé.....	154
4.3.7 Corrélation du profil post- CMP avec les résultats électriques.....	161
4.3.8 Corrélation entre le profil du CMP et les défauts générés.....	162
4.3.9 Conclusion.....	164

4.4 Le polissage mécanico-chimique et la gravure humide	165
4.4.1 Introduction.....	166
4.4.2 Contexte de l'étude.....	168
4.4.3 Le procédé de la gravure humide.....	169
4.4.4 Analyses multivariées.....	172
4.4.5 Modélisation du procédé.....	174
4.4.6 Premiers résultats de la pré- production.....	182
4.4.7 Surveillance de la boucle de régulation.....	183
4.4.8 Conclusion.....	185
Conclusion et perspectives	191

Liste des figures

- Figure 1.1 Pyramide représentant les différentes techniques de contrôle
Figure 1.2 Evolution du rendement en fonction du temps et du prix des puces
Figure 1.3 Architecture de la FAB automatisée
Figure 1.4 Enchaînement des opérations de fabrication des circuits
Figure 1.5 Représentation temporelle de l'évolution de la pression pendant le procédé de dépôt chimique en phase vapeur
Figure 1.6 Schéma de deux boucles de régulation Feed back et Feed forward
Figure 1.7 Le taux de déploiement de la FDC par type d'équipements
Figure 1.8 Taux de déploiement de la FDC en fonction du temps
Figure 1.9 Les différents types des outils de la FDC en fonction de temps
Figure 1.10 Stratégie de l'étude du déploiement d'une boucle de régulation
Figure 1.11 Exemple de représentation d'un box plot
Figure 1.12 a) Cartographie en 49 points b) Profil de la plaque après une opération de dépôt CVD mesuré en 49 points
Figure 1.13 Principe de la régression ACP pour un système à 3 dimensions.
Figure 1.14 Représentation des observations dans le nouvel espace
Figure 1.15 Représentation des observations lors d'une analyse PLS
Figure 1.16 Les coefficients de corrélations des différentes variables sur les deux composantes
Figure 1.17 Résultats d'une analyse à partir de Simca-P
Figure 1.18 Plan composite pour deux facteurs
Figure 1.19 Représentation de temps total d'attente.
Figure 2.1 Evolution des produits en fonction du temps suivant la loi de Moore (augmentation de la densité d'intégration et réduction des dimensions critiques)
Figure 2.2 Opérations élémentaires dans la technologie CMOS
Figure 2.3 Coupe d'un circuit après toutes les étapes de fabrication.
Figure 2.4 Représentation schématique d'un transistor bipolaire
Figure 2.5 Une structure CMOS faisant apparaître l'existence de deux transistors montés en série.
Figure 2.6 Schématisation des premiers modèles d'isolation
Figure 2.7 Schématisation de l'approche de l'oxydation sélective
Figure 2.8 Différentes étapes de fabrication pour la réalisation du LOCOS : a) Dépôt du padox et du nitre b) Après la gravure c) Oxydation d) Retrait du nitre de silicium et du padox
Figure 2.9 Schématisation de phénomène de bec d'oiseau
Figure 2.10 Représentation du modèle de Kooi
Figure 2.11 Schématisation de la méthode d'oxydation locale scellée latéralement
Figure 2.12 Représentation du procédé POP SILO
Figure 2.13 Schématisation de la méthode BOX
Figure 2.14 Schématisation des étapes élémentaires pour la réalisation du BOXES
Figure 2.15 Représentation de la fabrication du STI
Figure 2.16 Procédé du dépôt dans la tranchée : a) Oxydation thermique des bords de tranchée b) Remplissage par HDP-CVD c) Remplissage par dépôt TEOS
Figure 2.17 Procédés de polissage mécanico-chimique : a) Etat Initial b) Gravure de l'oxyde avec un motif inversé par rapport à la gravure active c) Polissage d'oxyde jusqu'au nitre

- Figure 2.18 Etapes du procédé de la gravure humide : a) STI après CMP b) Désoxydation de la surface de nitrure et ajustement de la marche c) Retrait Nitrure
- Figure 2.19 Schématisation de la hauteur de marche
- Figure 2.20 Représentation de la fabrication du PB-STI
- Figure 2.21 Comparaison de deux structures ayant différentes valeurs de la distance « a »
- Figure 2.22 Compression de la zone active par croissance d'oxyde
- Figure 2.23 Localisation des contraintes maximales en bord du STI.
- Figure 3.1 Coupe schématique d'un transistor MOS avec une vue de dessus
- Figure 3.2 Caractéristique de transfert d'un transistor MOS avec une tension V_{DS} constante
- Figure 3.3 Structure de test employée pour la mesure de la résistance carrée Nwell STI dans une structure PMOS
- Figure 3.4 Evolution de la résistance carrée Nwell STI en fonction de l'épaisseur de STI
- Figure 3.5 Evolution de la résistance carrée Nwell STI en fonction de la hauteur de marche
- Figure 3.6 Evolution de la résistance carrée Pwell STI en fonction de a) l'épaisseur STI et b) la hauteur de marche
- Figure 3.7 Coupe 1D de la structure pour les transistors PMOS et NMOS
- Figure 3.8 Profil des dopants dans le Nwell STI
- Figure 3.9 Profil des dopants dans le Pwell STI
- Figure 3.10 Simulations de la résistance carrée de la couche implantée sous le STI.
- Figure 3.11 Effet de translation du pic d'implantation avec des énergies différentes lors d'implantation de Nwell
- Figure 3.12 Calcul de l'intégral des concentrations des atomes après implantation d'un Nwell et d'un Pwell avec différentes énergies d'implantation
- Figure 3.13 Evolution de courant I_{DS} en fonction de la hauteur de marche pour les transistors NMOS($W=0,25\mu m$ $L=10\mu m$)
- Figure 3.14 Schématisation du transistor
- Figure 3.15 Structure de test conçue pour le monitoring de la tension de seuil V_t
- Figure 3.16 Evolution de la tension de seuil en fonction : a) de l'épaisseur STI b) de la hauteur de marche STI
- Figure 3.17 a) Conditions de test b) La caractéristique $I_d(V_g)$
- Figure 3.18 Direction et plan associé à une contrainte
- Figure 3.19 Contraintes d'isolement normales aux interfaces zone active/STI
- Figure 3.20 Diagramme des bandes d'énergie a) sans contraintes b) sous contraintes mécaniques
- Figure 3.21 Effet des contraintes du STI sur les parties actives
- Figure 3.22 Structure de test pour surveiller le stress
- Figure 3.23 Variation de la mobilité en fonction de la distance « a » pour les transistors NMOS et PMOS dont les dimensions ($L=0,18\mu m$ et $W=0,28\mu m$) par rapport à une référence non stressée mécaniquement.
- Figure 3.24 Dispositif expérimental : a) Machine de flexion b) mesures de flexion 4 points
- Figure 3.25 Variation de la mobilité en fonction de la contrainte appliquée
- Figure 3.26 Représentation de l'orientation de champ électrique et la contrainte appliquée
- Figure 3.27 Extraction de niveaux de contraintes en fonction de la distance « a » pour les transistors NMOS & PMOS
- Figure 3.28 Représentation d'un transistor contraint dans les 3 directions (X,Y,Z).
- Figure 3.29 Structure de base pour la simulation de la contrainte mécanique, utilisation du maillage adaptatif.
- Figure 3.30 Répartition du stress dans la zone active pour une valeur de « $a=0.2\mu m$ »

- Figure 3.31 Représentation de stress en fonction de la longueur (suivant la direction x du canal) pour différentes valeurs de « a » au milieu du canal.
- Figure 3.32 Représentation de stress en fonction de la profondeur de tranchée
- Figure 3.33 Variation du niveau des contraintes en fonction de la distance « a » pour les transistors NMOS (pour les transistors W=10µm)
- Figure 3.34 Représentation schématique de l'effet des contraintes du STI lorsque nous réduisons la longueur de grille.
- Figure 4.1 Dispersion de la hauteur de marche de lot à lot
- Figure 4.2 Dispersion de la tension de seuil.
- Figure 4.3 Récapitulatif des opérations de fabrication de la brique STI
- Figure 4.4 Un exemple d'étude de variabilité des paramètres de mesures du module STI
a)de lot à lot b) de plaque à plaque
- Figure 4.5 Boucles de régulations envisagées sur le module STI (Direct STI)
- Figure 4.6 Boucles de régulations envisagées sur le module STI (STI standard)
- Figure 4.7 Exemple de simulation sur le gain en dispersion de la hauteur de marche pour le Direct STI suite aux actions proposées.
- Figure 4.8 Remplissage des tranchées de STI (dépôt HDP-CVD + TEOS)
- Figure 4.9 Schématisation des principaux indices géométriques caractérisant le dépôt
- Figure 4.10 (a)Vue en coupe d'un dépôt PECVD classique (b) apparition d'un trou dans la tranchée
- Figure 4.11 Mécanisme du procédé (a) dépôt (b) gravure (c) dépôt HDP-CVD.
- Figure 4.12 Dépôt d'oxyde en deux étapes
- Figure 4.13 Configuration de l'équipement HDP-CVD.
- Figure 4.14 Schéma d'une chambre de dépôt
- Figure 4.15 Schématisation de la boucle de régulation R2R
- Figure 4.16 Evolution de l'épaisseur déposée lot à lot pour différents sites de mesures.
- Figure 4.17 Evolution de l'épaisseur totale pendant les 2 steps.
- Figure 4.18 Evolution de l'épaisseur d'oxyde avant et après le nettoyage de la chambre
- Figure 4.19 Exemple illustrant la variabilité intra-plaque de l'épaisseur d'oxyde déposé
- Figure 4.20 Résultats de l'analyse des composantes principales
- Figure 4.21 Résultats de l'analyse PLS
- Figure 4.22 Evolution de la pression avant et après une maintenance préventive
- Figure 4.23 Evolution de la pression avant et après le nettoyage de la chambre
- Figure 4.24 Evolution de la puissance RF Bias avant et après la tache de qualification
- Figure 4.25 Evolution de l'épaisseur d'oxyde avant et après la tache de qualification
- Figure 4.26 Cartographie de mesures en forme de couronne de 49 points
- Figure 4.27 Evolution de l'épaisseur déposée en fonction de la RF Bias et du silane side pour une valeur de silane top de 25 Sccm.
- Figure 4.28 Exemple d'une simulation à l'aide du logiciel SAS.
- Figure 4.29 Comparaison entre les données de simulation et les données expérimentales
- Figure 4.30 Evolution de l'épaisseur déposée en fonction du Silane Top et du Silane side pour une valeur de la RF Bias de 3500W
- Figure 4.31 Profil de dépôt en 3-D mesuré en 49 points
- Figure 4.32 Représentation des points mesurés et points non mesurés
- Figure 4.33 Evolution de l'étendue d'épaisseur avec le temps (mesures issues des 9points de mesures et les résultats de l'interpolation des 9 points de mesures)
- Figure 4.34 a) La répartition symétrique des injecteurs des gaz. b) profil après dépôt mesuré avec une cartographie de 49 points (présentant le problème d'injecteurs encrassés)
- Figure 4.35 Profils d'oxyde après CMP

- Figure 4.36 Représentation schématique de la machine de polissage
- Figure 4.37 Représentation de l'interface plaquette- tissu de polissage.
- Figure 4.38 a) Structure d'un film avant CMP b) Après CMP : oxyde résiduel reste dans certains endroits.
- Figure 4.39 Modélisation du procédé CMP
- Figure 4.40 Représentation schématique des marches initiales et finales
- Figure 4.41 Profil de la profondeur de tranchée après gravure
- Figure 4.42 Hauteur de marche finale après polissage
- Figure 4.43 Schématisation de l'interface plaque/pad
- Figure 4.44 Distribution de la pression à travers la plaque
- Figure 4.45 Coupe d'une machine de polissage CMP
- Figure 4.46 a) cartographie de mesure de 460 points b) cartographie de mesure réduite à 109 points
- Figure 4.47 Exemple de cartographie de mesure de type radial des mesures de la hauteur de marche
- Figure 4.48 Variation de la hauteur de marche au niveau du produit suivant les valeurs de pression de l'anneau de maintien
- Figure 4.49 Evolution de l'épaisseur d'oxyde pré- CMP au sein de la plaque
- Figure 4.50 Evolution de l'épaisseur d'oxyde post- CMP au sein de la plaque (sur les 25 derniers mm)
- Figure 4.51 Corrélation entre le profil d'oxyde post CMP et les résultats de test électrique sur le paramètre de tension de seuil
- Figure 4.52 Comparaison de profil d'oxyde après CMP avec le profil de la densité des défauts
- Figure 4.53 Condensation des particules de l'abrasif dans certaines zones du tissus de polissage en contact avec la plaque
- Figure 4.54 Représentation simplifiée des étapes des procédés STI
- Figure 4.55 Schéma synoptique de la boucle CMP- gravure humide
- Figure 4.56 Variation de la hauteur de marche en fonction de l'épaisseur de nitrure post-CMP
- Figure 4.57 Variation de la hauteur de marche en fonction de l'épaisseur de nitrure post-CMP (X_1) dans le cas du direct STI
- Figure 4.58 Comparaison de deux structures « box plot » représentant la distribution de l'épaisseur de nitrure post-CMP à travers la plaque dans le cas du STI standard et du Direct STI
- Figure 4.59 Gravure de deux lots par voie humide
- Figure 4.60 Les différents bains constituant la machine de gravure humide
- Figure 4.61 Résultats de l'analyse PLS
- Figure 4.62 Comparaison des profils : a)après CMP b) après gravure humide
- Figure 4.63 Comparaison entre les valeurs de mesure et de simulation
- Figure 4.64 Illustration de l'épaisseur polie après le procédé CMP
- Figure 4.65 Représentation schématique du modèle.
- Figure 4.66 Comparaison entre les données mesurées et la simulation
- Figure 4.67 Evolution de la tension de seuil en fonction de l'épaisseur STI
- Figure 4.68 Schéma synoptique du superviseur

Liste des tableaux

- Tableau 2.1 Evolution des technologies en fonction du temps
Tableau 3.1 Valeurs des épaisseurs STI pour les six plaques
Tableau 3.2 Densité de dopage du substrat en fonction de l'épaisseur ST
Tableau 3.3 Données de la distance « a »
Tableau 4.1 Temps du procédé par plaques
Tableau 4.2 Récapitulatif des conditions expérimentales du plan d'expériences
Tableau 4.3 Valeurs de P-value obtenues sous SAS.
Tableau 4.4 Valeurs de P-value obtenues sous SAS
Tableau 4.5 Conditions expérimentales
Tableau 4.6 Ecart type normalisé de la hauteur de marche des plaques
Tableau 4.7 Conditions expérimentales des plaques sélectionnées pour le plan d'expériences
Tableau 4.8 Ecart type normalisée de l'étendue d'épaisseur d'oxyde avant CMP.
Tableau 4.9 Exemple des conditions de procédé.
Tableau 4.10 Mesures avant et après gravure H₃PO₄
Tableau 4.11 Conditions expérimentales visées pour le plan d'expériences.
Tableau 4.12 Résultats expérimentaux
Tableau 4.13 Valeurs de (P-value) obtenus sous SAS
Tableau 4.14 Les résultats des mesures sur les lots de la pré- production

Liste des abréviations

APC	Advanced Process Control
BPSG	Boron Phosphorus Silicate Glass
BOX	Buried Oxide
CD	Critical Dimension
CMOS	Complementary Metal Oxide Silicon
CMP	Chemical Mechanical Polishing
CVD	Chemical Vapor Deposition
DCS	Distributed Control System
DOE	Design of experiments
EWS	Electrical Wafer Test
FDC	Fault Detection and Classification
HDP-CVD	High Density Plasma CVD
IMD	Inter Metal Dielectric
LCL	Lower Control Limit
LOCOS	Local Oxidation of Silicon
LPCVD	Low Pressure CVD
LSL	Lower Specification Limit
PBL	LOCOS Poly Buffered LOCOS
PECVD	Plasma Enhanced CVD
PMD	Pre Metal Dielectric
PSG	Phosphorus Silicate Glass
PT	Parametric Test
R2R	Run to Run
RTO	Real Time Optimisation
SILO	Sealed Interface Local Oxidation
SPC	Statistical Process Control
STI	Shallow Trench Isolation
TEOS	Tetra Ethyl Orto Silicate
UCL	Upper Control Limit
USL	Upper Specification Limit
WIWNU	Within Wafer Non Uniformity
ZA	Zone Active

Introduction générale

Le marché des semi-conducteurs représente un total de plus de 250 milliards de dollars et se maintient avec une croissance annuelle allant de 8 à 10%. Ces dernières années, les applications des semi-conducteurs se sont beaucoup diversifiées et touchent désormais plusieurs secteurs: la téléphonie mobile, les télécommunications, internet... Le marché des semi-conducteurs est constitué à plus de 95% de la technologie CMOS (Complementary Metal Oxide Semiconductor).

Depuis plus de 40 ans, les industriels augmentent ainsi la densité de leur circuits. Pour cela, ils réduisent la taille des éléments qui les constituent, autrement dit les transistors. L'association de l'industrie des semi-conducteurs : SIA (Semiconductor International Association) publie chaque année une feuille de route : ITRS (International Technology Road map for Semiconductor). Cette feuille de route définit les spécifications en terme de dimensions, de composants, de matériaux et aussi d'équipements à intégrer pour la fabrication. Pour respecter les spécifications exigées, il faut alors contrôler avec rigueur les procédés de fabrication.

Généralement, les procédés sont incapables de reproduire toujours à l'identique un produit. Lorsqu'on effectue le contrôle d'une des caractéristiques d'un produit, on observe une dispersion des valeurs mesurées autour de la valeur cible (visée), on parle alors de variabilité. Pour mieux contrôler les procédés, il est nécessaire de réduire ou d'éliminer si possible cette variabilité mais avant tout, il faut comprendre l'origine de la variabilité et le type de variabilité.

L'industrie du semi-conducteur utilise depuis plusieurs années les techniques de contrôle des procédés tel que le SPC (Statistical Process Control). Le contrôle SPC est basé sur l'utilisation de cartes de contrôle. On considère un procédé sous contrôle statistique

lorsque toutes les valeurs des échantillons mesurés se situent à l'intérieur des limites de tolérance du procédé définies par l'intermédiaire de cartes de contrôle. Tant que la valeur est dans ces limites, il n'y a pas de raison d'agir sur le procédé. On risquerait de décentrer un procédé bien centré. Si une valeur est en dehors de ces limites, il y a une forte probabilité que le procédé ne soit plus maîtrisé et qu'une variation anormale en soit la cause. Il faut alors le recentrer.

Le contrôle par la technique SPC classique n'est plus adéquat avec les nouvelles technologies car ces dernières imposent des limites de contrôle très serrées. Un contrôle plus rigoureux est donc nécessaire d'où le recours à la technique de contrôle avancé des procédés ou APC (Advanced Process Control). Le contrôle avancé des procédés associe trois techniques dont le SPC classique et le contrôle des équipements connu sous le nom de FDC (Fault Detection and Classification). La troisième composante est une technique de régulation de type Run à Run (Run to Run : R2R) permettant l'ajustement de certains paramètres de procédés en temps réel afin de s'approcher de la valeur cible.

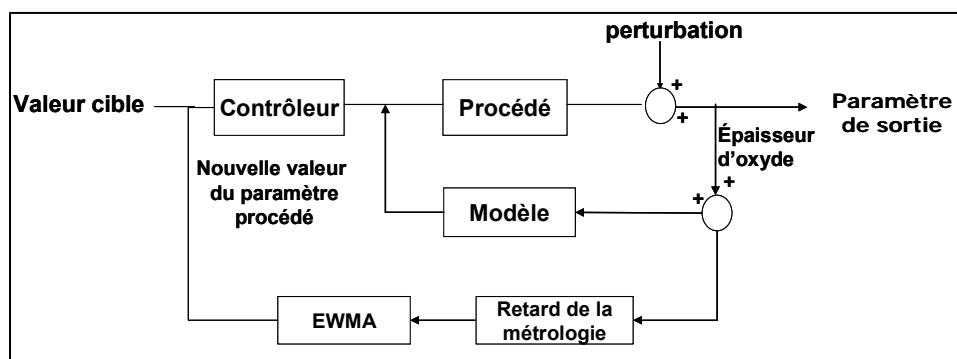


Figure 1. Exemple d'un schéma synoptique d'une boucle de régulation de type R2R

La réduction des dimensions entraîne également des changements de procédés au niveau des briques principales des circuits (isolation, grille, métal et interconnexions). La brique qui sera abordée tout au long de ce manuscrit est la brique concernant l'isolation entre les transistors. Pour les technologies $<0,25\mu\text{m}$, le procédé d'isolation employé est basé sur des tranchées d'isolations peu profondes ou bien le STI (Shallow Trench Isolation). En effet, les procédés d'isolation étaient déjà connus bien avant l'introduction de la technique d'isolation par tranchées, pour les problèmes qu'ils engendraient, comme par exemple, la création de dislocations. Le STI, basé sur des tranchées d'oxydes peu profondes, offre une bonne isolation entre les transistors.

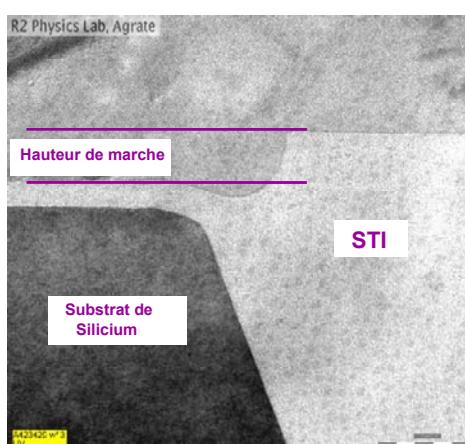


Figure 2. Représentation schématique du STI

Le STI est caractérisé généralement par une grandeur géométrique appelée hauteur de marche comme cela est illustré sur la figure 2. Cette hauteur de marche est le résultat de différentes étapes technologiques que nous décrirons ultérieurement. Si cette hauteur dépasse quelques nm, les effets résultants seront néfastes sur le rendement final et ne peuvent être constatés que lors des tests électriques et du test final pour la sélection des bonnes puces.

Des variabilités (dispersion de la mesure par rapport à la valeur cible) de la hauteur de marche ont été observées. Les variabilités peuvent se manifester de lot¹ à lot, de plaque à plaque et même au sein d'une plaque (variabilité intra-plaque). La variabilité de la hauteur de marche est la conséquence des différentes variabilités observées au niveau des paramètres de sortie (paramètres de mesures physiques) des procédés associés à la brique STI. Les analyses de variabilité permettent de conclure quels procédés/paramètres contribuent plus à la dispersion de la hauteur de marche.

Nous allons, tout au long de ce manuscrit, nous pencher sur les différents procédés identifiés comme paramètres contributeurs à la dispersion de la hauteur de marche. Le défi est d'arriver à mieux contrôler cette dispersion.

Les solutions aux problèmes de dispersion des mesures peuvent être abordées selon deux volets :

- le changement des procédés impliquant un certain nombre de modifications sur les recettes et les étapes des procédés. Ceci est une bonne voie mais elle ne s'inscrit pas dans le cadre de notre travail, donc nous ne focaliserons notre attention que sur les procédés considérés comme matures.
- le contrôle des procédés qui consiste à apporter des solutions pour stabiliser le procédé et éviter des dérives (atteinte ou dépassement des limites de contrôle). C'est sur ce deuxième volet que nous allons travailler.

L'objectif de ces travaux de thèse est d'obtenir une meilleure compréhension des variabilités générées au niveau du module STI, de les quantifier et de mesurer leur impact sur le rendement. On pourra justifier ainsi les actions de contrôle avancé des procédés menées en vue de réduire la dispersion de la hauteur de marche et proposer des solutions telles que les boucles de régulation de type R2R. Les produits technologiques qui sont concernés par cette étude sont des produits mémoires (longueur de grille des transistors est de 0,18µm). Le choix de ce type de produits est justifié par le fait que ces produits sont plus sensibles au STI que les produits logiques. Cette forte sensibilité est due aux procédés thermiques additionnels comme le procédé de recuit au niveau des opérations post-STI.

Dans le chapitre 1, nous rappellerons les méthodes de contrôle utilisées dans l'industrie en général et les enjeux économiques importants favorisant l'introduction de telles actions. Ensuite, nous dresserons un état de l'art détaillé sur les différents travaux industriels et académiques sur l'APC dans le milieu des semi-conducteurs. Suite à cet état de l'art, nous exposerons la stratégie de l'étude proposée et nous décrirons tous les outils statistiques utilisés au cours de la thèse. Ce premier chapitre va surtout servir de référence et nous y ferons régulièrement appel tout au long des autres chapitres.

Le chapitre 2 sera consacré au STI. Au début du chapitre, nous rappellerons quelques bases élémentaires sur la technologie CMOS, la loi de Moore et une brève description des procédés de fabrication nécessaires pour la réalisation des circuits. Nous ne nous focaliserons

¹ Un ensemble de plaques, généralement un lot est composé de 25 plaques.

par la suite que sur la brique technologique d'isolation. Avant de décrire le STI, nous rappellerons les techniques employées avant l'introduction du STI. Nous verrons bien que les changements de procédés effectués jusqu'au STI répondent souvent à un souci généralement lié aux performances électriques. Après ce rappel, nous décrirons tous les procédés de fabrication permettant la réalisation de la brique STI et enfin nous discuterons de l'impact du STI sur le rendement final. Nous mettrons l'accent en particulier sur l'influence du STI sur certains paramètres électriques (résistance carrée, tension de seuil...) et également sur les contraintes mécaniques. Cette partie sera importante pour situer notre travail dans cet axe de recherche et éventuellement, y apporter notre contribution. Nous évoquerons également dans ce chapitre, un nouveau mode de STI ou STI amélioré : le direct STI, nous verrons que c'est un mode qui présente un enjeu économique en permettant de se passer de deux étapes de fabrications. En plus de ça, il permet d'avoir une bonne uniformité de la hauteur de marche.

Une grande partie expérimentale sera décrite dans le chapitre 3 pour illustrer les différentes expériences menées en vue de démontrer et de quantifier l'impact du STI sur les paramètres électriques et les contraintes mécaniques, suivra une partie statistique sur l'analyse des données de mesures électriques. Nous allons appuyer ces travaux sur des simulations et des caractérisations pour mieux expliquer physiquement les résultats obtenus. Nous terminerons ce chapitre par des propositions et des conclusions, parmi lesquelles l'obtention d'un indicateur électrique pour le suivi des boucles de régulations R2R.

Après ce travail d'analyses et de corrélations, nous aborderons le chapitre 4. Nous exposerons les analyses de variabilité effectuées sur les procédés de fabrication au niveau de la brique STI. Cette analyse nous permettra d'identifier les paramètres critiques qui contribuent le plus à la variabilité de la hauteur de marche. Ce chapitre est composé de trois grandes parties correspondant aux trois boucles de régulation proposées pour la réduction de la dispersion de la hauteur de marche.

La première partie traite de la réalisation de la boucle de régulation de type « feed back » de la boucle de dépôt HDP-CVD (High Density Plasma Chemical Vapor Deposition). Nous rappellerons les principes physiques du HDP-CVD. Nous détaillerons ensuite les analyses multi-variées effectuées qui serviront à identifier les étapes/paramètres critiques. Ces résultats seront interprétés par la suite pour définir les plans d'expériences. Deux modèles de procédés sont recherchés: le modèle d'épaisseur déposée et le modèle de l'étendue de l'épaisseur d'oxyde. Ces modèles seront simulés pour tester leur robustesse de prédiction. Nous proposerons une méthode originale basée sur des algorithmes d'apprentissage statistique pour éliminer l'effet des injecteurs de gaz.

La deuxième partie de ce chapitre sera consacrée à la boucle de régulation de type « feed back » sur le procédé polissage mécano- chimique CMP (Chemical Mechanical polishing). Cette boucle ne concernera que le STI standard marqué par une non uniformité de la hauteur de marche, résultat de la non uniformité des paramètres de mesures post- CMP. Nous rappellerons brièvement les définitions de base du procédé CMP ainsi que les différents travaux de recherches effectués pour modéliser le CMP. Nous situerons notre travail dans l'axe « modélisation de l'uniformité sur la plaque ». Nous allons voir que la pression est un paramètre très critique sur l'uniformité de l'oxyde post-CMP sur la plaque et en particulier sur les bords des plaques. L'approche proposée consiste à modifier la pression au niveau des bords de plaque. Nous détaillerons notre approche et les différents plans d'expériences effectués pour retrouver le modèle du procédé. Nous terminerons enfin cette partie par l'interprétation de tous les résultats.

La dernière partie de ce chapitre traite la boucle de régulation de type « feed-forward » entre les procédés de CMP et de gravure humide (WET). Au début de ce chapitre,

nous insisterons sur l'originalité de ce travail. L'équipement de gravure humide sera décrit en détails. Ensuite, nous exposerons les résultats de l'analyse multivariée pour identifier les paramètres critiques qui nous permettront par la suite de concevoir et de réaliser le modèle du procédé. Le modèle établi sera proposé à la production afin de le tester. Nous exposerons tous les résultats expérimentaux et de simulations.

Nous terminerons ce chapitre par une proposition concernant la surveillance des boucles de régulation R2R. Cette surveillance est basée sur un superviseur. Dans le but d'éliminer l'effet produit et les variations induites lors des changements de procédés, une calibration périodique de certains éléments des boucles s'impose. Les valeurs cibles de certains paramètres des procédés doivent donc être actualisées pour assurer à la fin de bonnes performances électriques.

Enfin, nous pourrons conclure sur les différents travaux réalisés, et nous proposerons les perspectives à venir.

Chapitre I

Le contrôle des procédés

Le contrôle des procédés dans le domaine industriel constitue un enjeu économique très important. Le développement des techniques de contrôle est favorisé par la croissance des vitesses des microprocesseurs. Ce développement a accompagné la création de nouveaux outils et logiciels plus performants. Ces derniers ont contribué à la naissance d'une technique de contrôle plus robuste et plus rapide appelée « contrôle avancé des procédés » ou APC : Advanced Process Control [BONAVITA 1998]. Les industries chimiques et pétrochimiques sont de loin les industries leaders en terme d'usage des techniques de contrôle avancé des procédés. L'industrie pharmaceutique utilise aussi ces techniques avancées de contrôle. Contrairement à l'industrie des semi-conducteurs, les lignes de production des autres industries sont caractérisées par le fait que les opérations des procédés suivent un flux continu et leur nombre est limité. Dans l'industrie des semi-conducteurs, le nombre d'opérations pour réaliser des circuits est de l'ordre de 350 à 400 pour les technologies actuelles. Le flux des procédés est discontinu et le délai de réalisation des dispositifs est en moyenne de 2 à 3 mois. Ces différences majeures ont fait que l'industrie des semi-conducteurs a enregistré un retard pour l'application des techniques de contrôle avancé des procédés. Le défi principal est de surmonter les contraintes liées à cette industrie.

Dans les paragraphes suivants, l'évolution des techniques de contrôle dans le domaine de l'industrie pétrolière sera décrite et nous dresserons un état de l'art des techniques de contrôle des procédés dans le domaine des semi-conducteurs. Nous donnerons également des exemples d'industrialisation des systèmes de contrôle avancé des procédés.

1.1 Les techniques de contrôle

1.1.1 Les méthodes de contrôle

Beaucoup de techniques de contrôle ont été proposées pour contrôler les procédés. Le succès de certaines méthodes de contrôle dans une industrie suscite beaucoup d'intérêts pour le généraliser à d'autres industries. C'est le cas du Système de Contrôle Distribué (DCS) qui a eu un grand succès dans l'industrie pétrolière dans les années 70 et qui a ouvert la voie à l'utilisation des techniques d'asservissement telle que : la cascade, le « Feed Forward² », le « Feed Back³ », le contrôle non linéaire, etc ... Nous reviendrons longuement dans ce manuscrit sur l'application des techniques de « Feed Forward » et de « Feed Back ».

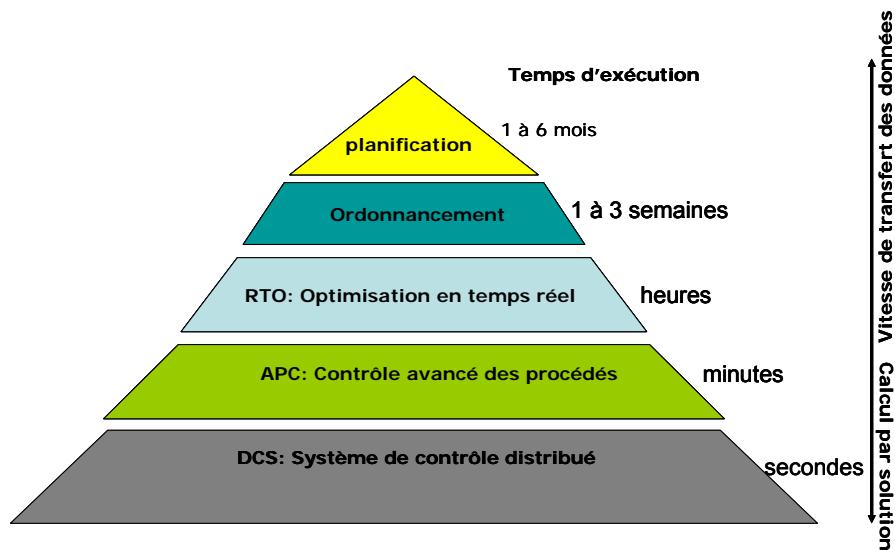


Figure 1.1- Pyramide représentant les différentes techniques de contrôle [APC network 2007].

Dans les années 80, un nouveau type de contrôleurs a vu le jour. Il s'agit de contrôleur prédictif multivarié⁴ qui non seulement assure des techniques telles que le « Feed Forward » et le « Feed Back » mais qui a permis de franchir une autre étape plus avancée. Si des interactions entre les paramètres du procédé existent, ce contrôleur est beaucoup plus efficace. Par contre, les capacités de calcul de ce contrôleur sont plus importantes que ceux du DCS [APC network 2007].

Depuis la fabrication des ordinateurs puissants, capables de surveiller une usine en collectant des données en temps réel, cette invention a ouvert la voie à l'apparition des technologies complémentaires comme les réseaux de neurones. Un autre exemple de contrôle avancé est la technologie des systèmes. Cette technologie a vu le jour à la fin des années 80. Le principe de cette technologie repose sur l'exploitation des idées des meilleurs collaborateurs. Ensuite, les données sont complétées en temps réel. Cette technologie a connu un succès dans la l'industrie pétrolière.

² Contrôle par anticipation

³ Contrôle par rétroaction

⁴ Modèle capable de prédire les paramètres de sortie d'un système ayant en entrée plusieurs paramètres

Dans les années 1990, les entreprises ont manifesté un besoin d'avoir des interfaces robustes entre les ingénieurs et les opérateurs. Les fournisseurs de la technologie DCS ont inclus dans leurs matériels des contrôleurs de prédiction par modèles. Cette technique est connue sous le nom de technique d'optimisation en temps réel (RTO : Real Time Optimisation). Au départ, cette technologie n'a pas réussi à s'imposer dans les applications industrielles à cause de temps de calcul des différentes équations jugés très long. Il a fallu donc attendre l'arrivée sur le marché d'ordinateurs plus puissants pour réduire les temps de calculs. La technologie RTO a réussi à résoudre des cas compliqués dans l'industrie pétrolière. Il est à souligner que la RTO ne remplace pas la programmation linéaire qui est adoptée depuis plus de 50 ans [APC network 2007].

1.1.2 **Les enjeux du contrôle des procédés face aux nouvelles technologies**

1.1.2.1 **La vie d'une technologie**

Un circuit intégré se compose de plusieurs couches minces. Chaque couche a ses propres propriétés chimiques pour réaliser au final les caractéristiques électriques désirées. Les dispositifs subissent lors de la fabrication différentes opérations physiques et chimiques. Le nombre des opérations diffère d'une technologie à une autre. Pour les technologies actuelles produites sur le site ST-Rousset, ce nombre avoisine les 350 étapes. Pour élaborer l'ensemble de ces couches, on utilise une plaque de silicium de type monocristallin. Le diamètre des plaques est de 100 à 300mm. Sur le site de ST-Rousset, le diamètre des plaques utilisées est de 200mm. Avec les nouvelles technologies (réduction des dimensions des circuits), des problèmes liés aux procédés de fabrication apparaissent et impactent négativement le rendement. Des efforts sont à déployer sur les 3 principaux domaines incontournables de la réalisation des dispositifs en vue d'assurer un bon rendement :

- au niveau de la conception (design) en proposant de nouvelles architectures adéquates aux nouvelles technologies.
- au niveau de la fabrication en introduisant de nouveaux équipements et de nouveaux procédés pour assurer la reproductibilité des fonctions des produits.
- au niveau de l'industrialisation et de la mise sous contrôle en proposant de nouvelles méthodes de surveillance des procédés et des équipements afin de mieux quantifier et contrôler les variations des mesures sur les produits observées dans le temps.

Dans ce manuscrit, nous nous intéressons uniquement à ce dernier domaine et plus particulièrement aux méthodes de contrôle des procédés de fabrication qui sont devenues un secteur primordial dans le fonctionnement d'une usine de fabrication des semi-conducteurs (FAB⁵). La figure 1.2 représente grossièrement la vie d'une technologie. Nous représentons le rendement final de fabrication en fonction du temps avec et sans utilisation des techniques de contrôle avancé des procédés. La première période, appelée « période pilote » est marquée par

⁵ Dans l'industrie de fabrication de semi-conducteur, nous désignons couramment sous le nom de FAB (de l'anglais *fabrication plant*). Le cœur d'une *fab* réside dans sa salle blanche, une zone où l'environnement est contrôlé pour éviter toute micropoussière et où la moindre micro-vibration est proscrite. La salle blanche est remplie d'équipements très coûteux nécessaires aux nombreuses étapes de fabrication

le développement de nouveaux modules procédés. Les modules, issus des générations antérieures qui ne répondent plus aux besoins sont identifiés. Dans la seconde période (cycle rapide), on intègre les nouveaux modules dans une séquence complète de fabrication. Après plusieurs itérations, la séquence de fabrication est affinée et optimisée pour limiter la sensibilité de fonctionnement aux paramètres des procédés de fabrication. A la fin de cette étape, la capacité de production et le rendement sont augmentés. La dernière période correspond à la production de masse des circuits.

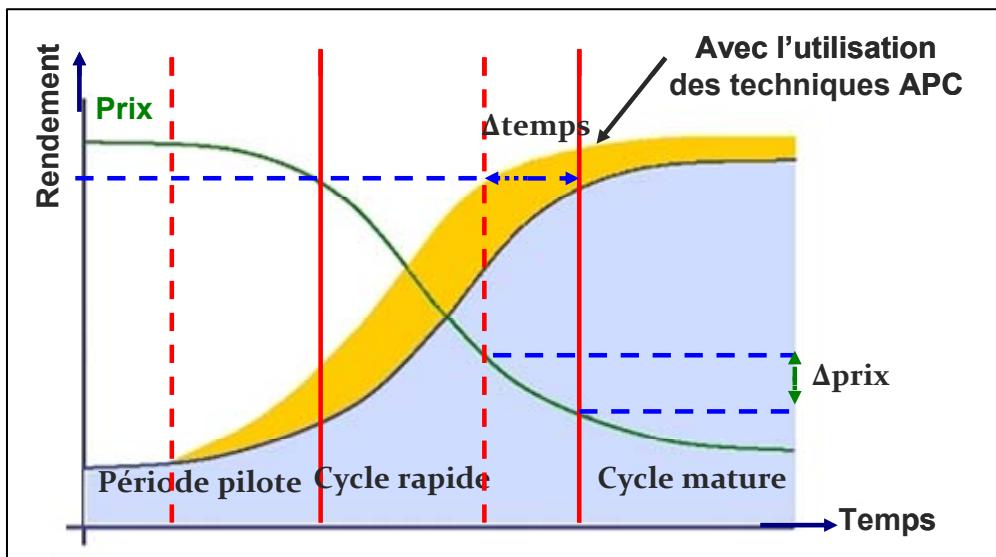


Figure 1.2- Evolution du rendement en fonction du temps [KRIGER 2005]

Au début le prix du produit et la demande sont hauts. Ces deux quantités diminuent plus ou moins rapidement avec l'arrivée des concurrents sur le marché. Après un certain temps (quelques mois), la technologie des produits les plus avancés est dépassée par de nouvelles générations et les prix continuent à chuter comme cela est illustré sur la figure 1.2. Le déploiement des techniques de contrôle avancé des procédés permet de limiter le temps de la période de cycle rapide et d'améliorer le rendement pendant cette période et pendant la période mature. Pour déployer ce type de techniques, il est impératif que la FAB dispose d'outils modernes, on parle souvent d'une FAB moderne.

1.1.2.2 Une usine de fabrication des semi-conducteurs automatisée

Typiquement une usine de fabrication de semi-conducteurs (FAB) moderne a les caractéristiques suivantes disponibles pour l'analyse et le contrôle :

- les données en temps réel fournies par les différents capteurs qui reflètent l'état de santé de l'équipement.
- la métrologie intégrée et les données de métrologie en ligne disponibles après plusieurs étapes de procédés de fabrication afin de remédier aux problèmes de retards de la métrologie. Le retard de la métrologie peut être défini comme étant le temps d'attente pour avoir l'information sur les résultats de métrologie suite au procédé de fabrication.
- une base de données des paramètres électriques au niveau du test final des puces.

Ces trois conditions constituent les blocs importants de l'architecture du système de la FAB automatisée. Une première version de l'architecture de la FAB automatisée a été présentée par T.Sonderman & J.Qin en 2002 [Qin 2006] comme cela est illustré sur la figure 1.3.

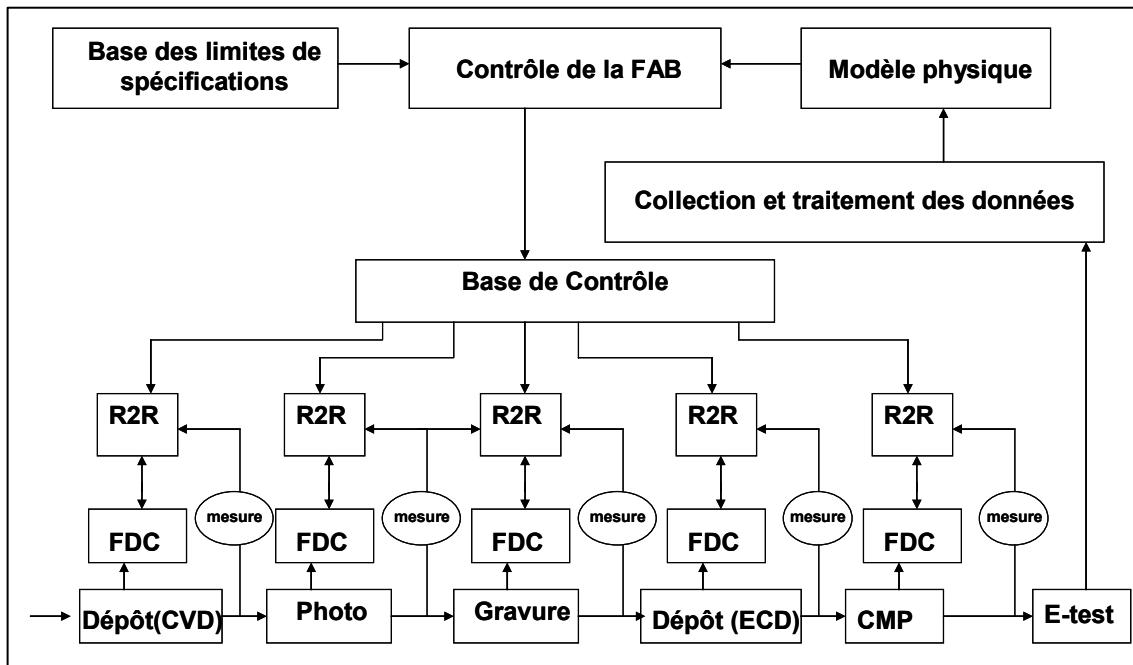


Figure 1.3- Architecture de la FAB automatisée [SONDERMAN 2005] [Qin 2006]

Cette représentation de la FAB automatisée est le résultat de l'analogie FAB semi-conducteurs – usine de chimie. Nous distinguons sur la figure 1.3 plusieurs niveaux de contrôle. Le premier niveau de contrôle concerne la maîtrise de l'équipement. Il s'agit généralement de la mise en place des systèmes de contrôle pour s'assurer du bon fonctionnement de l'équipement. Le second niveau est un système de contrôle des procédés. Ce système utilise les données de mesure et les paramètres des procédés. Nous reviendrons en détail sur les deux premiers systèmes de contrôle dans les prochains paragraphes. Le troisième niveau est une base de données qui regroupe toutes les données fournies par les deux premiers systèmes de contrôle. Le dernier niveau collecte et traite les données de tests électriques et ensuite les compare avec les données théoriques (calculées à partir de modèles physiques déjà établis). Suivant les spécifications souhaitées, des recommandations peuvent être formulées pour les systèmes de contrôle. L'enjeu économique de la modernisation d'une FAB est de réduire le temps de cycle, ce dernier impacte directement le rendement de la production.

1.1.3 Le temps de cycle

Le temps de cycle⁶ de fabrication constitue une partie des coûts de fabrication d'une ligne de production. Le défi est de réduire au maximum ce temps de cycle, car, plus ce temps est long, plus les coûts de fabrication sont élevés.

⁶ Le temps de cycle de production, appelé aussi temps de séjour ou délai de fabrication, est le temps total nécessaire pour fabriquer un lot de plaques. Il mesure le temps écoulé entre le moment où le lot entre dans la fab

En général, le temps de cycle dépend [ROBINSON 2005, ROBINSON 2007] :

- du temps d'utilisation de l'équipement.
- du nombre des équipements qualifiés à chaque opération.
- de la variabilité⁷ ou dispersion qui provient de l'absence de la répétitivité parfaite des mesures.

Sur le premier point, on peut améliorer le temps de cycle en augmentant le temps de disponibilité de l'équipement. Sur le deuxième point, il est possible d'augmenter le nombre des équipements qualifiés⁸ et s'assurer que chaque opération de fabrication est réalisable sur au moins deux équipements qualifiés. Sur ce dernier point, il faut chercher à identifier la source de variabilité et trouver une solution pour l'éliminer.

1.1.4 Les sources de variations dans la FAB

Il y a beaucoup de facteurs qui contribuent à la variabilité des paramètres de sortie (mesures).

1.1.4.1 Variabilité due à l'équipement :

- le temps d'attente au niveau de l'équipement (fabrication ou mesure)
- le temps de transport des lots (d'un atelier à un autre)
- le problème lié à la gestion des lots (le choix des équipements pour les prochaines opérations).
- le temps de maintien des lots. Pour des raisons d'inspection ou d'analyse des données, certains lots restent inactifs pendant un certain temps pour valider les résultats des inspections.

1.1.4.2 Variabilité due au temps du procédé :

- le rejet des plaques. Quand les résultats des mesures après un procédé sont hors des plages des limites de tolérances (limites de spécifications), les plaques ne vont pas continuer les opérations suivantes et par conséquent elles seront recyclées.
- le re-traitement des plaques. Si, dans certains cas, les mesures des plaques après une opération de fabrication ne sont pas satisfaisantes, si le degré de criticité du procédé n'est pas très important, il est possible de re-traiter (re-processer) les plaques. Le phénomène est appelé généralement « Re-work » ou « Re-travail ».
- le mixage des recettes⁹ des différents produits procédés sur le même équipement

et le moment où il en sort. Le temps de cycle inclut le temps effectif de fabrication (temps de process), le temps de transfert entre les opérations, le temps d'attente devant les machines, et le temps d'attente pour le transfert.

⁷ Généralement, les procédés sont incapables de reproduire toujours à l'identique un produit. Lorsqu'on effectue le contrôle d'une des caractéristiques d'un produit, on observe une dispersion des valeurs mesurées autour de la valeur cible (visée), on parle alors de variabilité.

⁸ Tous les équipements n'ont pas la capacité de réaliser une opération donnée d'un lot. Pour qu'un équipement soit apte à réaliser une opération ou une recette on doit le qualifier pour cette recette. Un équipement qui n'est pas qualifié à un instant donné peut l'être à l'instant suivant moyennant les réglages nécessaires.

⁹ Fichier présent sur chaque équipement comportant les consignes des différents paramètres du procédé (pression, puissance RF, flux des gaz...)

- le mixage des produits procédés sur le même équipement
- l'arrêt d'une machine pour une maintenance
- l'occupation de la machine pour des expériences spécifiques¹⁰.
- le temps de prise de décision lors d'un problème.
- le temps entre deux étapes consécutives

Dans ce manuscrit, on va s'intéresser uniquement aux variations liées au procédé. L'ensemble des actions qui seront décrites dans les prochains chapitres vise à réduire la variabilité liée au procédé. L'objectif de ces actions est de réduire le nombre de rejet de plaques et donc, d'améliorer le rendement et le temps de cycle. Le développement des nouvelles techniques de contrôle pour la réduction de la variabilité surtout sur les nouvelles technologies est un moyen pour atteindre cet objectif.

1.1.5 **Le contrôle des procédés et des équipements**

Le concept de contrôle des équipements et des procédés est connu sous le nom de contrôle avancé des équipements/contrôle avancé des procédés (AEC/APC : advanced Equipment Control/Advanced Process Control). Pour une meilleure maîtrise de la FAB, le contrôle des équipements et des procédés est nécessaire. Le contrôle avancé des procédés est composé de trois composantes principales que nous allons décrire en détail dans les paragraphes suivants.

1.1.5.1 **Le contrôle statistique du procédé (SPC: Statistical Process Control)**

Depuis 1920, le SPC est le moyen de contrôle des procédés de fabrication adopté par la plupart des industries. Il est basé sur des indicateurs statistiques (moyenne, étendue, écart type....) [MONTEGOMERY 2001]. Nous détaillerons ces indices statistiques dans la troisième partie de ce chapitre. Les mesures prélevées après l'étape de fabrication sont comparées avec certaines limites fixées par les technologues. Ce type de contrôle est assisté par des cartes de contrôle.

1.1.5.1.a **Les principes de base des cartes de contrôle**

Les cartes de contrôle sont les outils indispensables pour réaliser un pilotage rationnel du procédé de fabrication. Une application rigoureuse de cette méthode permet d'améliorer de manière significative la capabilité¹¹ du procédé.

Il existe généralement deux types de limites :

- les limites de spécifications : ce sont les limites qu'il ne faut pas franchir, sinon les plaques seront rejetées car il y a un risque de dysfonctionnement. Ces limites sont souvent symbolisées par « LSL » et « USL » limite basse et limite haute.

¹⁰ Expériences faites par des ingénieurs soit pour explorer un phénomène physique, soit pour qualifier un procédé.

¹¹ Nous utilisons généralement la notion de « capabilité des procédés » (ou C_{pk}) pour analyser la variabilité des procédés et des mesures électriques en regard des spécifications associées.

- les limites de contrôle sont les barrières qualifiant le bon fonctionnement d'une opération. Ces limites sont souvent symbolisées par « LCL » et « UCL » limite basse et limite haute.

1.1.5.1.b Echantillonnage des mesures

Dans un milieu de production, les plaques ne peuvent pas toutes être mesurées. Seuls, quelques lots sont mesurés et seules quelques plaques d'un même lot sont mesurées également. Cette méthode est appelée échantillonnage des mesures. Cette méthodologie s'applique à 3 types de mesures comme cela est illustré sur la figure 1.4.

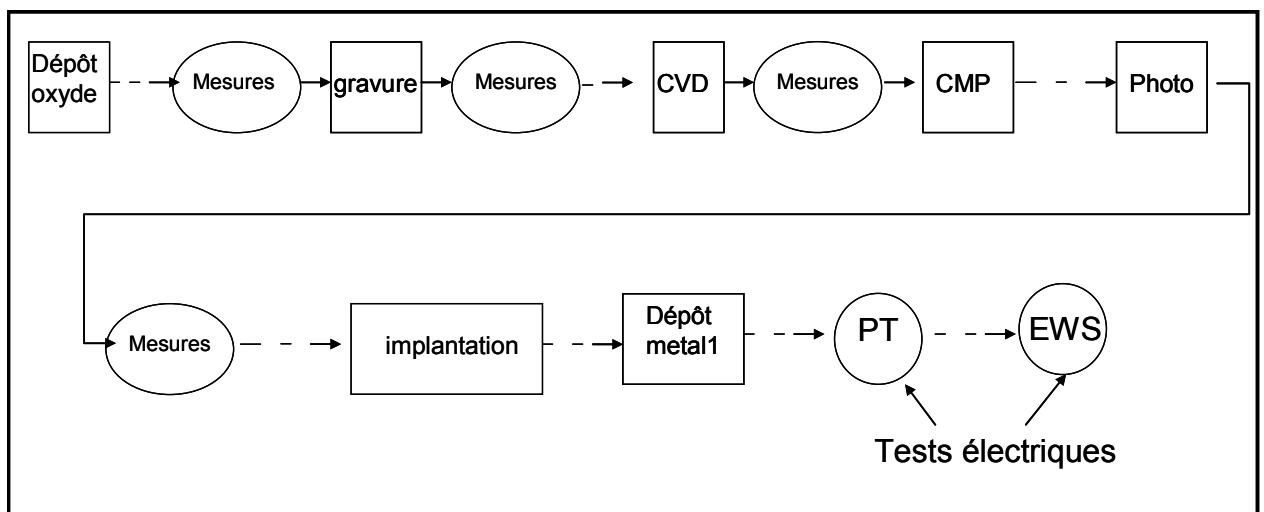


Figure 1.4- Enchaînement des opérations de fabrication des circuits

Nous pouvons distinguer :

- métrologie ou mesures physiques : à la fin de chaque opération de fabrication, les plaques sont mesurées pour s'assurer du bon déroulement de l'opération. Généralement, les mesures se font sur des structures de test et non sur le produit pour ne pas l'abîmer.
- tests paramétriques (PT: parametric test) : à la fin de la formation de la grille et des interconnexions, des tests sont faits sur chacune des plaques. Ces tests sont réalisés sur des structures de test. Le test PT est considéré comme un test préliminaire pour la sélection des bonnes plaques avant le test électrique final.
- test électrique Final (EWS: Electrical Wafer Sort): le test EWS¹² est le dernier maillon de la chaîne de mesures de fabrication des composants. Le test permet de tester la qualité des circuits.

1.1.5.1.c Maintien du procédé sous contrôle

¹² Chaque circuit est testé individuellement sur les paramètres qui le caractérisent, il s'agit des tests électriques. Si un test est hors spécification, alors le circuit est déclaré mauvais. Le pourcentage de puces bonnes (qui fonctionnent) après ce test correspond au rendement. C'est une donnée économique essentielle et un critère décisif de compétitivité de l'unité de fabrication.

En analysant finement la dispersion d'un procédé, on peut extraire deux causes essentielles de dispersion. Il s'agit de perturbations communes qui sont liées à des phénomènes aléatoires et de perturbations non aléatoires (ou spéciales) qui sont des causes de dispersion identifiables. Ces perturbations peuvent être modélisées. Contrairement aux perturbations aléatoires, les perturbations non aléatoires nécessitent une intervention sur le procédé.

1.1.5.2 Détection et classification des défauts (FDC : Fault Detection and Classification)

La méthode classique pour suivre le bon fonctionnement des équipements de fabrication consiste à surveiller les équipements après une étape de fabrication. Ceci induit un temps de cycle énorme et un risque de rejet de plaques important. Donc, ces méthodes de contrôle de surveillance des équipements ne sont plus adaptées aux nouvelles technologies où la perte liée au nombre de plaques rejetées constitue un enjeu économique très important. En vue de détecter les dérives (écart entre les valeurs des mesures et les valeurs cible) des équipements en temps réel et d'éventuellement déduire les origines de ces dérives et de classifier les anomalies, un contrôle en temps réel s'impose. C'est le principe de la FDC. Les outils FDC utilisent des systèmes d'analyses univariées et multivariées. Un exemple d'une analyse univariée est illustré sur la figure 1.5. On observe l'évolution temporelle de la pression de l'équipement avec le temps durant un procédé de dépôt chimique en phase vapeur. Cette méthode de contrôle nous permet de surveiller ce paramètre pour chaque étape de dépôt. Si les valeurs de la pression dépassent les limites de contrôle fixées, une alarme sera générée et l'équipement sera arrêté.

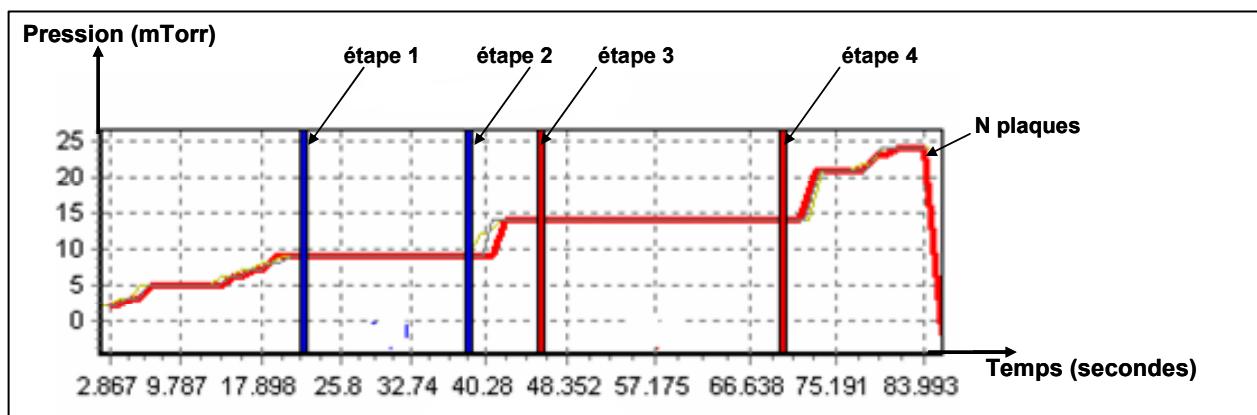


Figure 1.5- Représentation temporelle de l'évolution de la pression pendant le procédé de dépôt chimique en phase vapeur

1.1.5.3 Boucles de régulation Run to Run (R2R)

Chaque étape du procédé de fabrication induit une variabilité au niveau des plaques. Cette variabilité peut être traduite de batch¹³ à batch σ_{batch} , de lot à lot σ_{lot} (variabilité inter-lot), de plaque à plaque σ_{plaque} (variabilité inter-plaque), de site à site σ_{site} (variabilité intra-plaque), des variabilités liées à l'équipement de mesure $\sigma_{équipement}$ de mesure et enfin des

¹³ Le batch est défini comme un ensemble de deux ou plusieurs lots, par exemple lors d'une opération de diffusion, 6 lots sont disposés ensemble dans le four. On parle donc d'un batch de 6 lots)

variabilités liées à la nature de produit $\sigma_{produit}$. La variabilité non expliquée constitue les résidus¹⁴. On définit la variabilité totale alors comme étant la somme :

$$\sigma_{Tot}^2 = \sigma_{batch}^2 + \sigma_{Lot}^2 + \sigma_{Plaque}^2 + \sigma_{sites\ de\ mesures}^2 + \sigma_{Equipement\ de\ fabrication}^2 + \sigma_{Equipement\ de\ mesure}^2 + \sigma_{produit}^2 + Résidus \quad (1.1)$$

Afin de réduire ces variabilités, des boucles de régulations sont conçues. Le principe est de réajuster un ou plusieurs paramètres du procédé en temps réel en fonction des mesures physiques ou électriques des lots précédents ; ceci, suivant un modèle établi à partir des expériences ou à partir d'un historique de données regroupant des mesures de métrologie et des données équipements. Il faut noter qu'on ne peut réguler que sur les perturbations non aléatoires (perturbations qui peuvent être modélisées). On peut distinguer deux types de boucles de régulation :

- les boucles de régulation de type « Feed Forward ». Ce type de boucle est utilisé pour rattraper la variabilité (la dispersion) observée sur l'étape de fabrication N en modifiant les paramètres de la recette de l'étape (N+1). Ce type de boucles est connu aussi sous le nom de « Boucles inter-process » quand il s'agit d'agir sur le procédé (N+1) connaissant la dispersion du procédé N. Par exemple, dans le cas de la boucle entre le procédé de dépôt chimique en phase vapeur et le procédé de polissage mécano-chimique (CVD – CMP) comme c'est illustré sur la figure 1.6. La vitesse d'enlèvement est modifiée de plaque à plaque suivant les valeurs des épaisseurs de dépôt sur chaque plaque.

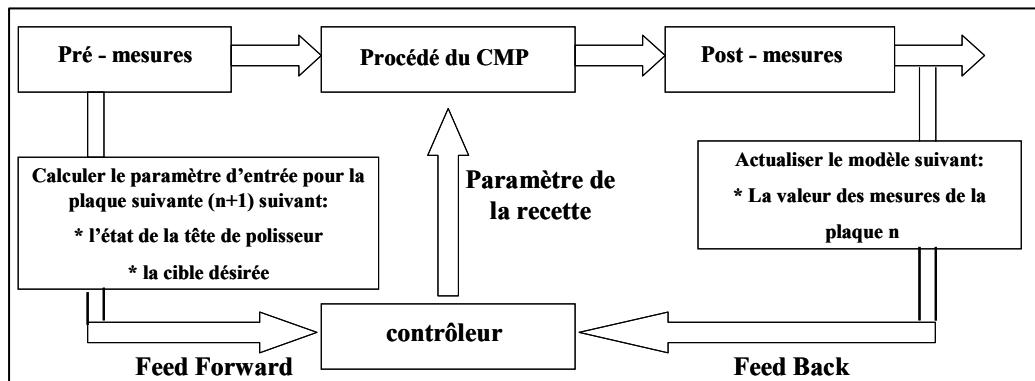


Figure 1.6- Schéma de deux boucles de régulation Feed back et Feed forward

- les boucles de régulation de type « Feed Back » sont déployées généralement sur un même équipement à partir des mesures de métrologie des lots précédents (N-1). L'écart de la valeur mesurée de la plaque du lot (N-1) permet d'estimer la valeur du paramètre équipement pour le lot (N). La figure 1.6 montre un exemple de boucle « Feed back » au niveau du procédé de polissage mécano-chimique (CMP). La valeur de l'épaisseur d'oxyde enlevée est renvoyée au contrôleur pour calculer un nouveau temps de polissage (une nouvelle vitesse d'enlèvement).

Ce concept de déploiement des boucles de régulation a fait ses preuves dans les sites de production dans le monde des semi-conducteurs. AMD a été le premier fabricant de composants à semi-conducteurs à se lancer dans l'automatisation de la FAB en implémentant des boucles de régulation (R2R) au niveau de l'étape de fabrication de polissage mécano-

¹⁴ La différence entre la valeur de la variabilité totale observée et la variabilité totale estimée.

chimique (CMP). Le rendement de fabrication au niveau de cette opération de fabrication a été amélioré grâce à l'implémentation de ces boucles.

Toutefois avant de déployer une telle boucle, il faut identifier les étapes critiques¹⁵. Ceci passe généralement par une étude de la variabilité de l'état actuel des procédés et l'estimation du gain attendu avec le contrôle.

Nous verrons dans l'état de l'art qui suit les différents travaux relatifs aux techniques de contrôle dans le milieu académique ainsi que le succès qu'ont eu ces techniques dans le milieu industriel.

1.2 Etat de l'art sur le contrôle et la modélisation des procédés

Dans cette partie, nous exposerons un état de l'art détaillé sur le contrôle et la modélisation des procédés au niveau académique et industriel. Suite à cet état de l'art, une stratégie sera définie pour déterminer les actions à effectuer en vue du déploiement des boucles de régulation R2R

1.2.1 Travaux sur le contrôle et la modélisation des procédés

Les problèmes de contrôle des procédés en microélectronique peuvent être classés en quatre catégories [EDGAR 2000]:

- la gestion ou le management de l'usine (FAB)
- le contrôle des contaminations (défectivité)
- la manipulation des matériaux
- le contrôle des opérations de fabrication

Sur ce dernier point, beaucoup d'efforts ont été déployés au niveau de l'ordonnancement des lots et au niveau du flux de transfert des lots sur les différents équipements qualifiés. Par contre, moins d'efforts ont été mis en œuvre dans le cadre de l'amélioration du contrôle des différentes opérations de fabrication.

Les principales opérations de fabrication sont : l'oxydation, le dépôt (de diélectriques et de métaux), la diffusion, le dépôt physique en phase vapeur, l'implantation ionique, la photolithographie¹⁶, la gravure¹⁷ (sèche ou humide) et le polissage mécanico-chimique. Les premières études de modélisation et de contrôle ont été focalisées sur les opérations de photolithographie, de dépôt et de gravure sèche. Des travaux sur la modélisation et le contrôle des opérations de fabrication (photolithographie, dépôt et gravure sèche) ont été présentés par Badgwell en 1995[BADGWELL 1995]. Dans le milieu industriel, les procédés de dépôt et de gravure sèche sont modélisés de manière empirique. Cette modélisation est souvent faite avec l'aide de la connaissance approfondie des mécanismes physiques/chimiques régissant les procédés. Comme les limites de spécifications deviennent de plus en plus serrées pour les

¹⁵ Etapes ayant une influence importante sur la suite des opérations.

¹⁶ Une technique qui permet de délimiter des zones à la surface de la plaque pour les traiter différemment du reste de la surface lors d'une opération ultérieure (gravure, dopage).

¹⁷ Une technique qui permet d'attaquer un matériau présent sur la plaque dans des zones non protégées par un masque.

nouvelles technologies, ce manque de connaissances approfondies des mécanismes physiques/chimiques régissant les procédés va poser problème.

Le contrôle des procédés de fabrication de composants à semi-conducteur a été étudié par à un certain nombre d'universités et de laboratoires de recherche industrielle parmi lesquels, U.C. Berkeley [LEE 1995] et [MAY 1991]. Leur travail portait sur la modélisation statistique et le contrôle des équipements de gravure plasma. Les équipes de l'université du Michigan [HAMBY 1998] travaillaient sur les boucles R2R en utilisant une modélisation multivariée et un contrôle en temps réel. Le laboratoire MIT [BONING 1995, BONING 2004] s'est focalisé sur l'étude des capteurs et sondes pour les équipements et les systèmes de contrôle. Les groupes de recherche de l'université du Maryland [ADIVIKOLANU 2000, BARAS 1996, ZAFIRIOU 1995] ont développé beaucoup de travaux relatifs aux algorithmes R2R. De son côté, SEMATECH¹⁸ [BAKSHI 1997], a publié plusieurs papiers relatifs aux problèmes rencontrés par la FDC sur les équipements de gravure sèche.

Le contrôle non-linéaire des opérations de R2R a été proposé par [DEL CASTILLO 1998]. À l'université d'Austin ont été développées des boucles de régulation à base de cartes de contrôle de type moyenne mobile avec pondération exponentielle (EWMA : Exponentially Weighted Moving Average) et de double EWMA qui tiennent compte du retard de la métrologie [Good 2002] et [Good 2005]. Le but de ces travaux est d'apporter une robustesse et une stabilité¹⁹ aux boucles. Ils ont aussi travaillé sur la surveillance statistique multivariée des équipements RTA (Rapid Thermal Annealing) ou bien (Recuit Thermique Rapide), sur les équipements de gravure [Yue 2000] et [Yue 2001] ainsi que sur le contrôle statistique multivarié des données de métrologie des dimensions critiques de la grille du transistor, ces dimensions sont généralement symbolisées par CD (Critical Dimensions) [Cherry 2003]. D'autres applications de contrôle avec la FDC sont rapportées par Del Castillo et Hurwitz [DEL CASTILLO 2000] et par James Moyne [MOYNE 2001].

1.2.2 Le contrôle avancé des procédés (APC : Advanced Process Control) dans le milieu industriel

1.2.2.1 Chez ST Microelectronics

En 2000, une boucle de contrôle a été mise en place sur les équipements de diffusion²⁰. Les équipements sont des fours fonctionnant à de hautes températures. Pendant la même année, une autre boucle a été réalisée sur les équipements de photolithographie au niveau de la fabrication de la grille du transistor. L'objectif de cette boucle était de contrôler le CD et l'indice de recouvrement appelé autrement Overlay au niveau de la fabrication de la grille du transistor. Le logiciel qui gère la gestion de cette boucle est « EPCC ». Ce programme a été développé sur ST-Crolles. En 2002, une boucle de régulation au niveau du procédé de polissage mécanico-chimique (CMP) a été réalisée. Le principe de cette boucle était de réajuster le temps de polissage afin de garantir la même épaisseur enlevée. Le réajustement ne se fait pas de lot à lot mais après chaque tâche de qualification²¹ sur l'équipement. En 2005, une boucle sur les équipements de dépôt d'oxyde (CVD) a été réalisée. L'objectif de cette boucle était d'automatiser les tâches de qualifications qui se

¹⁸ SEMATECH un consortium réunissant les principaux fabricants de semi-conducteur

¹⁹ Le bon fonctionnement d'un procédé de contrôle actif est conditionné par la stabilité comportementale du système en boucle fermée. Il est courant d'étudier les pôles du système afin d'en déterminer ses conditions de stabilité.

²⁰ Procédé thermique permettant l'obtention des films d'oxydes

²¹ Les premières opérations pour qualifier un procédé.

faisaient avant manuellement. En 2006, une boucle de type feed-forward a été réalisée sur le module de la grille et en 2007, les boucles sur le module STI faisant partie de mon sujet de thèse ont été mises en pré-production. Nous reviendrons en détails sur le principe de ces boucles dans le chapitre 4.

Sur le volet du déploiement de la FDC, le taux de déploiement des outils de la FDC a touché la plupart des équipements utilisés dans la FAB (200mm) comme c'est illustré sur la figure 1.7.

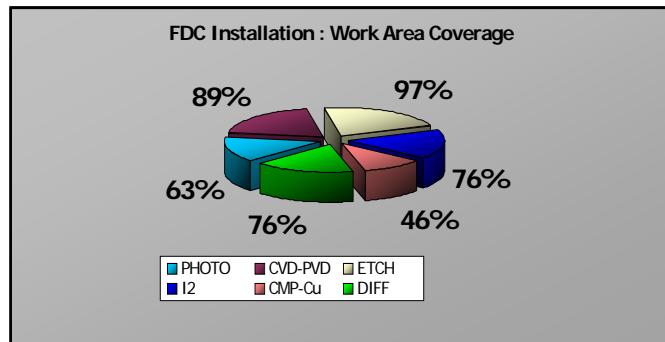


Figure 1.7- Le taux de déploiement de la FDC par type d'équipements

On constate que sur les équipements de gravure (etch), 97% des équipements sont équipés de la FDC. Le déploiement de la FDC a commencé en 2004. Le taux de déploiement est en forte progression, comme cela est illustré sur la figure 1.8. Il est évident qu'avant de lancer la FDC sur un équipement donné en production, il faut déjà installer l'outil FDC pour collecter le maximum de données. Si le mécanisme de collecte de données et le nombre de variables choisies à surveiller sont satisfaisants, il faut ensuite passer à l'étape de la mise en production. Ce protocole de la mise en production explique l'écart de pourcentage entre les outils de la FDC installés et les outils matures²² comme cela est aussi illustré sur la figure 1.8.

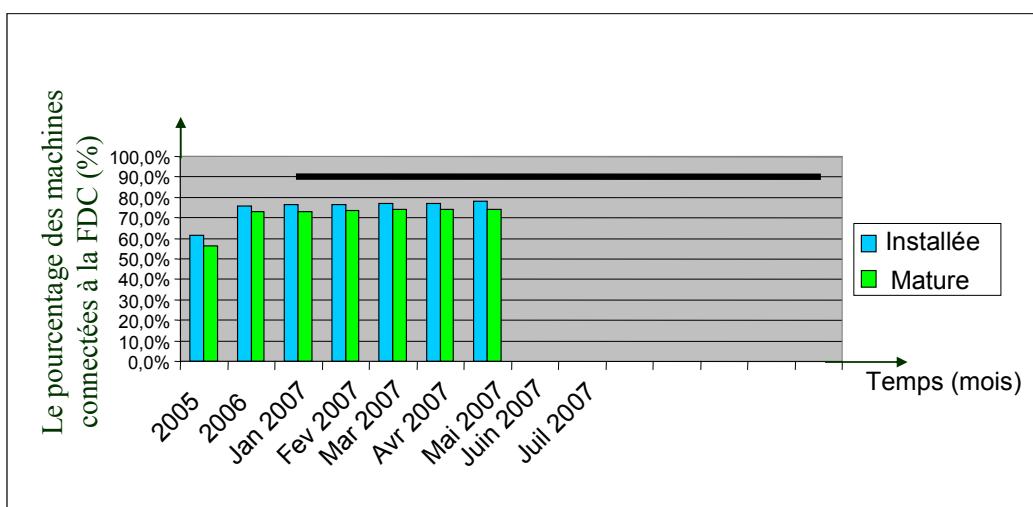


Figure 1.8- Taux de déploiement de la FDC en fonction du temps

Les outils de FDC à installer diffèrent d'un équipement à un autre car certains fabricants d'équipements développent leurs propres outils de FDC et par conséquence ils n'acceptent pas que d'autres outils soient connectés sur leurs équipements. On peut citer : la

²² Mise en production.

société TEL (spécialisée dans la fabrication des fours) qui a développé un logiciel « Ingénio ». Un autre fabricant, Lam Station a développé un logiciel pour le suivi de la FDC sur les équipements de gravure sèche. Les autres équipements sont connectés pour la plupart par les outils « Maestria » développés par la société PDF Solutions. On peut voir cette tendance en fonction du temps sur la figure 1.9.

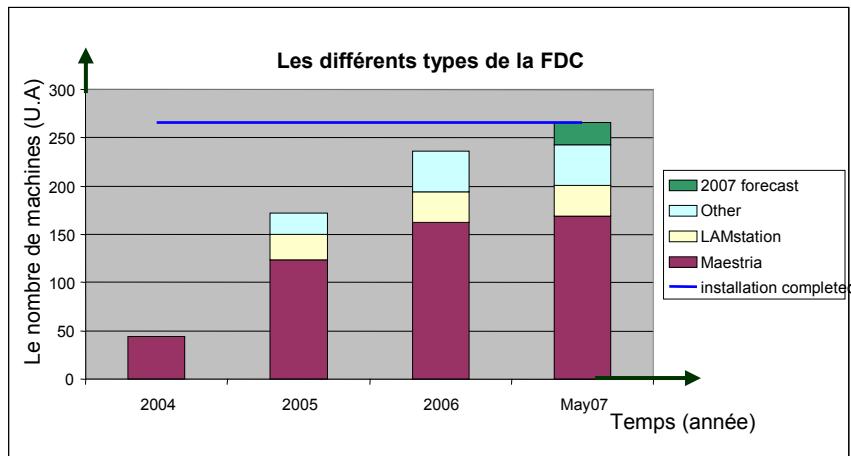


Figure 1.9- Les différents types des outils de la FDC en fonction de temps

En mai 2007, le nombre d'équipements équipés de la FDC est évalué à 241 sur un nombre total de 308 équipements. Comme pour toutes les autres compagnies de fabrication des semi-conducteurs, l'objectif est d'atteindre la barre des 100% [LAHAV 2005]. ST Microelectronics continue ses actions pour améliorer le taux d'équipements équipés par la FDC pour atteindre à terme la barre des 100%.

1.2.2.2 Les autres compagnies

Les réalisations APC dans les autres compagnies de fabrication de semi-conducteurs

- INTEL : le numéro 1 mondial dans le domaine des semi-conducteurs a commencé ses activités APC en 1996 (premiers travaux pour le déploiement de la FDC suivis des premières études sur le R2R). Pendant la période 2000-2002 des tests intensifs ont été effectués sur les logiciels de gestion des boucles de régulations. Les équipements dédiés aux technologies 130/90/65 nm ont vu un taux de déploiement de 75 à 85%. Le taux des boucles de régulation type R2R est estimé à 50-65%. Les technologies concernées par les boucles de régulations sont les plus modernes (130/90/65 nm). Intel a transféré également sur ses sites de production les technologies avancées avec les modèles associés.
- AMD : le numéro 2 mondial dans le domaine de la fabrication de microprocesseurs a commencé ses activités en APC en 1996 en développant le logiciel CATALYST. AMD est en concurrence avec INTEL sur les dispositifs rapides. AMD [SONDERMAN 2005] a réalisé en 2001 les premières boucles de régulation en production sur les blocs critiques de fabrication des circuits (grille, STI...) sur les technologies 180/130 nm. Par exemple, une boucle R2R a été mise en place au niveau du bloc Photolithographie où, 35% de réduction de l'overlay a été observée. AMD a réussi avec succès son transfert des technologies sur sa FAB à Dresden.
- IBM a commencé ses activités pendant la période 1996/1997. Pour les applications FDC, IBM s'est associé avec SCHNEIDER ELECTRIC pendant 4-5 ans. Cette

collaboration a donné un résultat supérieur à 150 applications (150 recettes sur équipements). Pour la gestion des différentes boucles de régulation, IBM a testé CATALYST (développé par AMD) et PROCESSWORK (développé par ADVENTA). Pour le déploiement de la FDC, IBM a atteint un taux de déploiement de FDC allant de 55 à 85% sur les équipements dédiés aux technologies 130/90/65 nm. Pour ces mêmes technologies, le taux de réalisation de boucles de régulation R2R est de 45 à 75%.

- Texas Instruments (TI): la division de l'automation de TI composée de 250 à 300 personnes a commencé les travaux de déploiement des techniques APC depuis 1996. TI dispose de FABs de 200 et 300mm. La FAB DMOS4 (200mm) qui produit des dispositifs (technologies 0,25- 0,35µm) et la FAB DMOS5 (200mm) qui produit des dispositifs (technologies 0,13- 0,18µm) ont vu un taux de déploiement de boucles de régulation R2R gérées par le logiciel PROCESSWORK. La FAB DMOS6 (300mm) présente un taux de déploiement des boucles de régulation R2R de 60 à 75%. Pour la FDC le déploiement est de 75 à 90%.

Nous venons de voir dans les paragraphes précédents que, la FDC est assez développée par rapport à une autre composante du contrôle avancé des procédés à savoir les boucles de régulations (R2R). Plusieurs actions doivent être déployées pour développer les axes de recherches relatifs aux boucles de régulations. Nous allons développer certains de ces grands axes de recherches.

1.2.3 **La conception de contrôleur R2R**

Le contrôleur R2R supervise les analyses de(s) lot(s) précédent(s) pour adresser de nouvelles valeurs de paramètres d'entrées pour les prochains lots. Il y a plusieurs manières de construire le contrôleur R2R. Cependant, la majorité des contrôleurs construits ont presque tous une structure semblable.

Généralement, les contrôleurs sont basés sur un modèle du procédé (process model) [EDGAR 1999],[Mullins 1997]. Le dernier élément du contrôleur R2R est lié à la loi de contrôle qui spécifie comment et quand il faut actualiser la recette du procédé. La régression linéaire et les modèles avec réponse de surface sont souvent employés pour le contrôle basé sur le modèle du procédé. Nous allons détailler ce type de modèle dans les prochains chapitres. La majorité des modèles sont étudiés en régime permanent²³. Des modèles dynamiques ont été étudiés par un certain nombre de chercheurs (Box.1993, Mullins.1997) dans le but de modéliser la dérive intervenue sur des lots processés consécutivement.

Quand une boucle de régulation R2R de type feedback par exemple est affectée par du bruit, il est nécessaire de rajouter un bloc appelé observateur pour estimer l'état actuel du procédé. Le rôle de l'observateur est d'estimer les erreurs qui sont corrélées et surtout d'ignorer les variations aléatoires. Le type de l'observateur peut être :

- simple : moyenne des erreurs consécutives
- sophistiqué comme le filtre de Kalman.

L'observateur opère généralement dans un de ces deux modes :

²³ Le régime de fonctionnement d'un système lorsqu'il est soumis à une excitation invariante dans le temps, la dérive est supposée lente.

- soit dans le mode graduel ou progressif. Ce mode est employé pour les procédés dérivant lentement. Dans ce mode on suppose que la variance des paramètres de sortie est attribuée à une variation naturelle du procédé.
- soit dans le mode rapide. Ce mode est employé quand un procédé subit une dérive (écart entre la mesure et la cible) assez significative.

Dans certains cas particuliers, le mode rapide est nécessaire, par exemple dans le cas de l'étape faisant suite au nettoyage d'un équipement. Le problème qui se pose est comment alterner les deux modes ? En d'autres termes, quel mode faut-il déclencher ? Plusieurs chercheurs ([Box.1992, Montgomery.2001]) ont étudié l'utilisation de cartes de contrôle SPC. Cette technique repose sur la surveillance des résiduels du modèle prédictif. Si les résiduels²⁴ sont à l'intérieur de la bande de contrôle, le mode progressif est déclenché. Dans le cas contraire, ça veut dire pour un état « hors contrôle » le mode rapide est déclenché. D'autres chercheurs (Sachs, 1991) ont suggéré l'emploi d'un système automatique. Ce système est employé dans le mode rapide pour estimer où la dérive s'est produite. Dans le domaine des semi-conducteurs, les contrôleurs sont souvent basés sur les cartes de contrôle EWMA. Généralement le modèle du procédé suit une régression linéaire du type :

$$Y_k = (1 - \lambda)Y_{k-1} + \lambda \bar{x}_k \quad (1.2)$$

Où : Y_k est la moyenne pondérée de paramètre de sortie.
 λ est un coefficient $0 < \lambda < 1$
 Y_{k-1} est la valeur de mesure correspondant au run (n-1)
 \bar{x}_k est la moyenne de toutes les observations.

Dans toutes nos études relatives à la conception de contrôleur R2R, le type d'observateur adopté est l'estimateur EWMA.

Après avoir dressé un état de l'art détaillé et les objectifs devant principalement répondre à un besoin industriel, une stratégie a été développée en vue du déploiement de boucles.

1.3 Stratégie de l'étude et outils statistiques

Dans cette partie nous présenterons la démarche suivie pour l'étude du déploiement des boucles de régulation. Cette étude a été adoptée pour une première boucle au niveau du dépôt. Les bons résultats de ce travail nous ont motivé à généraliser cette stratégie sur l'ensemble des opérations susceptibles d'être concernées par le déploiement de boucles de régulation.

²⁴ Ecart entre la valeur prédictive et la valeur de la mesure

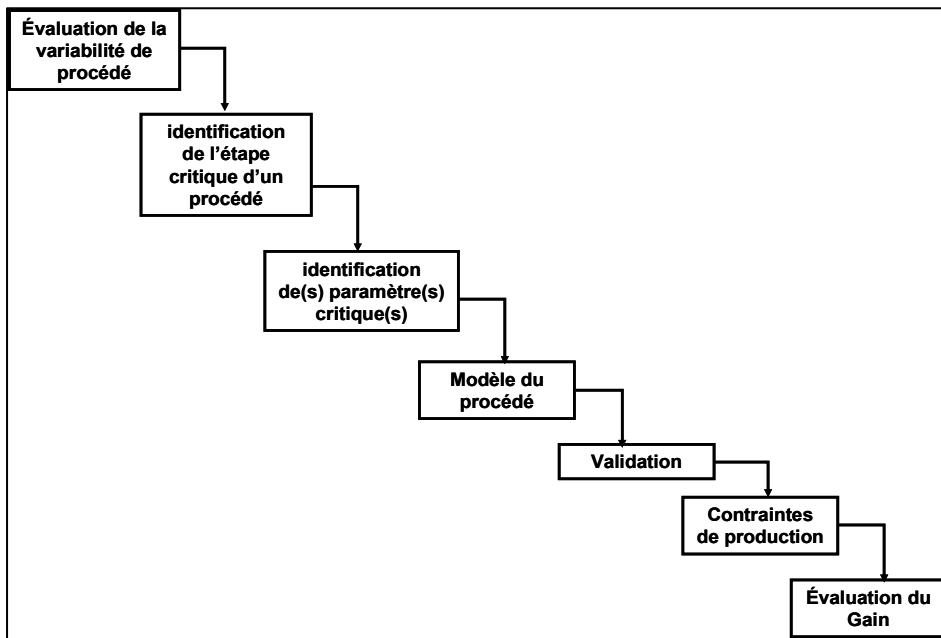


Figure 1.10- Stratégie de l'étude du déploiement d'une boucle de régulation

Cette stratégie a été développée pour exprimer le besoin en termes d'outils statistiques indispensables pour l'accomplissement de cette étude. Dans les paragraphes suivants nous détaillerons tous les blocs illustrés sur la figure 1.10 et nous décrirons rapidement les outils statistiques utilisés tout au long de cette thèse

1.3.1 Evaluation de la variabilité du procédé

Pour comprendre les phénomènes physiques des procédés de fabrication en microélectronique, il est important d'analyser les données de production [BERGERET 2004] afin de comprendre l'évolution et l'impact d'une ou de plusieurs variables du procédé. Le terme « analyse des données» désigne les étapes nécessaires pour la réalisation d'une étude de données [BESSE]. Avant de se lancer dans des analyses de données existantes durant cette étape, il est primordial de définir l'ensemble des paramètres à étudier.

1.3.1.1 Collecte des données

Dans le cadre de cette thèse, les paramètres en question sont :

- les mesures de métrologie
- les mesures électriques
- les données FDC provenant des équipements
- le suivi du cheminement des plaques tout au long des étapes de fabrication. Cette information contextuelle indispensable pour toute analyse nous renseigne notamment sur la position de chacune des plaques dans les équipements de fabrication. Cette information permet de garantir qu'aucune inversion de plaques n'est réalisée à l'intérieur d'un même lot et de connaître la position de chaque plaque dans les équipements de fabrication multi- plaques

Suivant le type de données, des bases de données sont disponibles et accessibles pour une exploration des données sur un historique limité dans le temps. Les données ne sont pas gardées généralement plus de 6 mois pour ne pas saturer les bases de données.

Sur le site de ST-Rousset on distingue :

- une base de données mesures physiques
- une base de données mesures électriques appelée autrement PT Analysis.
- une base de données sur les mesures électriques après le test final des puces

L'extraction des données sur ces bases de données est directe. Toutes ces bases de données sont accessibles via le logiciel SAS. Il faut noter que pour les paramètres FDC de chaque équipement, la base de données est en cours de construction. Pour exploiter ces données, l'extraction est possible uniquement sur les stations (ordinateurs) des différents équipements.

Chaque base de données doit être accessible par les différents outils d'analyses afin que ceux-ci puissent importer les données choisies dans leur environnement propre. On distingue ainsi deux types d'accessibilité aux données :

- accessibilité directe : les outils d'analyse sont connectés aux différentes bases de données existantes. L'extraction des données est réalisée directement dans l'environnement de l'outil, aucun extracteur intermédiaire n'est nécessaire.
- accessibilité indirecte : les outils d'analyse ne sont pas connectés aux bases de données. Une interface intermédiaire est nécessaire pour extraire et structurer les données des différentes bases.

1.3.1.2 Analyses des données

A la fin de l'étape de la collecte des données, des tableaux sont constitués. Chaque colonne représente un paramètre équipement (variable) et chaque ligne représente une observation (la valeur du paramètre équipement). Ces données doivent être vérifiées et donc mises en forme en supprimant les valeurs aberrantes (dues par exemple à une erreur de saisie). Ensuite, il faut évaluer la distribution de chacune des variables.

1.3.1.2.a Répartition des données

Après la collecte des données, la première phase consiste à analyser la répartition de l'ensemble des données. Cette analyse est souvent faite d'une manière rapide grâce aux « box plot ». Le box plot est un graphique synthétique permettant de représenter sur un même graphique des statistiques tels que la médiane (barre à 50 % de la population), les quantiles (25% et 75%) et les déciles (10% et 90%). Il rend aussi compte du niveau d'asymétrie, de la dispersion et des valeurs extrêmes de la distribution. La figure 1.11 illustre un exemple d'un box plot regroupant un historique de données de mesures d'un paramètre (hauteur de marche²⁵).

²⁵ C'est un paramètre de mesure qui caractérise l'isolation des transistors.

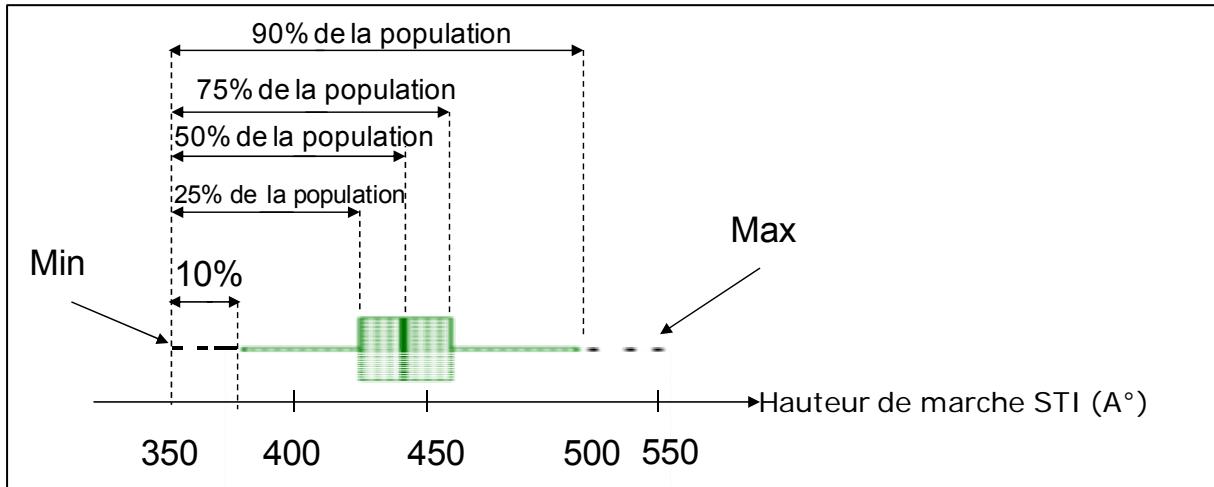


Figure 1.11- Exemple de représentation d'un box plot

1.3.1.2.b Prétraitement des données

Avant toute analyse de données, il est nécessaire de pré-traiter les données. Généralement, cette phase de prétraitement de données comporte deux étapes :

Normalisation des données :

Les variables observées sont rarement à la même échelle. On traite par défaut toutes les variables de façon équivalente pour ne se concentrer que sur les écarts. Pour chacune des variables, les valeurs sont divisées par l'écart type calculé sur cette variable, puis sont centrées. On dispose ainsi de données comparables pour la construction des modèles.

Distribution des données :

En général, les algorithmes utilisés nécessitent une distribution gaussienne. Afin de vérifier cette hypothèse, deux indicateurs sont généralement employés : le skewness et le kurtosis. Ces deux indicateurs nous renseignent sur l'écart existant entre la distribution du paramètre et une loi gaussienne.

- l'indicateur statistique Skewness ou coefficient d'asymétrie nous renseigne sur le manque de symétrie de la distribution des données. Ce paramètre est défini par :

$$Skewness = \frac{\sum_{i=1}^n \frac{(x_i - \bar{x})^3}{n}}{\sigma^3} \quad (1.3)$$

Avec x_i : valeur de l'observation i

\bar{x} : valeur moyenne de l'ensemble des observations.

n : nombre d'observations

σ : écart type de l'ensemble des observations.

Si le skewness est égal à 0, la distribution des données est symétrique. Plus cette valeur est proche de 1 en valeur absolue, plus la distribution des données est asymétrique.

- l'indicateur Kurtosis ou coefficient d'aplatissement nous renseigne sur la forme de la distribution des données. Ce paramètre est défini par :

$$Kurtosis = \frac{\sum_{i=1}^n \frac{(x_i - \bar{x})^4}{n}}{\sigma^4} \quad (1.4)$$

Avec x_i : valeur de l'observation i

\bar{x} : valeur moyenne de l'ensemble des observations.

n : nombre d'observations

σ : écart type de l'ensemble des observations.

Le coefficient d'aplatissement est égal à 3 pour une loi parfaitement gaussienne. Lorsque le coefficient de kurtosis est inférieur à 3 alors la forme de la distribution est plus aplatie que celle d'une loi gaussienne, tandis que lorsque ce coefficient est supérieur à 3, la distribution est plus concentrée.

1.3.1.3 Etude la variabilité

Les premières investigations visent à quantifier la dispersion et la nature de la dispersion. Il est recommandé de disposer d'un maximum d'observations sur une durée de temps importante incluant les différentes interventions sur la machine (comme les maintenances préventives). Quand la dispersion est très importante, il est nécessaire de quantifier son impact sur le rendement final pour justifier le déploiement des boucles de régulation.

Chaque étape du procédé de fabrication induit une variabilité au niveau des plaques. Cette variabilité peut être traduite de batch à batch σ_{batch} (le batch est défini comme un ensemble de deux ou plusieurs lots, par exemple lors d'une opération de diffusion, 6 lots sont disposés ensemble dans le four. On parle donc d'un batch de 6 lots), de lot à lot σ_{lot} (variabilité inter-lot), de plaque à plaque σ_{plaque} (variabilité inter-plaque), de site à site σ_{site} (variabilité intra-plaque), des variabilités liées à l'équipement de mesure et enfin des variabilités liées à la nature de produit $\sigma_{produit}$. La méthode d'analyse statistique prédictive d'analyse de la variance permet de décomposer la variance totale suivant les différents effets possibles (σ_{batch} , σ_{lot} , ...).

1.3.1.4 Alignement des sites de mesures

1.3.1.4.a Cartographie de mesures

Le nombre de plaques à mesurer répond à un échantillonnage adopté et le nombre de sites de mesures diffère d'une opération à une autre. À ST-Rousset, la cartographie²⁶ de 9 points est la plus souvent utilisée. Pour des tâches de qualification ou des expériences spécifiques une cartographie complète, incluant plusieurs points allant jusqu'au bords de la plaque peut être utilisée (49 points par exemple). Un exemple de cartographie en 49 points est

²⁶ La recette de la mesure ou bien l'ensemble des points répartis sur la plaque définis pour la mesure.

illustré sur la figure 1.12a. La figure 1.12b illustre un exemple de topographie de dépôt d'oxyde obtenu avec l'aide d'une cartographie complète.

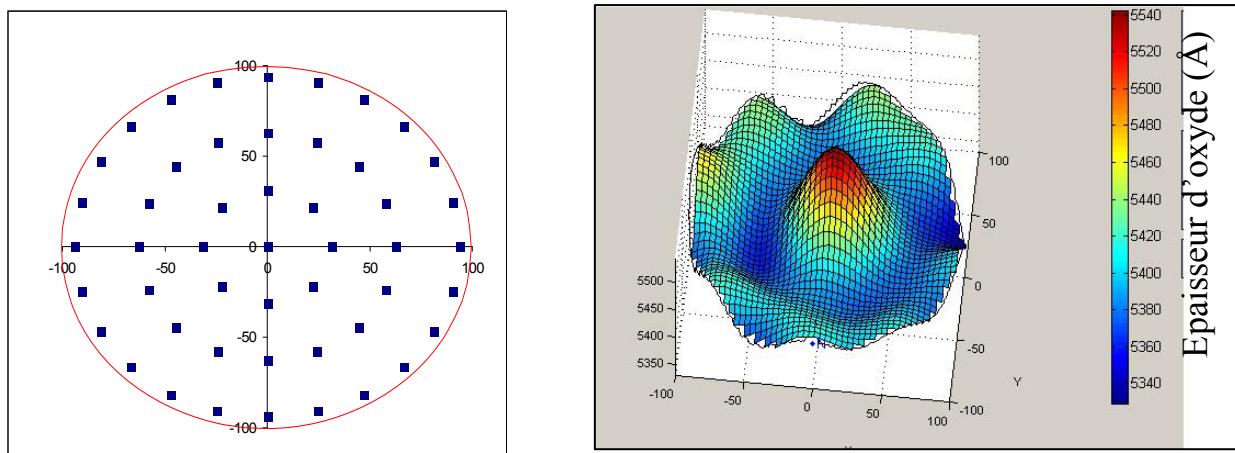


Figure 1.12- a) Cartographie en 49 points b) Profil de la plaque après une opération de dépôt CVD mesuré en 49 points

Certaines mesures prennent beaucoup de temps, telle que la mesure par AFM (microscopie à force atomique). Dans ce cas, le nombre de plaques à mesurer est réduit et le nombre de sites de mesure est réduit également. Pour l'opération de gravure sèche, une cartographie de 5 points est utilisée par exemple. Les cartographies de 9 points ne reflètent pas la bonne topographie du profil. Nous verrons plus tard dans le chapitre 4 que le fait de négliger des effets de bord peut être fatal.

Généralement, les mesures se font à l'aide de structures de test (mesures non destructives). La localisation des points de mesures diffère donc d'une opération à une autre. Pour obtenir de très bonnes corrélations entre les différentes opérations ainsi qu'une bonne étude de la variabilité, un alignement des sites de mesures et l'estimation des points qui ne sont pas mesurés est nécessaire [ALEGRET 2005]. L'extrapolation de la surface est généralement la méthode employée pour l'alignement des profils de mesures. L'extrapolation de la surface nécessite toutefois une excellente qualité de modélisation. Cependant, la plupart des logiciels disponibles n'offre pas cette qualité d'extrapolation car ils utilisent des algorithmes de type polynomiaux pour réaliser les interpolations. Lors de l'interpolation, chaque interpolation est traitée individuellement sans tenir compte de la forme globale de la surface [ALEGRET 2006]. Cette modélisation offre une bonne qualité de modélisation proche des points de mesure mais une faible qualité loin de ces zones.

Pour combler ce manque de qualité de la modélisation de la surface globale, une solution est d'utiliser des algorithmes d'apprentissage [ROUSSOS 2005], [NERUDA 2005]. Une application permettant de modéliser l'ensemble de la surface de la plaque à partir d'une cartographie d'entrée a été développée [ALEGRET 2005]. Le logiciel Wafer Fit, réalisé par le groupe MASA, permet de réaliser cette fonction, grâce aux algorithmes de reconstruction de surface.

1.3.1.4.b Algorithmes d'apprentissage de type RBF (fonction radiale de base)

Les fonctions radiales de base (RBF) sont des sommes pondérées de translations de fonctions symétriques augmentées d'un polynôme de degré k. Elles s'expriment :

$$S(x) = p(x) + \sum_{i=1}^n \lambda_i \phi(|x - x_i|) \quad (1.5)$$

Avec : $p(x)$: polynôme de degré k
 λ_i : poids attribué à la fonction ϕ
 ϕ : fonction de base
 $|x - x_i|$: distance entre le point extrapolé et le point effectivement mesuré.

Plus réguliers que les interpolations polynomiales, les algorithmes de type RBF permettent d'apprendre une forme de surface en ajustant les coefficients de la fonction S(x).

Plusieurs analyses ont été faites sur des plaques mesurées en 9 points et en 49 points. Wafer fit a été utilisé pour traiter les données. Les comparaisons montrent que les nouvelles valeurs des indicateurs statistiques (moyenne et écart type) après interpolation sont plus fiables que celles de 9 points et plus proches des résultats des mesures de 49 points [BELHARET 2006 a]. Grâce à cette bonne modélisation, ces nouveaux indicateurs peuvent remplacer à long terme le « SPC » classique. Ce type d'interpolation peut être appliqué aux boucles de régulation en tant que module de pré-traitement.

Nous avons utilisé également le logiciel Wafer-Fit pour optimiser les cartographies des mesures. Cette optimisation consiste à réduire un grand nombre de points de mesures en seulement quelques points tout en ayant la même qualité des indices statistiques (moyenne et écart type). Nous allons voir dans la deuxième partie du chapitre 4 que nous pouvons réduire le nombre de points de mesures d'une cartographie de mesure de 460 à 109 points de mesures.

1.3.2 Identification de l'étape et des paramètres critiques

Après avoir obtenu des tableaux avec de bonnes valeurs, afin de comprendre les phénomènes physiques, de déterminer les paramètres et les étapes critiques, une étude multivariée s'impose. L'outil utilisé est le logiciel Simca-P développé par le groupe UMETRICS. L'approche multivariée est spécialement adaptée pour la modélisation de phénomènes physiques [WISE 1991]. La particularité de ces modélisations tient en la possibilité d'étudier un nombre important de variables (plusieurs dizaines) avec peu d'observations.

1.3.2.1 Analyses sous Simca-P

Le simca-P est un logiciel d'analyses multivariées de données proposant des analyses descriptives (PCA) [Principal Component Analysis] et prédictives (PLS) [Partial Least Squares] capables de modéliser de larges tableaux de données avec valeurs manquantes. Ces techniques permettent aussi de déterminer les corrélations entre les variables.

Les logiciels tels que : SAS, Statgraphics ont l'inconvénient de ne pas traiter les valeurs manquantes dans les tableaux. L'algorithme NIPALS employé par SIMCA-P permet de résoudre ce problème. Il faut juste souligner que la présence de tableaux avec des valeurs manquantes est courante dans le domaine des semi-conducteurs. Ce manque de valeurs peut s'expliquer soit :

- par les mesures qui ne sont pas faites (échantillonnage imposé)
- par les mesures qui sont aberrantes (erreurs de mesure)

Autre avantage, le logiciel Simca-P accepte les fichiers (txt,xls,csv). Pour toutes les analyses multivariées faites au cours de cette thèse, Simca-p a été le logiciel choisi.

1.3.2.2 Les méthodes d'analyse multivariées

1.3.2.2.a L'analyse en composantes principales

Rappel sur la technique de l'analyse en composantes principales

L'analyse en composantes principales est une technique descriptive permettant d'étudier les relations qui existent entre des variables quantitatives, sans tenir compte, a priori, d'une quelconque structure, ni des variables, ni des individus. L'analyse en composantes principales consiste à générer des variables synthétiques par combinaison linéaire des variables initiales. La génération des variables est un problème équivalent à la minimisation des distances de projections. On peut générer plusieurs axes de régression, construits orthogonaux deux à deux, c'est à dire décorrélés entre eux. Ces axes sont nommés « composantes principales ».

Les domaines d'applications de cette méthode sont très variés et de nombreux exemples sont proposés, notamment par [WOLD 2004], [YUE 2004], [CHERRY 2003].

Visualisation des résultats suivant les deux premières composantes

Le point de départ d'une analyse ACP est un tableau de données (n lignes et p colonnes). Le calcul de la moyenne et de l'écart type donne, pour chaque variable des informations concernant l'ordre de grandeur et la dispersion des données [YINGHUA 2002]. De même, le calcul de la matrice de corrélation des variables donne des indications sur l'évolution simultanée des variables prises deux à deux. Ces éléments de statistique descriptive univariée et bivariée ne donnent cependant aucune information sur le problème lorsque les p variables sont considérées simultanément. Cette étude simultanée des variables est précisément le but de l'analyse en composantes principales.

L'ACP est une technique qui est utile pour la compression et la classification des données. Le problème consiste à réduire la dimensionnalité d'un ensemble des données (échantillon) en trouvant un nouvel ensemble de variables plus petit que l'ensemble original des variables, qui néanmoins contient la plupart de l'information de l'échantillon [SMITH 1991]. La variation présente dans l'échantillon est donnée par les corrélations entre les variables originelles. Les nouvelles variables, appelées composantes principales (PC), sont non corrélées et sont ordonnées par fraction de l'information totale que chacune contient. Donc, le principal usage de la PCA, est de réduire la dimensionnalité tout en conservant le plus d'information possible.

PC 1: correspond à l'axe qui contient la plus grande quantité de dispersion.

PC 2: pour ce deuxième axe, on réitère la régression mais en se restreignant à ce que le premier axe n'a pu expliquer.

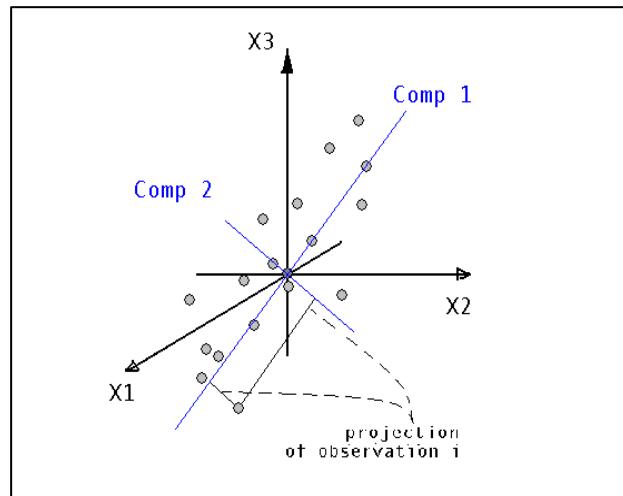


Figure 1.13- Principe de la régression ACP pour un système à 3 dimensions.

Comme cela est illustré sur la figure 1.13, on peut exprimer la régression de PC1 comme :

$$PC1 = \alpha.X_1 + \beta.X_2 + \gamma.X_3 \quad (1.6)$$

L'intérêt d'une telle analyse pour l'étude d'un phénomène complexe est de ne donner de valeur à une information que si elle se vérifie d'une façon globale.

Le cercle des corrélations

Les cercles des corrélations sont des graphiques visant à représenter géométriquement les variables dans le nouveau système de coordonnées.

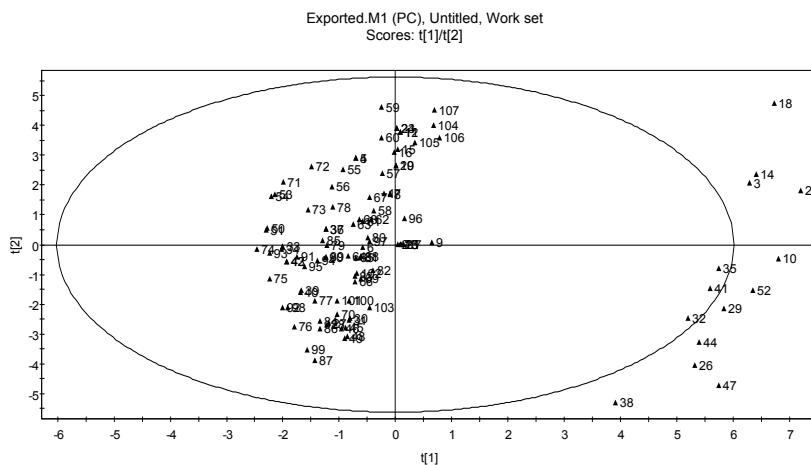


Figure 1.14- Représentation des observations dans le nouvel espace

La figure 1.14 montre la position de chaque point par rapport à la construction des composantes principales (axe t1 représente l'axe 1 et t2 représente l'axe 2) et par rapport à l'ensemble des données. Sur la figure 1.14, on distingue des points en dehors de l'ellipse, ces points sont des points aberrants.

1.3.2.2.b Méthodes des moindres carrés ou PLS (Partial Least squares)

Rappel sur la méthode des moindres carrés

La régression PLS généralise et combine les caractéristiques de l'analyse en composantes principales et de la régression multiple. On dispose de deux blocs de données, un bloc constitué de variables d'entrées explicatives et un bloc constitué de variables de sortie à expliquer. La régression PLS estime directement les coefficients d'influence sur Y_K attribuable à chaque variable X_i par le biais de composantes hiérarchisées qui matérialisent les faisceaux de relations constatées entre les tableaux des variables d'entrée et de sortie. De façon plus concrète, on cherche à définir les relations à l'intérieur d'un groupe de variables d'entrée mais à expliquer les variations constatées sur les variables de sortie en fonction des variations observées dans le tableau des variables d'entrée [TENENHAUS 1998].

Visualisation des résultats suivant les deux premières composantes

Géométriquement, pour ce type d'analyse, on représente l'ensemble des observations dans l'espace des variables X_i et l'espace des réponses Y_k . La figure 1.15 illustre une représentation des observations lors d'une analyse PLS où la première composante est une fonction qui caractérise au mieux les variations observées dans l'espace des X_i comme lors d'une ACP standard et qui définit aussi une bonne corrélation dans l'espace des Y .

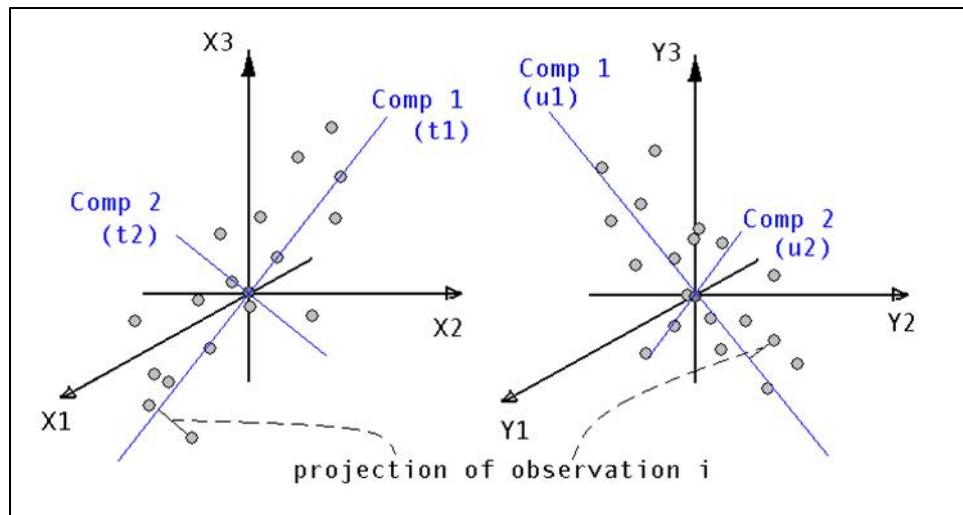


Figure 1.15- Représentation des observations lors d'une analyse PLS.

L'ensemble des observations t_1 des variables sur l'axe de régression CP1 est considéré comme une nouvelle variable. Cette variable est pondérée par un coefficient c_1 tel que

$$\hat{y}_{(1)} = c_1 * t_1 \quad \text{décrire au mieux les variations de } Y.$$

La différence entre la valeur estimée $\hat{y}_{(1)}$ et la valeur réelle Y constitue le résidu. Le résidu peut être défini comme étant la variation de Y non expliquée par la première composante définie dans l'espace des X . La deuxième composante est une fonction orthogonale à la première qui améliore autant que possible la description des variations des X donnant aussi une bonne corrélation dans l'espace des Y . L'ensemble des projections t_2 des

observations sur l'axe de régression CP2 est considéré comme une seconde nouvelle variable. Cette variable est pondérée par un coefficient c_2 tel que : $\hat{y}_{(2)} = c_2 * t_2$ décrire au mieux les variations du résidu obtenu après le calcul de la première composante.

Le cercle des corrélations

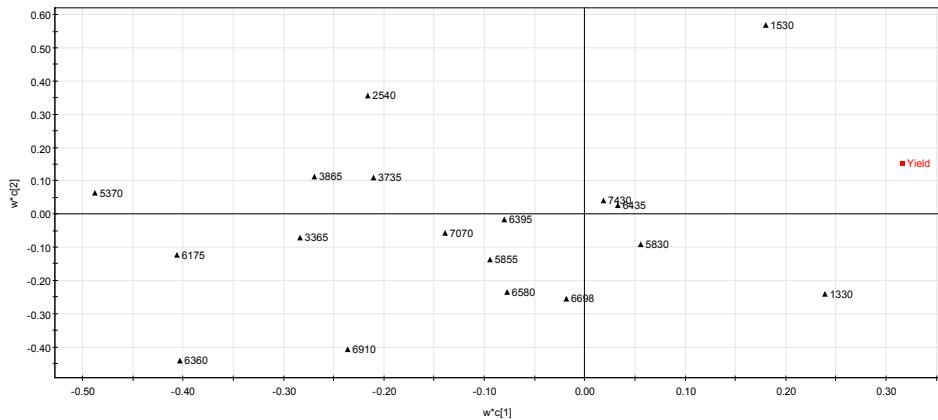


Figure 1.16- Les coefficients de corrélations des différentes variables sur les deux composantes

La figure 1.16 nous informe sur les coefficients de corrélation entre les paramètres d'entrée et la variable de sortie représentée dans cette figure « yield ». Les variables les plus éloignées du centre sont les plus influentes sur la variable de sortie. La variable 5370 présentant un coefficient de corrélation de (-0,5) est la plus corrélée sur la première composante, tandis que sur la deuxième composante, ce sont les variables 1530, 6360 et 6910 qui sont plus influentes. Avec des coefficients de corrélations sur les deux composantes très faibles, les variables 7430, 6435, 5830, 6395, 6698, 6580, 5855 n'influent pas sur la variable de sortie.

1.3.2.3 Indicateurs de qualité sous le logiciel SIMCA-P

Simca-P est un logiciel facile à utiliser. Toute la difficulté réside dans l'interprétation des résultats. Pour cela, il est très important de connaître les indicateurs de ce logiciel pour bien interpréter les graphes obtenus.

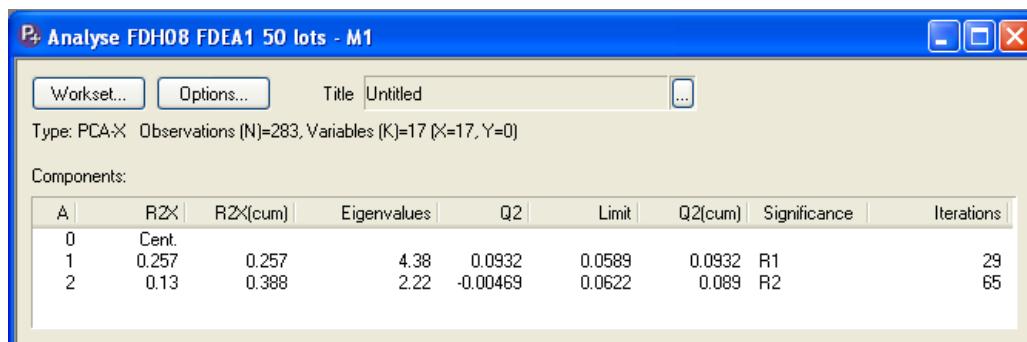


Figure 1.17- Résultats d'une analyse à partir de Simca-P

- R^2 est un facteur qui indique la qualité du modèle établi ou bien la fraction des X expliqués par le modèle. Une faible valeur de R^2 peut être due au bruit ou à des données non pertinentes ($0 < R^2 < 1$).
- Q^2 représente l'estimation de la capacité prédictive du modèle. Plus il est important, plus il est prédictif. Si Q^2 est négatif, le modèle n'est pas du tout prédictif ($0 < Q^2 < 1$).

En fonction des valeurs de R^2 et Q^2 , trois cas de figure peuvent se présenter :

- R^2 et Q^2 élevés : le modèle établi est explicatif et très prédictif. Dans ce cas, le modèle peut être appliqué à d'autres données pour la prédiction de la variable à expliquer à partir des variables explicatives.
- R^2 et Q^2 faibles : le modèle est peu explicatif et peu prédictif. Le fait que le modèle est peu explicatif signifie que les variables explicatives n'ont pas d'influence sur la variable à expliquer.
- R^2 élevé et Q^2 faible : le modèle est explicatif, mais le fait qu'il est peu prédictif rendra difficile la généralisation de ce modèle sur d'autres données.

1.3.3 Modèle du procédé

Le modèle du procédé représente le cœur du contrôleur. Il est recommandé d'avoir un modèle simple (linéaire) pour des raisons de simplicité mais aussi pour faciliter le transfert du modèle sur d'autres machines et éventuellement vers d'autres FABs. On distingue deux sortes de modèles :

1.3.3.1 Les modèles issus des données de production

Lorsqu'on dispose d'un nombre important de données sur un procédé donné, après traitement approfondi de ces données, il est alors facile d'obtenir un modèle. Le grand inconvénient dans ce type de modèle est que la variabilité des paramètres contrôlables n'est pas très importante. Il sera donc difficile de modéliser le comportement du procédé en dehors des plages habituelles des données (de production) observées.

1.3.3.2 Les modèles issus de plan d'expériences (DOE : Design Of Experiments)

On choisit a priori une fonction mathématique qui relie la réponse aux facteurs. On prend un développement limité de la série de Taylor-Mac Laurin. Les dérivées sont supposées constantes et le développement prend la forme d'un polynôme de degré plus ou moins élevé :

$$Y = a_0 + \sum a_i X_i + \sum a_{ij} X_{ij} + \dots + \sum a_{ii} X_i^2 + a_{ij...z} X_i X_j \dots X_z \quad (1.7)$$

Où :

- Y est la réponse ou la grandeur d'intérêt. Elle est mesurée au cours de l'expérimentation et elle est obtenue avec une précision donnée.
- X_i représente le niveau attribué au facteur i par l'expérimentateur pour réaliser un essai. Cette valeur est parfaitement connue. On suppose même que ce niveau est déterminé sans erreur (hypothèse classique de la régression).
- a_0, a_i, a_{ij}, a_{ii} sont les coefficients du modèle mathématique adopté a priori. Ils ne sont pas connus et doivent être calculés à partir des résultats des expériences.

L'intérêt de modéliser la réponse par un polynôme est de pouvoir calculer ensuite toutes les réponses du domaine d'étude sans être obligé de faire les expériences.

Chaque point expérimental permet d'obtenir une valeur de la réponse. Cette réponse est modélisée par un polynôme dont les coefficients sont les inconnues qu'il faut déterminer. A la fin du plan d'expériences, on a un système de n équations (s'il y a n essais) à p inconnues (s'il y a p coefficients dans le modèle choisi à priori). Ce système s'écrit d'une manière simple en notation matricielle :

$$Y = aX + e \quad (1.8)$$

- Y est le vecteur des réponses.
- X est la matrice de calcul, ou matrice du modèle, qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé.
- a est le vecteur des coefficients.
- e est le vecteur des écarts

Ce système possède un nombre d'équations inférieur au nombre d'inconnues. Il y a n équations et $p + n$ inconnues. Pour le résoudre, on utilise une méthode de régression basée sur le critère des moindres carrés. On obtient ainsi les estimations des coefficients que l'on note : \hat{a}

$$\text{Le résultat de ce calcul est : } \hat{a} = (X'X)^{-1} X'Y. \quad (1.9)$$

Formule dans laquelle la matrice X' est la matrice transposée de X . De nombreux logiciels exécutent ce calcul et donnent directement les valeurs des coefficients. C'est ce type de modèles qui est recommandé car il met en considération les mécanismes physiques des procédés mais également les variations des paramètres en dehors des plages habituelles des données (de production) observées.

1.3.3.2.a Choix du modèle

Le choix du modèle diffère selon le phénomène à étudier. Ensuite les limites des plages des paramètres à étudier doivent être définies [FARON 2004]. Il existe plusieurs types de plans d'expériences :

- les Plans à deux niveaux (plans factoriels fractionnaires à deux niveaux, les plans de Koshal, plans de Rechtschaffner, plans de Plackett et Burmann, tables de Taguchi, plans sursaturés).
- les Plans à plusieurs niveaux (plans complets à trois niveaux, carrés latins, carrés de Youden, Plans à niveaux mixtes)
- les plans pour surfaces de réponse (plans composites, plans de Doehlert, plans de Box-Behnken, plans hybrides, plans de Mozzo, plans de D-optimaux)

Il faut noter que d'autres plans d'expériences existent tels que les plans de mélange et les plans booleens.

Dans le cadre de cette thèse, seuls les plans pour surface de réponse seront utilisés pour deux raisons, la première est liée à la spécificité des procédés étudiés et la deuxième raison est que ce modèle est le plus utilisé dans le domaine des semi-conducteurs [ADIVIKOLANU 1998].

1.3.3.2.b Les plans composites

Les plans pour surface de réponse sont appelés également plans du second degré. Ils permettent d'établir des modèles mathématiques du second degré. Ils sont utilisés pour les variables continues. Pour deux facteurs, on a :

$$Y = a_0 + a_1 X_1 + a_2 X_2 + a_{12} X_1 X_2 + a_{11} X_1^2 + a_{22} X_2^2 + e \quad (1.10)$$

Où X_1 et X_2 sont deux variables et $X_1 X_2$ représente l'interaction entre ces deux variables.

Dans ce manuscrit, un seul type de plans de second ordre a été utilisé, il s'agit des plans composites. Un plan composite est constitué de trois parties :

1. Un plan factoriel dont les facteurs prennent deux niveaux.
2. Au moins un point expérimental situé au centre du domaine d'étude.
3. Des points axiaux. Ces points expérimentaux sont situés sur les axes de chacun des facteurs

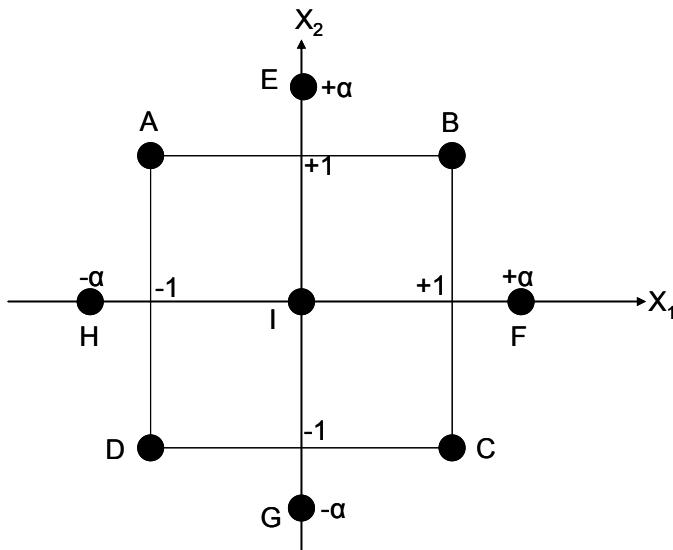


Figure 1.18- Plan composite pour deux facteurs

La figure 1.18 représente un plan composite pour deux facteurs. Les points A, B, C et D sont les points expérimentaux d'un plan 2^2 . Les points E, F, G et H sont les points axiaux. Le point I est le point central. Ce point peut être répliqué une ou plusieurs fois. Si on réalise 9 essais (expériences), 6 coefficients doivent être déterminés. Il faut donc résoudre un système de 9 équations à 6 inconnues.

1.3.3.2.c Analyse statistique

Après l'exécution des plans d'expériences, les résultats des mesures des expériences sont traités. Les résultats des indices statistiques sont déterminants pour juger la qualité du DOE effectué. Le logiciel SAS permet une analyse statistique des résultats obtenus. Cette dernière met en évidence l'erreur du modèle et les éventuelles corrélations que la construction du plan ne permet pas d'expliquer. Pour quantifier cette analyse, deux grandeurs statistiques sont étudiées :

- le coefficient de détermination ajusté (R^2 ajusté: adj- R^2) traduit le taux de variations expliquées par les effets retenus dans le modèle.
- la probabilité que le résultat est significatif. Nous désignons la valeur de la probabilité par (P-value) telle que P-value < une certaine de valeur de seuil calculé par le programme statistique. Sous le logiciel SAS, la valeur de seuil est de 0,05. Lorsque la probabilité pour un résultat est inférieure à 0,05, nous dirons que le résultat est significatif.

1.3.4 Validation du modèle

La dernière étape est la déduction du modèle qui sera valable uniquement dans la plage des variations des facteurs étudiés. Le modèle établi doit être testé de manière physique et statistique :

- statistique : pour voir si le modèle établi est valable lorsque l'on étudie une autre base de données.
- physique : en regardant si le modèle mathématique est conforme à la réalité physique.

1.3.5 Contraintes de production

Le site de ST Rousset est un site de production. Il n'est pas évident de mesurer toutes les plaques. Dans ce cas là, un échantillonnage est donc adopté suivant les opérations et leur degré de criticité.

1.3.5.1 Echantillonnage des mesures

Suivant la criticité des procédés, la lenteur de la mesure et la capacité de charge des outils de mesures, le nombre de lots mesurés diffère d'un procédé à un autre (exemple 1 lot/4 pour les mesures d'oxyde ; 100% des plaques lors de test électrique). Pour les mêmes raisons, l'échantillonnage au niveau des plaques varie d'une étape à une autre. Sur le site de Rousset une cartographie de 9 points est adoptée lors des mesures physiques et une cartographie de 5 points lors des mesures électriques.

1.3.5.2 Retard lié aux étapes de métrologie

Une grande partie des équipements de la FAB ST-Rousset n'est pas équipée de la métrologie intégrée. L'ajustement des paramètres se fait normalement en temps réel quand on travaille avec des boucles de régulation. Ceci malheureusement ne sera pas le cas avec le déploiement des boucles de régulation R2R sur certains procédés. Si les équipements ne sont pas équipés de la métrologie intégrée le temps d'attente impactera négativement le temps de cycle.

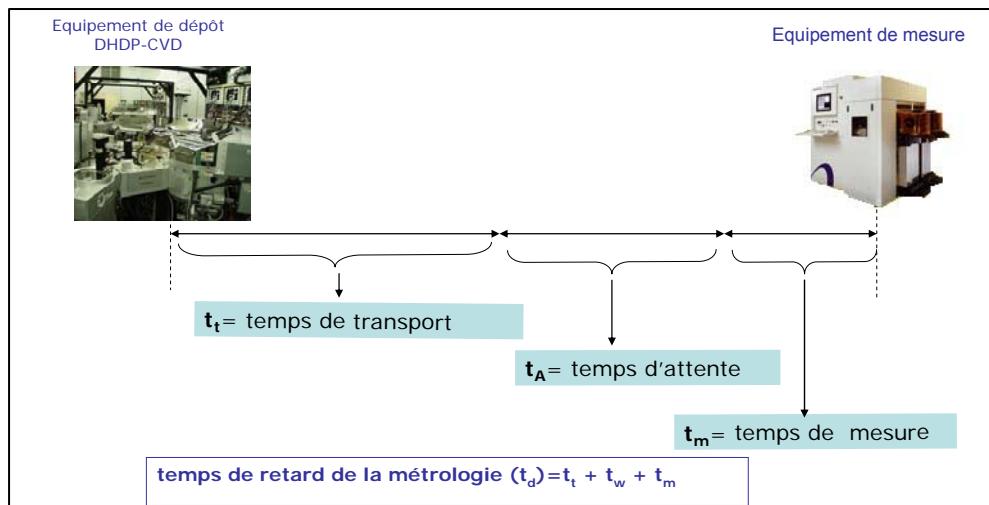


Figure 1.19- Représentation de temps total d'attente.

La figure 1.19 montre un exemple de modélisation du temps total d'attente au niveau métrologie, après le process HDP-CVD. Le lot est envoyé à la station de mesure. On tiendra compte du temps de transport et du temps d'attente au niveau de l'équipement qui dépendent du nombre de lots qui sont en attente pour des mesures ainsi que du temps de mesure qui dépend du nombre de plaques à mesurer [BELHARET 2006b].

1.3.5.3 Effet Produit

Les produits fabriqués sont des produits mémoires et logiques. Les équipements voient donc une variété de produits avec des recettes différentes. Pour une même recette et pour des lots d'une même technologie, les produits sont parfois sensibles à la recette (paramètres d'entrée, limites de contrôle....). Un exemple : dans le procédé polissage mécanico-chimique (CMP), l'«effet produit» est un effet récurrent. En fait, l'épaisseur enlevée durant le procédé CMP dépend de plusieurs paramètres parmi lesquels la densité d'intégration. Cette dernière différencie un produit d'un autre.

1.3.6 Evaluation du gain sur la réduction de la variabilité

Le déploiement des boucles de régulation de type R2R doit apporter sa contribution à l'amélioration du rendement. Par conséquence, cet apport doit être quantifié pour bien voir l'apport de ce type de contrôle et pour éventuellement le généraliser sur d'autres modules de fabrication. Nous utilisons généralement la notion de «capacité des procédés» (ou C_{pk}) pour analyser la variabilité des procédés et des mesures électriques en regard des spécifications associées. Le choix du paramètre électrique ou physique à optimiser doit reposer sur l'analyse de l'indicateur C_{pk} . L'évaluation du gain doit être estimée au niveau :

- des paramètres électriques:** il s'agit généralement d'évaluer le gain de la réduction sur la variabilité des valeurs de certains paramètres électriques. L'estimation du gain peut être réalisée à partir de simulations. Un paramètre électrique (courant, tension, résistance,...) doit être défini comme indicateur de suivi des boucles de régulations. Dans notre cas d'étude, nous présenterons dans le chapitre III, l'indicateur électrique qu'on a défini pour le suivi des boucles de régulation.

- **Du test final des puces:** le test final des puces est considéré comme une référence du rendement électrique. C'est pendant cette dernière étape qu'on connaît le pourcentage des bonnes puces (fonctionnelles). L'apport des boucles sur l'amélioration du rendement final est apprécié pour justifier leur déploiement.

1.4 Conclusion

Dans les paragraphes précédents, nous avons mis en avant la nécessité de contrôler les procédés face aux variabilités rencontrées notamment dans les nouvelles technologies. Le succès du déploiement de boucles de régulation R2R sur ST Microelectronics et dans les compagnies concurrentes a suscité une course contre la montre dans le déploiement de boucles afin de garantir un bon rendement et de réduire le temps de cycle. Les outils statistiques de type multivarié permettent d'identifier les paramètres et les étapes critiques. Sans ces outils, il sera difficile de traiter le grand nombre de données collectées par les différents équipements de procédés et de mesures. La connaissance des mécanismes des procédés (aspects physiques, chimiques...) est un atout pour valider les corrélations constatées entre les différents procédés et valider aussi les résultats des analyses multivariées. Les modèles de procédés sont les tâches difficiles à accomplir. Une partie expérimentale est généralement nécessaire pour accomplir cette tâche.

Le bloc d'isolation des transistors a été identifié comme un module critique sur le rendement notamment dans les technologies mémoires. Nous allons revenir en détail sur ce point dans le chapitre III. C'est donc sur ce module que se sont portés les travaux de cette thèse. Dans les chapitres qui suivent, on exposera en détail comment on a pu appliquer la méthodologie de travail élaborée et décrite dans ce chapitre. La méthodologie établie est le fruit de l'état de l'art de recherches bibliographiques académiques et industrielles mais aussi de la disponibilité des moyens existants. Il faut noter qu'une FAB 200mm comme celle de ST- Rousset est moins automatisée qu'une FAB 300mm.

Dans le chapitre suivant, on va exposer le bloc d'isolation sur lequel on va travailler. Il est très important de comprendre les aspects physiques et technologiques des procédés de fabrication de ce module pour bien interpréter l'origine des variabilités observées.

Références bibliographiques

- [ADIVIKOLANU 2000] S. Adivikolanu, E. Zafriou, “Extensions and performance/robustness tradeoffs of the EWMA run-to-run controller by using the internal model control structure”, *IEEE Transactions on Electronics Packaging Manufacturing*, Vol. 23, p. 56–68, 2000.
- [ALEGRET 2006] Cyril Alegret, “Développement de méthodes génériques de corrélation entre les mesures électriques & physiques des composants et les étapes élémentaires de fabrication”, *thèse de doctorat*, Novembre 2006.
- [ALEGRET 2005] Alegret, C., F. Pasqualini, R. Fernandez , “Machine learning for metrology applications in front end manufacturing: Adaptive surface reconstruction and abnormal profile detection”. *2nd ISMI Symposium on Manufacturing Effectiveness*, Austin, 2005.
- [APC network 2007] www.apc-network.com, 2007
- [BADGWELL 1995] Badgwell, T. A., Breedijk, T., Bushman, S. G., Butler, S. W., Chatterjee, S., Edgar, T. F., Toprac, A. J., & Trachtenberg, I.. “Modeling and control of microelectronics materials processing”. *Computers and Chemical Engineering*, 19(1), pp. 1-41. (1995).
- [BAKSHI 1997] V. Bakshi, “Benchmarking of commercial software for fault detection and classification FDC of plasma etchers for semiconductor manufacturing equipment”, *Proceedings of the American Control Conference*, Albuquerque, New Mexico, pp.1579–1582, 1997.
- [BARAS 1996] J.S. Baras, N.S. Patel, “Designing response surface model-based run-by-run controllers: A worst case approach”, *IEEE Transactions on Components, Packaging, and Manufacturing Technology —Part C* Vol.19, p. 98–104, 1996.
- [BELHARET 2006 a] Djaffar BELHARET, P.Collot, J.Pinaton, “Run to run process model for HDP-CVD”, *7th European conference AEC/APC*, Aix en provence, Avril 2006.
- [BELHARET 2006 b] Djaffar Belharet, Philippe Collot, Pascal Fornara, Jacques Pinaton, Joel Cholvy, Agnès Roussy, “OXIDE HDP-CVD MODELING FOR SHALLOW TRENCH ISOLATION”, *IFAC workshop on advanced process control for Semiconductor Manufacturing*, Singapore, December 2006.
- [BERGERET 2004] F. Bergeret, Y. Chandon, C. Le Gall, “De la statistique dans l’industrie: Un exemple à FREESCALE”, *Journal de la société française de statistique*, Vol.145 (1), pp. 71-95, 2004.
- [BESSE] PHILIPPE BESSE, “Data mining : II. Modélisation Statistique &Apprentissage”, www.lsp.ups-tlse.fr/Besse/
- [BONAVITA 1998] Nunzio Bonavita, Raffaele Tomasi, “Improvements in Process Control through Model-Based a Control System Vendor’s Perspective Techniques”, *Proceedings of the 1998 IEEE International Conference on Control Applications*, Trieste, Italy 1-4 September 1998.
- [BONING 1995] D. Boning, W. Moyne, T. Smith, “Run by run control of chemical-mechanical polishing”, *IEEE/CPMT International Electronics Manufacturing Technology Symposium*, pp. 81–87, 1995.
- [BONING 2004] D.S. Boning, W.P. Moyne, T.H. Smith, J. Moyne, R. Telfeyan, A. Hurwitz, S. Shellman, J. Taylor, “Run by run control of chemical-mechanical polishing”, *IEEE Transactions on Semiconductor Manufacturing* 9 (1996) 307–314. computer aided analysis of complex multiscale systems, in: *AIChE Journal*, pp. 1346–1355, 2004.
- [Box 1993] Box, G. E. P. “Process adjustment and quality control”, *Total Quality Management*, 4(2), pp. 214-227(1993).
- [Cherry 2003] G. Cherry, S.J. Qin, “Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis”, *IEEE Trans. Semi. Manuf., Journal of Chemometrics*, Vol. 17, p.408–502, 2003.

- [DEL CASTILLO 1998] E. Del Castillo, J.Y. Yeh, “An adaptive run-to-run optimizing controller for linear and nonlinear semiconductor processes”, *IEEE Transactions on Semiconductor Manufacturing* 11 (1998) 285–295.
- [DEL CASTILLO 2000] E. Del Castillo, A. Hurwitz, “Run-to-run process control: Literature review and extensions”, *Journal of Quality Technology* 29 (1997) 184–196. 10 (2000) 245–250.
- [EDGAR 1999] T.F. Edgar, W.J. Campbell, C. Bode, “Model based control in microelectronics manufacturing”, *Proceedings of the Conference on Decision and Control*, 1999, vol. 38, pp. 4185–4191.
- [EDGAR 2000] Thomas F. Edgar, Stephanie W. Butler, W. Jarrett Campbell, Carlos Pfeifer, Christopher Bode, Sung Bo Hwang, K. S. Balakrishnan, J. Hahn, “Automatic control in microelectronics manufacturing: Practices, challenges, and possibilities”, *Automatica*, Vol.36 , pp-1567-1603, 2000.
- [FARON 2004] K. Faron, M. Freeland, O. Krogh, S. Patel and G. Raghavendra, “Data analysis and modeling for process control”, *SPIE Proceedings series*, Vol. 5378, pp. 18-27, Ed. Tobin Kenneth W Jr, 2004.
- [Good 2002] R. Good, S.J. Qin, “Stability analysis of double EWMA run-to-runcontrol with metrology delay”, *Proceedings of the American Control Conference*, Anchorage, AK, 2002, pp. 2156–2161.
- [Good 2005] R. Good, S.J. Qin, “On the stability of MIMO EWMA run-to-run controllers with metrology delay”, *IEEE Transactions on Semiconductor Manufacturing*, revised, 2005.
- [HAMBY 1998] E.S. Hamby, P.T. Kabamba, P.P. Khargonekar, “A probabilistic approach to run-to-run control”, *IEEE Transactions on Semiconductor Manufacturing* 11 (4) (1998) 654–669.
- [KRIGER 2005] J.Kriger, “Detecting faults heuristically in time delayed data”, *European AEC/APC conference*, Dublin, 2005
- [LAHAV 2005] Orly Ben-Hamu Lahav, Amir Ben-TAL, “The journey towards making a 100% FDC coverage FAB solutions”, *European AEC/APC conference*, Dublin, 2005.
- [Lee 1995] S.F. Lee, C.J. Spanos, “Prediction of wafer state after plasma processing using realtime tool data”, *IEEE Transactions on Semiconductor Manufacturing* 8 (1995) 252–261.
- [May 1991] G.S. May, J. Huang, C.J. Spanos, “Statistical experimental design in plasma etch modelling”, *IEEE Transactions on Semiconductor Manufacturing* 4 (1991) 83–98.
- [MONTEGOMERY 2001] MONTEGOMERY Douglas C., “Introduction to statistical quality control”, *John Wiley & sons*, ISBN 0-471-31648-2, 2001
- [MOYNE 2001] J. Moyne, E. del Castillo, A.M. Hurwitz, “Run-to-Run Control in Semiconductor Manufacturing”, *CRC Press*, 2001.
- [Mullins 1997] Mullins, J. A., Campbell, W. J., & Stock, A.D. (1997). “An evaluation of model predictive control in run to run processing in semiconductor manufacturing”. In A. Ghanbari & A. Toprac, *Process, equipment, and materials control in integrated circuit manufacturing III. SPIE* , pp.182-189).
- [NERUDA 2005] NERUDA Roman, KUDOVA Petra, “learning methods for radial basis function networks”, *Future generation Computer Systems*, Vol.21, p.1131-1142, 2005
- [ROUSSOS 2005] ROUSSOS George, BAXTER Brad J.C., “Rapid evaluation of radial basis functions”, *Journal of computational and Applied Mathematics*, Vol. 180, p.51_701, August 2005
- [Qin 2006] S. Joe Qin, Gregory Cherry, Richard Good, Jin Wang, Christopher A. Harrison, “Semiconductor manufacturing process control and monitoring: A fab-wide framework”, *Journal of Process Control* Vol.16,pp. 179–191, 2006.
- [ROBINSON 2005] J. Robinson and F. Chance, “The Three Fundamental Drives of Fab Cycle Time,” *FabTime Newsletter*, Volume 6, No.5, 2005.

- [ROBINSON 2007] J. Robinson, “sources of variability in wafer FABs”, *FabTime Newsletter*, Volume 8, No.4, 2007.
- [SACHS 1991] Sachs, E., Guo, R., Ha, S., & Hu, A. (1991a). “ Process control system for VLSI fabrication”. *IEEE Transcations on Semiconductor Manufacturing*, 4, 104-107,1991.
- [SMITH 1991] Lindsay I Smith, *A tutorial on Principal Components Analysis*, February 26, 2002
- [SONDERMAN 2005] Tom Sonderman, “How Does AMD Achieve Truly Automated Control, The Integration of Production Scheduling and APC”, *European AEC/APC conference*, Dublin, 2005.
- [TENENHAUS 1998] M. Tenenhaus, La régression PLS: Théorie et pratique, Edition Technip, Aout 1998.
- [WISE 1991] BARRY M. WISE, “Adapting Multivariate Analysis for Monitoring and Modeling of Dynamic Systems”, *thèse de doctorat*, 1991.
- [WOLD 2004] Svante Wold, Lennart Eriksson, Johan Trygg, Nouna Kettaneh, “The PLS method -- partial least squares projections to latent structures and its applications in industrial RDP (research, development, and production)”, *PLS in industrial RPD - for Prague*
- [YINGHUA 2002] Yinghua Yang, Ningyun Lu, Fuli Wang, Liling Ma, Yuqing Chang, “Statistical Process Monitoring using Multiple PCA Models”, *Proceedings of the American Control Conference Anchorage*, AK May 8-10,2002
- [Yue 2000] H. Yue, S.J. Qin, R. Markle, C. Nauert, M. Gatto, “Fault detection of plasma etchers using optical emission spectra”, *IEEE Transactions on Semiconductor Manufacturing* 13 (2000) 374–385.
- [Yue 2001] H. Yue, S.J. Qin, J. Wiseman, A. Toprac, “Plasma etching endpoint detection using multiple wavelengths for small openarea wafers”, *Journal of Vacuum Science and Technology A* 19 (2001) 66–75.
- [YUE 2004] H. Henry Yue and Masayuki Tomoyasu, “Weighted Principal Component Analysis and its Applications to improve FDC Performance”, *43rd IEEE Conference on Decision and Control*, December 14-17, 2004, Atlantis, Paradise Island, Bahamas
- [ZAFIRIOU 1995] E. Zafriou, H.-W. Chiou, R.A. Adomaitis, “Nonlinear modelbased run-to-run control for rapid thermal processing with unmeasured variable estimation”, *187th Electrochemical Society Meeting*, 1995.

Chapitre 2

Isolation par les tranchées peu profondes (STI) dans la technologie CMOS

Plusieurs études ont démontré la criticité du module d'isolation sur le rendement final électrique des circuits. Dans ce chapitre, nous allons présenter le bloc d'isolation sur lequel on va travailler. Nous allons dresser un état de l'art détaillé sur les techniques d'isolation proposées. Nous verrons que le développement des nouvelles techniques d'isolation ne répond pas seulement à un besoin électrique pour bien isoler les transistors adjacents mais aussi à un besoin dimensionnel pour accentuer la densité d'intégration. A la fin de ce chapitre, nous discuterons en détail de l'impact des tranchées peu profondes sur certains paramètres électriques et sur la génération de contraintes mécaniques.

2.1 La technologie CMOS

2.1.1 La loi de Moore

Les premiers dispositifs électroniques à l'état solide datent de 1948 avec la réalisation du premier transistor bipolaire à base de germanium après la découverte de l'effet transistor par J. Bardeen, W. Shockley et W. Brattain. Depuis, les performances du transistor se sont améliorées avec le passage du germanium au silicium et la multitude des méthodes de fabrication. A partir de 1960, les premiers circuits intégrés (CI) apparaissaient grâce à la technologie « planar ». Cette technologie est caractérisée par le fait que toutes les étapes technologiques se réalisent collectivement pour tout le circuit. De même, toutes les connexions sortent sur un même plan. Les progrès des circuits intégrés ont continué avec la réalisation du transistor Métal/oxyde/semi-conducteur plus facile à intégrer. En 1965, G. Moore prédisait que la complexité des semi-conducteurs doublait tous les dix-huit mois. Cette prédiction est connue sous le nom de la « loi de Moore ». Pour que cette loi soit valable au fil du temps, les technologies de fabrication doivent constamment évoluer pour répondre aux problèmes générés lors de la réduction des dimensions des dispositifs [DEUNEVILLE.2002].

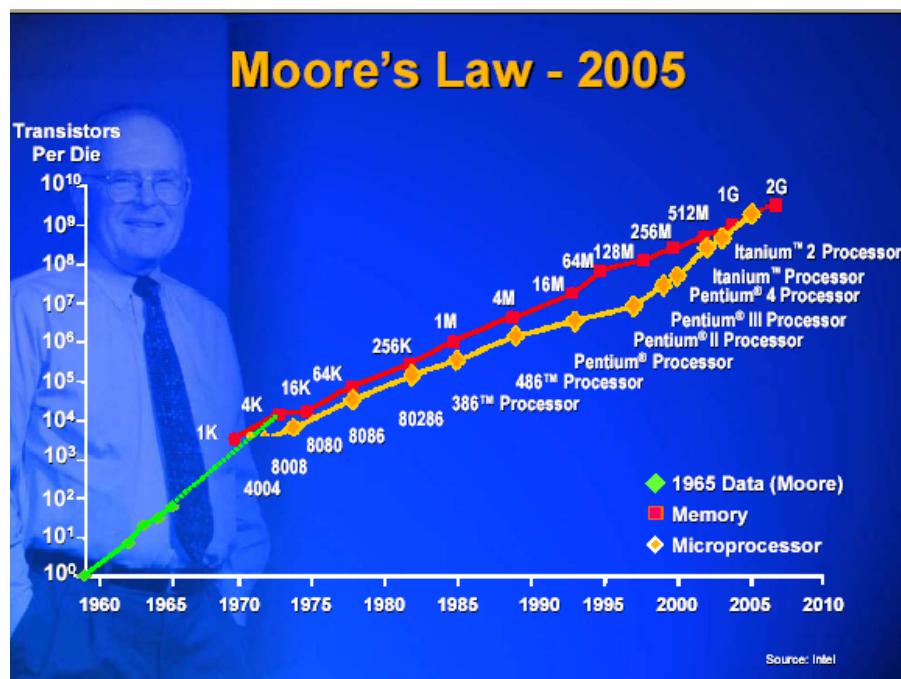


Figure 2.1- Evolution des produits en fonction du temps suivant la loi de Moore (augmentation de la densité d'intégration²⁷ et réduction des dimensions critiques) [www.intel.com]

La figure 2.1 illustre l'évolution des générations de produits fabriqués par INTEL²⁸ avec le temps. Cette évolution est caractérisée par le nombre de transistors implantés sur la plaque. Cependant, on constate un léger décalage dans l'évolution des produits mémoires avec les produits logiques (processeurs).

La complexité augmente d'un facteur deux chaque année. Depuis, le rythme de croissance ralentit un peu mais reste supérieur à un facteur deux tous les deux ans. Le tableau

²⁷ Le nombre de transistors implantés sur une surface d'une plaque de silicium.

²⁸ Entreprise leader de fabrication des semi-conducteurs.

ci-dessous résume les évolutions enregistrées et la vitesse de développement de la microélectronique

Année	1960	1975	1991	1995	2001	2006
Diamètre des plaques (cm)	5	10	15	20	30	30
Longueur de grille (μm)	25	5	0.5	0.35	0.18	0.065
Epaisseur d'oxyde de grille (nm)	500	120	12	7	3.5	1,2
Nombre de niveaux de métal	1	1	2-3	4-5	8	8-10
Métal de grille	Al	Poly Si Dopé++	Poly Si++ + TiSi_2	Poly Si++ + TiSi_2	Poly Si++ + TiSi_2	Poly Si++ + NiSi
Métal interconnexions	Al	Al	Al(Cu)	Al(SiCu)	AlCu	Cu
Techniques de dépôt métal	Evaporation	Evaporation	Pulvérisation cathodique	Pulvérisation cathodique	Pulvérisation cathodique	Electroplatine
Technique de dopage	Diffusion	Implantation ionique	Implantation ionique autoalignée	Implantation ionique autoalignée	Implantation ionique autoalignée	Implantation ionique autoalignée
Lithographie (μm)				0.248	0.248	0,193
Type de transistors	Bipolaire	N-Mos	C-MOS	C-MOS	CMOS-BICMOS	CMOS-BICMOS
Complexité (bits)	8	32K	16M	64M	1G	1-2G

Tableau 2.1- Evolution des technologies en fonction du temps

Au delà de la complexité et de la taille des composants, le tableau 2.1 présente les principales évolutions des technologies. La maîtrise et l'évolution des processus technologiques améliorent la reproductibilité et l'uniformité des éléments. Ils permettent la réduction des dimensions. Du coup, l'utilisation des surfaces des puces devient de plus en plus grande (de 0.1 à 8 cm^2 de 1970 à 1998 pour les mémoires DRAM) sur des tranches de Si de plus en plus larges (d'un diamètre de 5 à 20 cm de 1970 à 1998).

La réduction des dimensions de la grille entraîne le passage rapide de la diffusion à l'implantation ionique, puis à l'implantation ionique auto-alignée. Les relations entre les différentes grandeurs constituent une caractéristique importante des circuits intégrés, favorisant ainsi l'accroissement de la densité d'intégration. Suivant des lois d'échelles, la diminution de la longueur de la grille entraîne la diminution de l'épaisseur de la grille (7nm pour 0.35 μm), celle de la profondeur des jonctions (0.1 μm pour 0.35 μm) et l'augmentation des niveaux de dopage (quelques 10^{17}cm^{-3} pour 0,35 μm). Pour la grille, le passage de l'utilisation de l'aluminium au Si poly-cristallin dégénéré permet d'abaisser les tensions de seuil. La formation d'un siliciure au dessus permet de maintenir une résistance d'accès faible [DEUNEVILLE.2002]. Avec l'augmentation de la densité des composants, les densités de courant dans les lignes d'interconnexions augmentent considérablement. Le « vent » d'électrons pousse les atomes d'aluminium (déposés par évaporation) trop légers (« électromigration »). L'addition de quelques pourcents d'atomes plus lourds Si et /ou Cu résout le problème. L'augmentation de la densité de composants nous oblige à utiliser plusieurs niveaux de connexions métalliques pour permettre les croisements des métaux. Il

faut donc les séparer par des isolants déposés, puis il faut aplanir ces couches pour faciliter l'étape de la lithographie.

Malgré les annonces répétées de leur disparition imminente, les transistors bipolaires subsistent toujours pour des applications particulières (puissance, vitesse). Ils se développent à nouveau maintenant dans les technologies mixtes BICMOS²⁹. Dans cette thèse, nous n'allons nous concentrer que sur la technologie CMOS produite sur le site de ST- Rousset.

2.1.2 Introduction de la Technologie CMOS

La technologie CMOS est une technologie dite « planar », dédiée au développement des systèmes à très haute échelle d'intégration. Grâce aux propriétés électriques des transistors MOS complémentaires, cette technologie permet de réaliser des circuits à faible coût et à basse consommation. Cet avantage a fait de la technologie CMOS la technologie de pointe c'est à dire la plus avancée dans le domaine de la microélectronique. On peut distinguer dans la technologie CMOS 3 grands types :

1. la technologie CMOS à caisson n (utilisation d'un substrat en silicium de type p dans lequel est formé un caisson de silicium de type n).
2. la technologie CMOS à caisson p (utilisation d'un substrat en silicium de type n dans lequel est formé un caisson de silicium de type p). Les technologies NMOS et PMOS présentent une consommation en statique³⁰ non négligeable, due au fait que le transistor de charge n'est pas complètement fermé lorsque le transistor de commande est ouvert.
3. la technologie CMOS à double caisson (utilisation d'un substrat en silicium de type quelconque, sur lequel sont déposés des caissons de type n et de caissons de type p). La technologie CMOS est la technologie la plus répandue parmi toutes les technologies semi-conducteurs. Elle se caractérise par le fait que toutes les fonctions logiques dans cette technologie sont réalisées moyennant l'utilisation d'une paire de transistors MOS complémentaires, c'est-à-dire connectés en série, l'un au canal N et l'autre au canal P. Lorsque l'un conduit, l'autre est fermé, grâce à quoi une porte logique CMOS ne consomme de l'énergie qu'au moment de la commutation. Le circuit le plus simple et le plus connu est l'inverseur, composé de deux transistors NMOS et PMOS.

Grâce au dessin quasi symétrique entre les transistors, la technologie CMOS présente une meilleure adaptation à l'automatisation de la conception des circuits et à la miniaturisation, ce qui donne à la technologie CMOS un autre avantage. De plus, les circuits CMOS présentent, en outre, une meilleure immunité au bruit et au rayonnement ionisant. Les procédés de fabrication utilisés dans chaque technologie CMOS diffèrent d'un fabricant à un autre et évoluent continuellement pour réaliser des transistors plus performants et plus rapides. Nous allons présenter les grandes lignes des opérations élémentaires utilisées pour les technologies CMOS à caisson n. Par analogie, on peut déduire facilement que cette même méthodologie sera employée pour la fabrication de la technologie CMOS à caisson p et pour la technologie CMOS à double caisson.

²⁹ BICMOS (Binary CMOS) est une technologie mixte qui combine la technologie bipolaire et la technologie CMOS.

³⁰ Etat bloqué du transistor

2.1.3 Procédés de fabrication d'un circuit CMOS

La figure 2.2 illustre les principales opérations de fabrication du circuit CMOS à caisson n. La première opération consiste à déposer une couche d'oxyde épaisse sur un substrat de silicium. Cette couche jouera le rôle de masque pour les opérations suivantes. À l'aide de l'opération de lithographie on définit la zone du caisson. L'opération suivante est l'implantation du phosphore. Pendant cette opération, la profondeur de caisson et donc son niveau de dopage va dépendre de l'énergie et de la durée de l'implantation. Le caisson n ainsi réalisé va servir à réaliser des transistors à canal p. Après, la couche d'oxyde épaisse est enlevée complètement et une autre couche fine est créée par oxydation du silicium. On procède ensuite au dépôt d'une couche de poly-silicium, la grille en poly-silicium est formée en gravant le poly-silicium dans les régions indésirables. Pour la formation de la source et du drain de transistor à canal n, on implante du phosphore. Pour le transistor à canal p, la formation de la source et du drain est obtenue par l'implantation du bore. Après le dépôt d'une couche épaisse d'oxyde, une gravure verticale d'oxyde est faite pour créer les vias qui seront par la suite remplis par un métal.

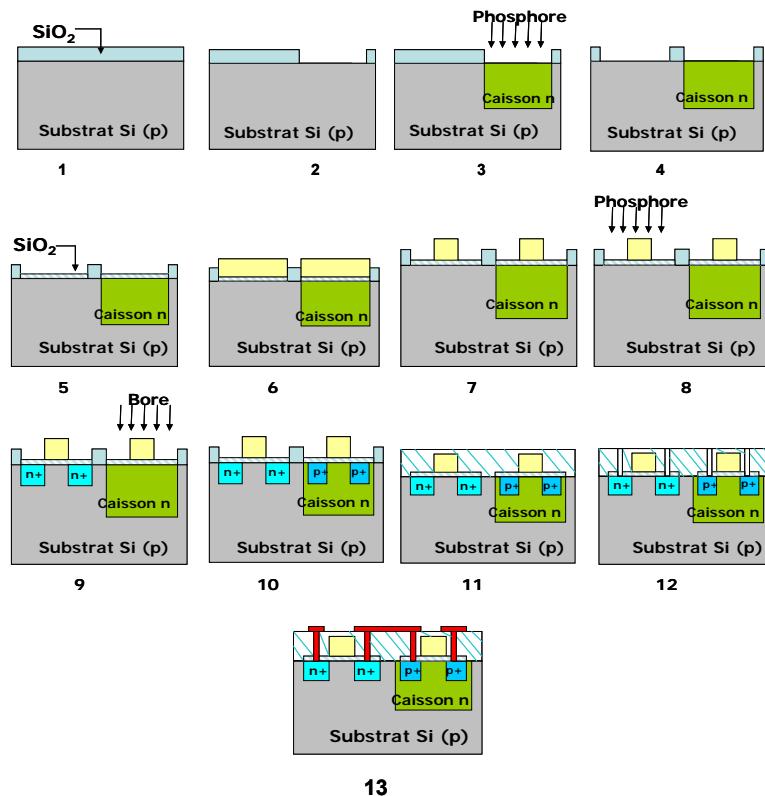


Figure 2.2- Opérations élémentaires dans la technologie CMOS

2.2 Techniques d'isolation pour les circuits intégrés

L'enchaînement des étapes du procédé de fabrication telles que la photolithographie, la gravure, le dépôt, l'implantation ionique, ..., permet de créer, dans le substrat, une mosaïque de zones et de modules. Ces modules peuvent se résumer de la manière suivante : substrat, isolement latéral, caisson, canal, oxyde de grille, grille, jonctions espaces, silicium, interconnexions locales, contacts et métallisations. Ces différents modules sont illustrés sur la figure 2.3.

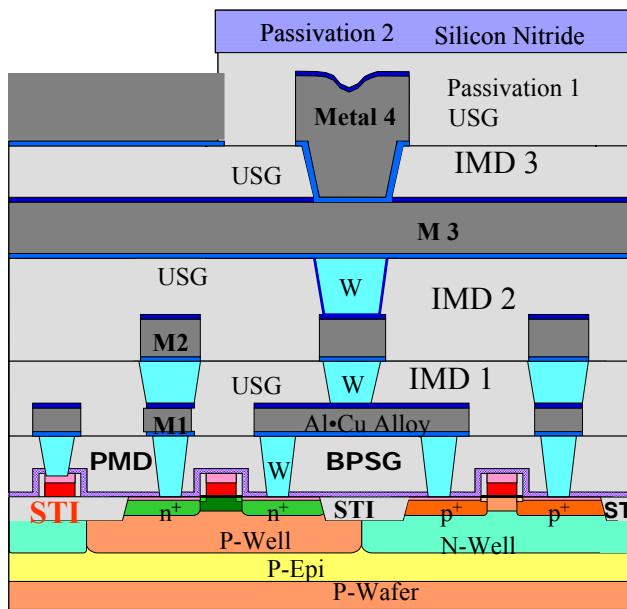


Figure 2.3- Coupe d'un circuit après toutes les étapes de fabrication.

Nous observons sur la figure 2.3 les différents niveaux. Le premier niveau est le niveau d'isolation symbolisé par les tranchées d'isolation peu profondes (STI). Ensuite, le niveau de la grille est formé. Nous distinguons également le niveau PMD (Pre-Metal dielectric), l'oxyde employé généralement est une couche de passivation dopée au bore et phosphore (BPSG : Borophosphosilicate glass). Enfin, le niveau métal. Dans cet exemple, on a 4 niveaux métalliques (M1, M2, M 3 et M4). Pour éviter des problèmes électriques entre les différents métaux, un diélectrique est déposé pour assurer cette fonction, il est appelé généralement IMD (Inter Metal Dielectric). Généralement, le verre de silice non dopé (USG: Undoped Silicate Glass) est souvent le plus utilisé.

Dans la suite de ce manuscrit, nous nous focaliserons sur le bloc d'isolation STI (Shallow Trench Isolation) qui fait l'objet de cette thèse.

2.2.1 Introduction sur l'isolation dans les circuits intégrés

La technologie d'isolation est un des aspects critiques de la fabrication des circuits intégrés. Une grande variété de techniques d'isolation a été développée. L'une des raisons vient du fait que chacune des technologies de circuits intégrés (par exemple NMOS, CMOS, Bipolaire) a ses propres besoins en termes d'isolation. En fait, ces technologies diffèrent suivant les spécificités de chaque technologie, en particulier, l'espace d'isolation, la planarité de la surface, la complexité du procédé et la densité de défauts générés [GOODWIN.1982, TAMAKI 1988] pendant la fabrication dans les structures d'isolation.

Avant l'invention du circuit intégré, seulement les dispositifs discrets comme les diodes, les transistors bipolaires, et les transistors à effet de champ (FET) pouvaient être fabriqués. Au début des années 50, on a observé un nouveau phénomène « le courant de fuite de la jonction » et une faible tension de claquage (résultat de la large densité des piéges à la surface du silicium). En 1958, un groupe de chercheurs du laboratoire « Bell Telephone » a observé que quand une couche de SiO_2 est déposée sur la surface du silicium, et que la jonction pn intercepte la surface, le courant de fuite diminue d'un facteur de 10 à 100. Plus tard, on a compris que l'oxyde de silicium réduit et stabilise les piéges au niveau de l'interface

entre le silicium et l'oxyde. Cet avantage de l'emploi de l'oxyde a été exploité par la suite pour améliorer les caractéristiques électriques des dispositifs. C.T.Sah a fait savoir que le succès du groupe Bell Labs à savoir, la stabilisation de la surface du silicium était l'avancée la plus importante dans le monde de la microélectronique dans les années 1950. Cela a contribué au développement des circuits intégrés. Avec l'arrivée des circuits intégrés, il est nécessaire d'isoler les transistors fabriqués sur une même plaque. Les circuits intégrés bipolaires ont été les premiers CI à être développés. La technique d'isolation consistait à isoler la région du collecteur du reste du dispositif. Cette technique est autrement appelée « isolation de la jonction » comme cela est illustré sur la figure 2.4.

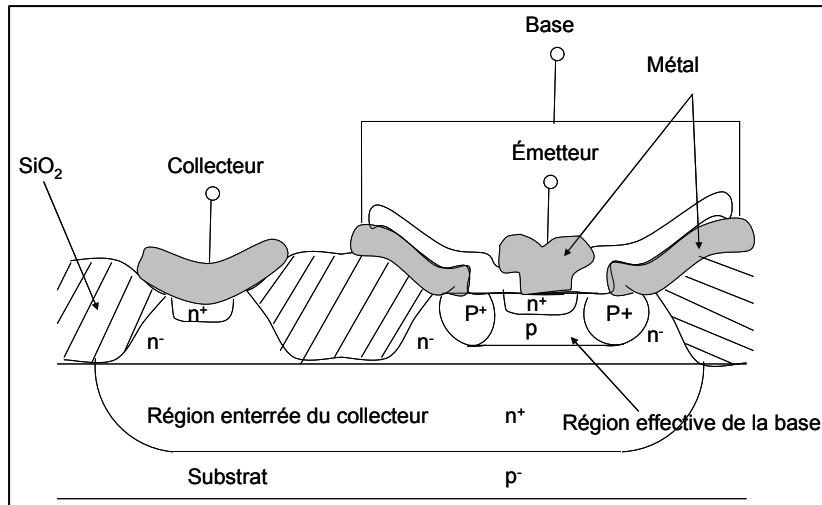


Figure 2.4- Représentation schématique d'un transistor bipolaire

Les CI NMOS et PMOS n'avaient pas besoin de l'isolation de la jonction, mais avaient besoin d'une structure d'isolation qui empêche l'établissement de canaux parasites entre les transistors adjacents. La plus importante technique développée était l'oxydation locale du Silicium ou LOCOS (Local Oxidation of Silicon). Cette isolation pose des problèmes électriques. Le plus grand problème électrique qui se manifeste dans le circuit est le phénomène de déverrouillage (ou latchup). La figure 2.5 illustre une structure CMOS associant des types de transistors différents (NMOS et PMOS). Le phénomène de déverrouillage est lié à l'existence de deux transistors bipolaires montés en série comme nous pouvons le constater sur la figure 2.5. Ce montage peut créer une amplification très importante des courants parasites ($\times 10000$). Les conditions sont donc généralement réunies pour que le thyristor parasite PNPN se déclenche et devienne passant avec une extrême sensibilité, mettant ainsi en court circuit l'alimentation. Cet incident aboutit généralement à la mise hors service de la cellule et parfois du circuit dans son ensemble.

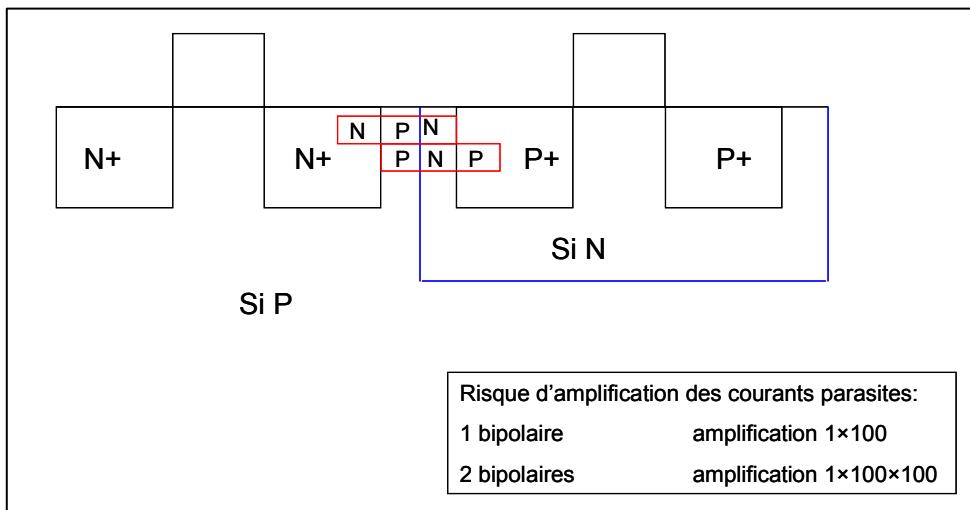


Figure 2.5- Une structure CMOS faisant apparaître l'existence de deux transistors bipolaires montés en série.

Pour éviter les passages de courants parasites, il y a deux possibilités :

- soit d'éloigner les transistors les uns des autres, ce qui limite l'intégration du circuit
- soit de disposer une zone isolante entre les transistors.

Pour maintenir et accroître la densité d'intégration, la deuxième solution est la plus efficace. Dans les prochains paragraphes, nous détaillerons toutes les solutions proposées pour la réalisation de cette zone isolante.

On commencera par rappeler la technologie d'isolation la plus utilisée au cours des deux dernières décennies.

2.2.2 La technologie LOCOS

La technique LOCOS a été utilisée principalement en raison de sa simplicité et de son moindre coût. Différentes méthodes ont été proposées pour la réalisation du LOCOS. Nous pouvons distinguer deux approches basiques.

- Une première approche (utilisée jusque dans les années 1970), consiste à faire croître une couche d'oxyde sur une surface de silicium. Ensuite, cet oxyde est gravé au niveau des zones actives. L'oxyde n'est présent que dans les zones de champ comme cela est illustré sur la figure 2.6.

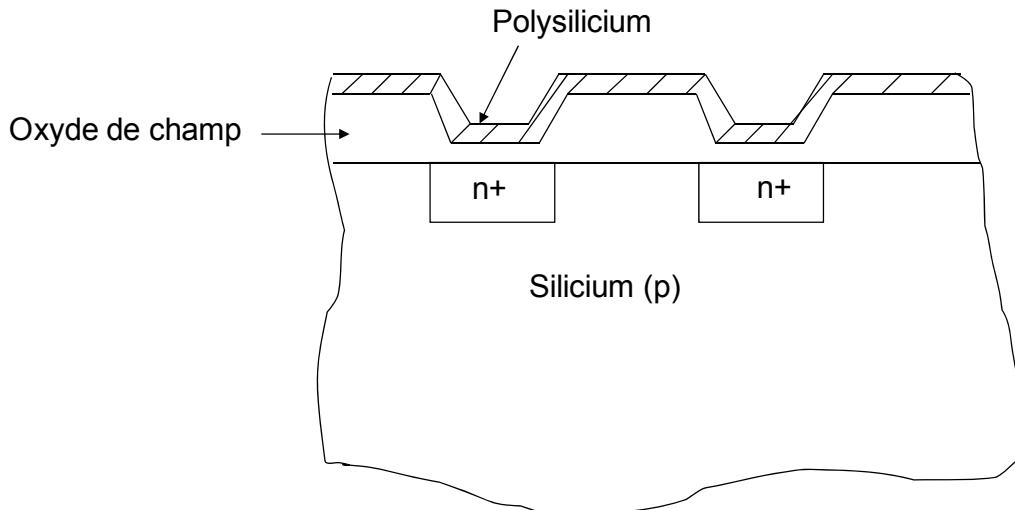


Figure 2.6- Schématisation des premiers modèles d’isolation

Ce procédé contenant deux opérations (croissance d’oxyde et gravure) n’est pas compliqué à faire. Par contre, il présente deux inconvénients qui ont empêché son application dans les composants à haute échelle d’intégration (VLSI) :

- 1) La marche de l’oxyde de champ est haute et les coins supérieurs sont pointus empêchant ainsi toute couverture des parties de circuit concernées par les interconnexions de métal.
 - 2) Le canal d’implant d’isolation (the channel stop implant) qui a pour rôle d’augmenter la tension de seuil des transistors³¹. Le canal d’implant d’isolation doit être achevé avant la croissance d’oxyde. Les zones actives doivent donc être alignées avec les zones implant d’isolation à l’aide d’une opération de lithographie. Le fait de rajouter une autre opération de procédé (dans ce cas la lithographie) constitue une contrainte de plus pour l’élaboration des circuits.
- La deuxième approche est l’oxydation sélective introduite par Appels&Kooi en 1970. Elle consiste à faire croître l’oxyde uniquement dans les régions de champs désirées. Le nitrate de silicium est déposé avant sur les zones actives pour empêcher toute oxydation de ces zones. A la fin de l’étape d’oxydation, le nitrate est enlevé. Dans ce cas de figure, l’implantation est sélective et les régions d’implant d’isolation sont auto - alignées avec l’oxyde de champ comme nous pouvons le constater sur la figure 2.7.

³¹ Deux méthodes sont généralement employées : soit l’augmentation de l’épaisseur d’oxyde de champ, soit l’augmentation de la densité de dopage au dessous de l’oxyde de champ). La combinaison des deux méthodes consiste à faire croître un oxyde de champ et à effectuer une implantation à travers cet oxyde pour augmenter le dopage. Le canal obtenu est appelé « canal d’implant d’isolations ».

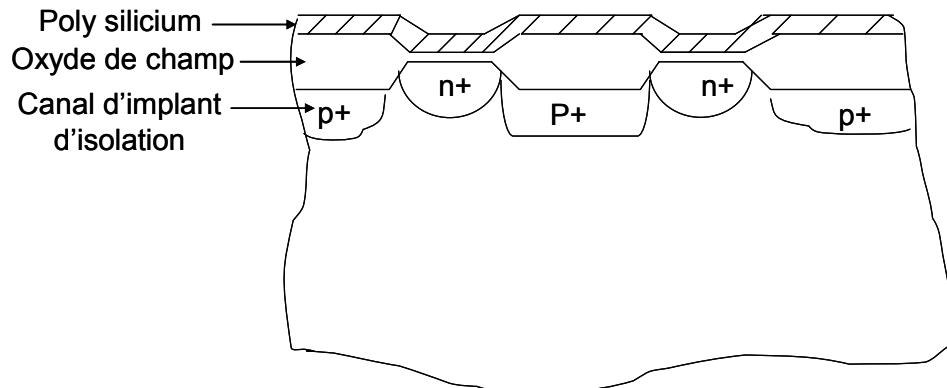


Figure 2.7- Schématisation de l'approche de l'oxydation sélective

La configuration illustrée sur la figure 2.7 est connue aussi sous le nom de LOCOS partiellement enfoncé (semi recessed LOCOS)[BASSOUS.1976]. Il faut noter que l'autre mode de LOCOS dit LOCOS complètement enfoncé (fully recessed LOCOS) [TSAI 1988] est utilisé principalement pour la fabrication des circuits intégrés bipolaires. Nous allons nous focaliser uniquement sur le processus de la réalisation du LOCOS partiellement enfoncé. Les différentes étapes de la formation du LOCOS sont illustrées sur la figure 2.8.

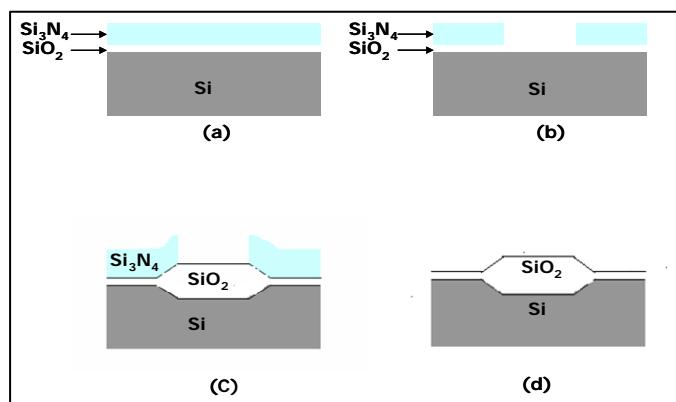


Figure 2.8- Différentes étapes de fabrication pour la réalisation du LOCOS : a) Dépôt d'oxyde piédestal (padox) et du nitrate b) Après la gravure c) Oxydation d) Retrait du nitrate de silicium.

Premièrement, une couche mince d'oxyde piédestal (padox) est déposée thermiquement sur la surface du silicium. Le rôle de cette couche est de diminuer les contraintes mécaniques entre le substrat de silicium [HUI 1982] et la couche de nitrate qui sera déposée plus tard (figure 2.8a). L'épaisseur de cette couche est de 20 à 60 nm. Ensuite, une couche de nitrate de silicium épaisse (Si_3N_4) est déposée par dépôt chimique en phase vapeur (CVD). Cette couche joue le rôle de masque d'oxyde. Cette couche est de 100-200 nm. Elle agit comme une barrière de diffusion de l'oxygène et empêche l'oxydation du silicium au-dessous du nitrate. La diffusion latérale de l'oxygène sous les bords nitrate cause au silicium une oxydation sous le nitrate. Pendant l'oxydation, seulement quelques dizaines de nm sont convertis en SiO_2 . L'un des critères du choix de l'épaisseur de nitrate à déposer est qu'elle soit supérieure à l'épaisseur convertie en SiO_2 . Après les étapes de photolithographie et l'opération de gravure, les zones d'isolation sont définies (figure 2.8b). Elles seront par la suite remplies d'oxyde. L'opération suivante est l'implantation à travers l'oxyde réalisé. Généralement, pour les NMOS, l'implant utilisé est le bore et dans le cas des PMOS, il s'agit

de l'arsenic. Après avoir défini la région de l'implant d'isolation, l'oxyde de champ est obtenu par le procédé d'oxydation pendant 2 à 4 heures sous une température de 1000°C. Pendant la croissance de l'oxyde, l'oxygène diffuse latéralement comme cela est illustré sur la figure 2.8c. Finalement, la dernière étape consiste au retrait du nitrule et de l'oxyde piédestal.

2.2.2.1 Inconvénients du LOCOS

2.2.2.1.a Bec d'oiseau

La figure 2.8c montre qu'on obtient une structure caractéristique. Cette structure est appelée «bec d'oiseau» [SHANKOFF 1980]. Le phénomène du «bec d'oiseau» provient de l'isotropie de l'oxydation du silicium. La diffusion de l'oxygène dans le silicium est un phénomène isotrope. Il y a donc oxydation du silicium sous le masque de nitrule. Cette diffusion est limitée par la contrainte mécanique induite par la couche de nitrule mais elle provoque néanmoins une extension du bouchon isolant au-delà du dessin initial. Cette diffusion latérale est proportionnelle à l'épaisseur d'oxyde formée. Elle est donc une limitation à l'augmentation de la densité d'intégration. La schématisation complète de ce phénomène est illustrée sur la figure 2.9.

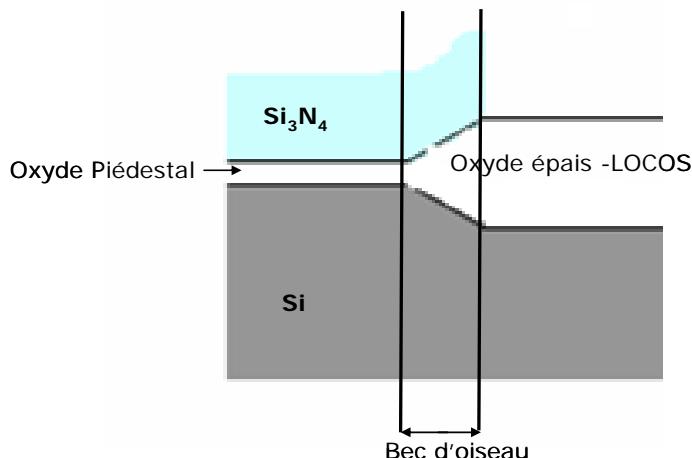


Figure 2.9- Schématisation de phénomène de bec d'oiseau

La formation du bec d'oiseau reste l'un des principaux inconvénients de la technologie LOCOS. Les inconvénients principaux du LOCOS sont liés essentiellement à l'apparition des problèmes d'uniformité pour la suite des opérations de fabrication à cause de l'élargissement de la zone d'isolement. Cet effet³² est dû à la diffusion latérale du bore vers la surface. [MIZUNO 1987]. En plus du problème lié à la diffusion latérale, un autre problème se pose lors des opérations qui suivent en particulier, au niveau du contact métal source et drain. Lors de l'étape de l'ouverture des vias, une sur-gravure est nécessaire pour atteindre les zones actives. Si, à la suite d'une sur-gravure, on atteint le substrat, la source sera court-circuitée par rapport au substrat causant ainsi le dysfonctionnement du circuit. Enfin, il faut dire que le LOCOS n'est pas adapté pour les composants submicroniques car l'épaisseur d'oxyde de

³² Ce phénomène apparaît dans les circuits NMOS où le bore est employé comme l'implant d'isolation. Quand la longueur du bec d'oiseau augmente, le bore a tendance à diffuser latéralement.

champ requise pour les technologies submicroniques est moindre que celle dans le cas des technologies avec un large espace.

2.2.2.1.b Ruban blanc

Un autre phénomène est observé par Kooi & al [KOOI 1976]. Il s'agit de la formation d'une couche de nitrate de silicium au dessous du LOCOS comme cela est illustré sur la figure 2.10. Ce phénomène est appelé « ruban blanc ». Pour comprendre ce phénomène, le modèle de KOOI suppose que durant la croissance du LOCOS, l'eau diffuse à travers l'oxyde. La réaction entre le H_2O et le masque de nitrate résulte dans la formation de l'ammoniaque (NH_3). Ce dernier peut diffuser latéralement et peut atteindre l'interface Si/SiO₂. Ainsi, une couche de Si₃N₄ est formée suite à la réaction de l'ammoniaque et de silicium.

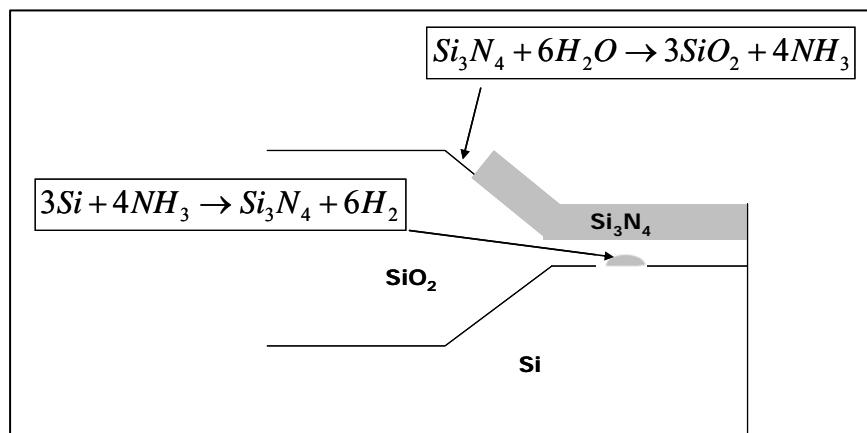


Figure 2.10- Représentation du modèle de Kooi.

La formation de cette couche a des conséquences néfastes par la suite pour les étapes qui suivent et en particulier pendant l'étape de la formation d'oxyde de grille. La vitesse de croissance d'oxyde sera ralentie dans les endroits où la couche de nitrate est formée et par conséquent, l'épaisseur d'oxyde de grille formée dans cette zone est plus mince que dans les autres régions. Cet effet va causer par la suite une basse tension de claquage. Une réduction du bec d'oiseau est cependant possible si l'on dépose un masque de nitrate plus épais, mais cela conduit à l'apparition de contraintes mécaniques très fortes dans le silicium au bord de l'isolation.

2.2.3 LOCOS modifié ou amélioré

Nous venons de voir que le principal inconvénient du LOCOS est celui de la perte dimensionnelle créée par la croissance du bec d'oiseau (la largeur de la zone isolante est supérieur à celle définie par l'ouverture dans la couche de nitrate de silicium. Il y a donc une baisse de la densité d'intégration des circuits). Pour limiter cet inconvénient, il faut limiter la croissance du bec d'oiseau et donc augmenter le rapport $\frac{\text{épaisseur de nitrate de silicium}}{\text{épaisseur d'oxyde piédestal}}$

Deux approches sont donc possibles : soit de réduire l'épaisseur de l'oxyde piédestal, soit d'augmenter l'épaisseur de nitrate. Ce sont ces deux types de constructions que nous allons décrire.

2.2.3.1 La structure SILO ou oxydation locale scellée SILO (Sealed-interface Local Oxidation)

La structure SILO est dérivée de la structure LOCOS. Le SILO réduit l'extension du « bec d'oiseau » par réduction de l'épaisseur de l'oxyde piédestal (limitation de la diffusion de l'oxygène à l'interface oxyde/silicium). La clef de la structure SILO est le scellement d'une couche de nitre à l'interface avec le silicium. Ce scellement est obtenu par nitruration rapide de silicium [HUI 1982, DEROUX-DAUPHIN 1985]. La couche de nitre peut être formée en utilisant un autre procédé à savoir l'implantation ionique (implantation de l'ammoniaque). D'autres structures SILO améliorées ont été proposées :

2.2.3.1.a Oxydation locale scellée latéralement (Laterally Sealed LOCOS)

Ghezzo & al ont développé un autre procédé [GHEZZO 1987] permettant de sceller l'interface durant la croissance de l'oxyde de champ. Le principe de ce procédé est de garder l'avantage du SILO à savoir la réduction du bec d'oiseau et l'amélioration des défauts engendrés par le contact du nitre avec le silicium. Le principe de cette méthode est l'emploi d'un masque « empilement actif » (active area stack). Ce masque est composé d'une première couche de padox 20nm, d'une couche de nitre 80 nm par CVD et d'une couche d'oxyde 200 nm réalisée par le procédé CVD assisté par plasma. L'ajout de cette couche épaisse d'oxyde, elle permet d'apporter un appui vertical pour la deuxième couche de nitre qui sera déposée et qui aura pour rôle de sceller les flans du masque. Une autre couche de 80 nm de nitre est déposée (par CVD) et finalement, une dernière couche d'oxyde de 220nm est déposée par CVD assisté par plasma (PECVD³³). Une gravure sèche anisotrope permet d'enlever les deux derniers films sur les régions actives. Le premier film d'oxyde PECVD est également enlevé. On obtient à la fin la forme d'un espaceur de nitre qui scellera les flans du masque. Finalement, on peut faire croître l'oxyde de champ. La schématisation complète de la méthode d'oxydation locale scellée latéralement est illustrée sur la figure 2.11.

³³ PECVD (plasma enhanced CVD) est une technique CVD assisté par plasma.

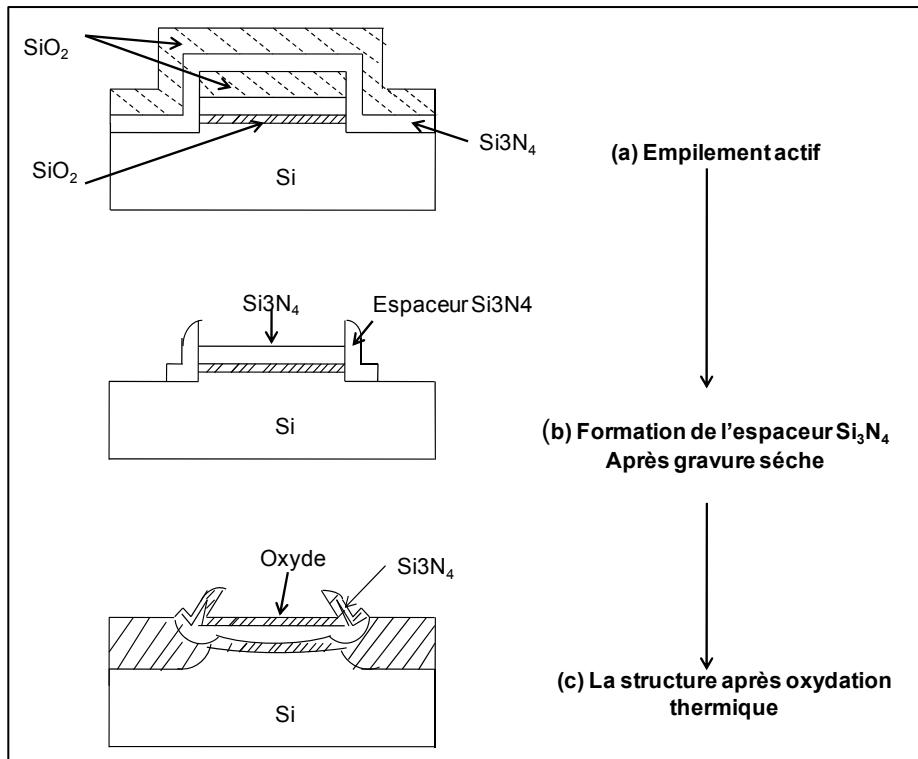


Figure 2.11- Schématisation de la méthode d'oxydation locale scellée latéralement

2.2.3.1.b Le procédé POP SILO

Le procédé décrit précédemment a apporté une bonne protection au niveau du scellage des flancs. Néanmoins, de grandes variabilités ont été observées pour mieux optimiser ce concept du procédé. Une autre approche a été proposée, appelée « POP-SILO ». Ce procédé est illustré sur la figure 2.12. Le principe de ce procédé est semblable au précédent. Il consiste en la croissance d'une couche d'oxyde de 12,5 nm suivie d'un dépôt (par CVD) d'une couche de nitrate de 80 nm. Finalement une couche d'oxyde de 150nm est déposée par LPCVD³⁴. Une autre couche de 250 nm de nitrate est déposée par CVD. Après l'opération de gravure pour définir la zone de croissance, on fait croître l'oxyde de champ.

³⁴ LPCVD (Low pressure CVD) est la technique CVD opérant à basse pression.

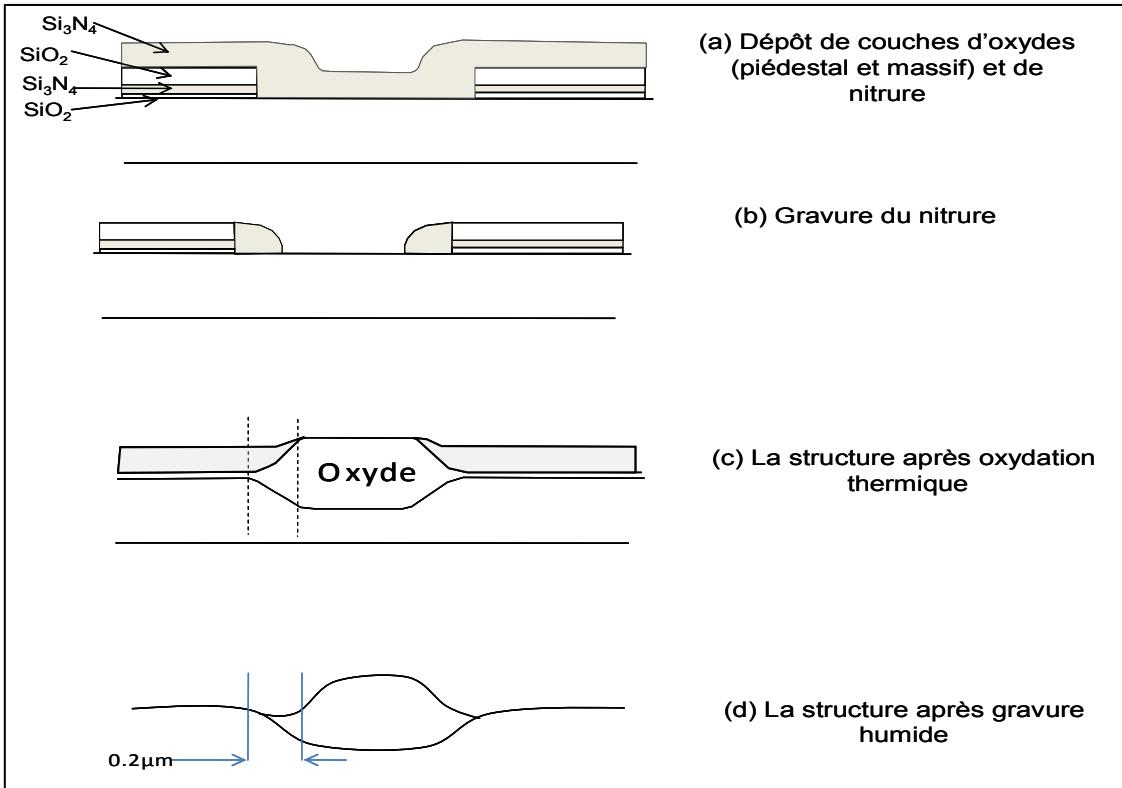


Figure 2.12- Représentation du procédé POP SILO

2.2.3.2 Le procédé PBL (Polybuffered LOCOS)

Ce procédé est légèrement différent du procédé LOCOS standard car l'oxyde piédestal (Padox) est remplacé par une couche dite « couche Pad-polybuffred » constituée de (polysilicium d'une épaisseur de 50 nm et d'oxyde d'épaisseur de 5 à 10 nm). Une couche épaisse de nitre (100 à 240 nm) est ensuite déposée [CHAPMAN 1987, HOSHI 1986]. L'isolement PBL permet de réduire les contraintes et d'adoucir les contraintes mécaniques induites par le nitre [JUANG 1999] grâce à la couche de silicium poly-cristallin qui sépare le nitre de silicium. L'utilisation de masques de nitre épais provoque une réduction considérable du bec d'oiseau (de 0,6 μm environ pour le LOCOS à 0,3 μm pour le PBL). Le succès de ce procédé dans les technologies 0,5μm a été rapporté par Nishihara [NISHIHARA 1988]

2.2.4 Technologies NON LOCOS - Tranchées d'isolation

Bien que les structures SILO et PBL aient amélioré le LOCOS, le bec d'oiseau, même s'il est réduit dans le meilleur des cas à 0,2μm ne permet pas la réalisation de zones actives (ZA) inférieures à 0,4μm, ce qui est incompatible avec les générations technologiques au-delà du 0,25μm [WOLF 1990]. Du point de vue électrique, l'isolement du type LOCOS ou PBL conduit à l'effet de canal court³⁵ qui augmente la tension de seuil V_{th} des transistors étroits (en

³⁵ Effet de canal court (NCE : Narrow Channel Effect), si on réduit les dimensions horizontales sans réduire les dimensions verticales, les lignes de potentiel ne sont pas parallèles à la grille c'est l'effet « canal court ». Ce dernier se traduit par un mauvais contrôle de la tension de seuil du transistor, donc un mauvais contrôle de ces caractéristiques.

fonction du type d'isolement, l'effet de canal court peut conduire aussi bien à une augmentation qu'à une diminution de la tension de seuil quand la largeur du transistor varie).

Pour éliminer tous les inconvénients du LOCOS et de ses structures dérivées, une autre approche d'isolation des dispositifs dans les technologies CMOS et BICMOS a été proposée. Il s'agit d'isolement par tranchées peu profondes (STI : Shallow Trench Isolation) [DAVARI 1988, CHANG 1999, LIEN 2002, FORSBERG 2004]. Son grand avantage est de ne pas présenter de bec d'oiseau. L'oxyde d'isolement étant déposé dans des tranchées précédemment gravées, on n'a plus besoin de l'oxydation localisée qui était la source du bec d'oiseau. Le STI permet également de supprimer l'effet de Latchup et l'effet de canal court inverse [NIITSU.1985, HORITA 2000]. Un autre avantage du STI est qu'il supprime les différences de niveaux entre les ZA et l'isolement. Il n'y a donc plus (ou beaucoup moins) de réflexion de lumière en bord d'isolement, ce qui permet de résoudre des motifs plus petits par photolithographie. Mais avant de détailler ce concept novateur, on va rappeler les premières techniques d'isolation par tranchées.

2.2.4.1 L'isolation par oxyde enterré (BOX: Buried Oxide)

Après avoir défini les tranchées, elles seront remplies par l'oxyde. Le principe de cette méthode [MIKOSHIBA.1984] repose sur la création de tranchées de 0,5 à 0,8 μ m. Ces tranchées sont définies, ensuite elles seront remplies par l'oxyde réalisé par CVD. Cet oxyde sera ensuite aplani par les procédés dépôt/pulvérisation comme cela est montré sur la figure 2.13.

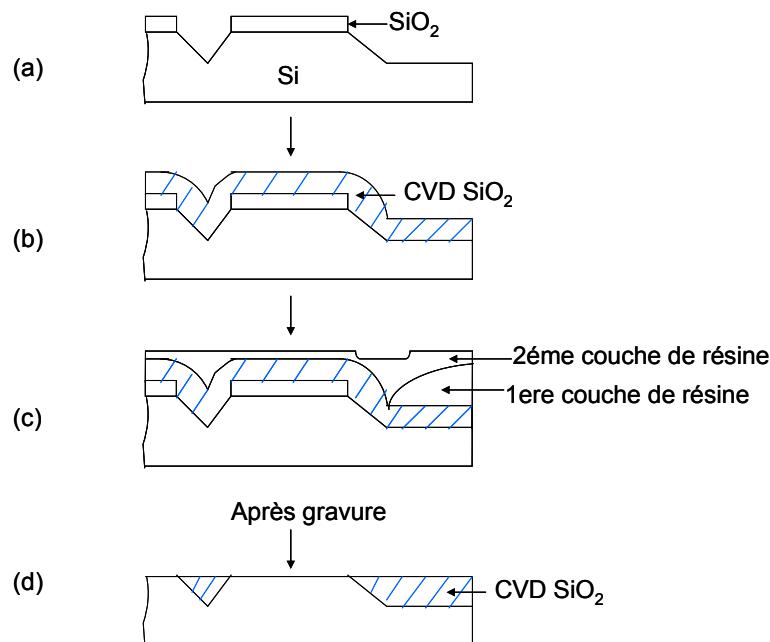


Figure 2.13- Schématisation de la méthode BOX

Comme cela est illustré sur la figure 2.13, cette méthode d'isolation permet d'éliminer « le bec d'oiseau » et permet en même temps d'avoir une surface plane. Malgré ces avantages, cette méthode souffre de quelques inconvénients :

- inversion du dopage au niveau des flancs des régions actives de type P lors de l'implantation du bore.
- création des trous dans les tranchées d'oxyde [FUSE.1987].

- le profil d’oxyde de champ diffère d’une zone à une autre (résultat de la différence de la vitesse de gravure d’une zone à une autre)

Pour remédier aux inconvénients cités auparavant, des actions ont été menées. Sur le premier point, l’élimination de l’inversion des régions actives de type p peut être réalisée en changeant l’angle d’implantation du bore (7° par rapport aux flancs). Une autre solution consiste à déposer par CVD une couche de verre boro-silicate (BSG) autour de la tranchée. Cette couche jouera le rôle de source de diffusion de Bore [WOLF 1990].

2.2.4.2 L’isolation par oxyde enterré avec un procédé de fin de gravure BOXES (Buried Oxide with etch stop process)

Pour empêcher l’exposition des flancs des régions actives, un autre procédé a été proposé. Il s’agit du BOXES (Burried oxide with etch stop process). Le principe de ce procédé est l’emploi d’une couche Molybdène/nitrure. Une couche d’oxyde sera ensuite déposée par CVD. Il faut noter que les conditions du procédé CVD choisies ne vont pas oxyder le Mo. Après les opérations de gravure pour planariser et les opérations de retrait du molybdène et du nitrure, on obtient la structure illustrée sur la figure 2.14.

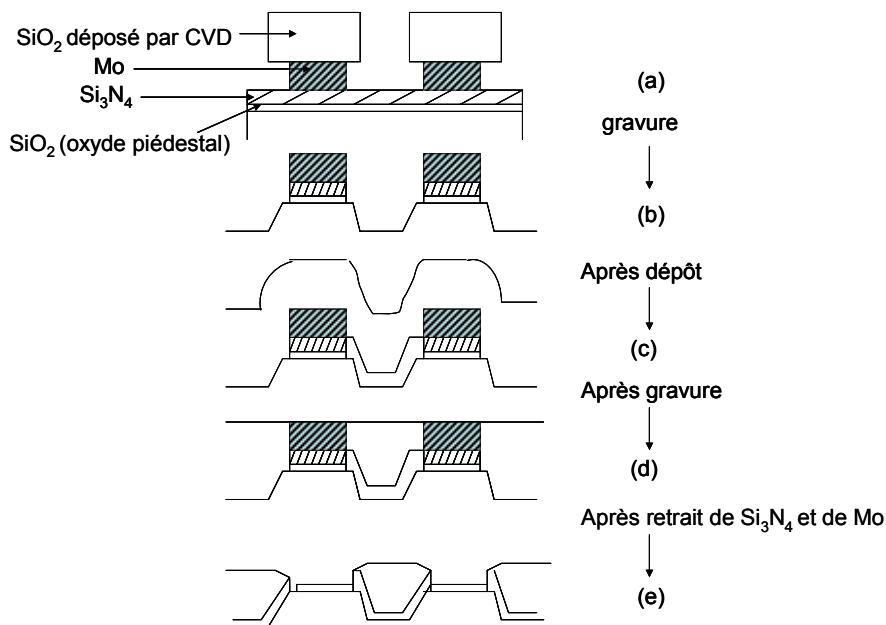


Figure 2.14- Schématisation des étapes élémentaires pour la réalisation du BOXES

Ce procédé amélioré a été efficace pour la technologie MOS (NMOS et PMOS) surtout pour isoler des transistors de même type. Par contre, il n'est pas capable d'isoler efficacement des transistors de types différents.

2.2.4.3 Le STI avec des tranchées peu profondes

Les dernières approches de réalisation du STI répondaient en partie aux besoins manifestés des technologies à savoir : la suppression de l’effet du bec d’oiseau, l’élimination du latchup et la planarité de la surface. Sur ce dernier point, le procédé utilisé est le dépôt/pulvérisation. Il n'est plus satisfaisant avec la réduction des dimensions. Un nouveau

procédé a vu le jour c'est le polissage mécanico-chimique (CMP : Chemical Mechanical Polishing). C'est un procédé qui est basé sur une combinaison de deux effets chimique et mécanique. Ce procédé qui offre une meilleure planarité de surface, a été inventé par IBM en 1983. Le CMP a été exploré par IBM au sein de sa division R&D. En 1989, IBM publie le premier papier sur ce sujet. Ce n'est qu'à partir de cette date que les avantages de ce procédé furent étudiés par les compagnies et les équipementiers. Les différentes applications du CMP sont le polissage d'oxyde et de métal. Nous reviendrons sur les mécanismes de ce procédé dans le chapitre IV.

Nous allons maintenant nous intéresser sur l'ensemble des opérations de fabrication de la brique STI. La figure 2.15 résume l'enchaînement de ces opérations. Premièrement on fait croître une fine couche d'oxyde à la surface du silicium. Une couche de nitride est déposée par CVD (Chemical Vapor Deposition) au dessus de cet oxyde. Par le jeu du couple photolithographie/gravure sèche nous gravons de manière anisotropique le substrat de silicium afin de définir les zones actives et les zones d'isolement du transistor. Nous pouvons alors remplir la tranchée avec un procédé de CVD plus rapide que la croissance d'oxyde thermique. Le surplus d'oxyde est ensuite planarisé jusqu'à la surface de nitride par le procédé CMP. Une épaisseur d'oxyde est gravée par l'acide HF et la barrière de nitride est gravée chimiquement de manière sélective par rapport à l'oxyde de la tranchée. Nous obtenons ainsi une zone active en silicium sur laquelle nous pourrons implémenter des transistors qui seront isolés les uns des autres par les tranchées d'oxyde [NONDAKUMAR.1998].

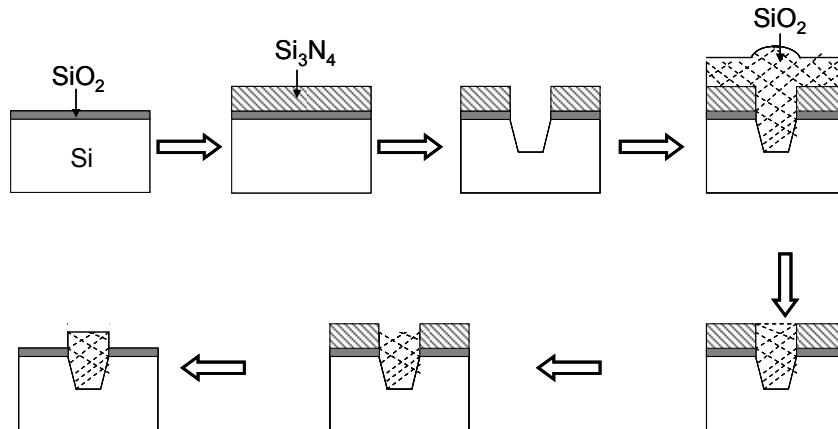


Figure 2.15- Représentation de la fabrication du STI

Les paragraphes qui suivent donnent un aperçu général des opérations qui seront décrites dans les chapitres suivants plus en détails :

2.2.4.3.a Remplissage de la tranchée

La première étape consiste à faire croître un oxyde thermique. Ce type d'oxyde nous permet de réduire au maximum le nombre de liaisons pendantes ainsi que la densité des dislocations. Le remplissage proprement dit de la tranchée se fait en deux temps. Un premier dépôt se fait par le procédé CVD assisté par une haute densité de plasma (HDP-CVD : High Density Plasma CVD) [NAG 1996]. Ce procédé permet de remplir la tranchée sans laisser de « trou » dans l'oxyde. Ensuite une couche d'oxyde est obtenue à partir de Tétrá-éthyloxysilane (TEOS : Tetra Ethyl Orto Silicate). Cet oxyde nous permet d'atteindre l'épaisseur escomptée. Une étape de densification par recuit thermique est nécessaire pour donner à l'ensemble des dépôts la même densité.

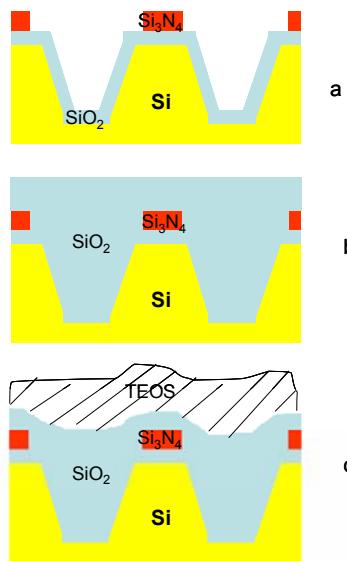


Figure 2.16- Procédé du dépôt dans la tranchée : a) Oxydation thermique des bords de tranchée
b) Remplissage par HDP-CVD c) Remplissage par dépôt TEOS

2.2.4.3.b Le polissage mécanico-chimique

Le CMP a tendance à reproduire des formes peu anguleuses au cours du polissage. L'uniformité est améliorée par gravure de l'oxyde avec un masque inversé par rapport à la gravure active. Cette étape est très importante pour la qualité du CMP et implique une importante épaisseur d'oxyde au dessus du silicium [GRAN.2000, PAN 1998]. Une fois l'oxyde gravé nous avons une surface beaucoup plus uniforme qui va être polie plus uniformément par le CMP.

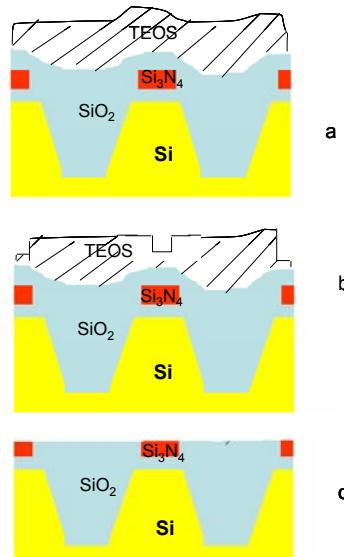


Figure 2.17- Procédés de polissage mécanico-chimique : a) Etat Initial b) Gravure de l'oxyde avec un motif inversé par rapport à la gravure active c) Polissage d'oxyde jusqu'au nitre

2.2.4.3.c La gravure humide

L’opération de la gravure humide (gravure WET) est une gravure chimique très sélective. Cette opération nous permet en effet dans un premier temps de désoxyder la surface (enlèvement d’une certaine épaisseur d’oxyde) et ensuite de retirer la totalité du film de nitrate. L’ensemble des opérations élémentaires de la gravure humide sont illustrées sur la figure 2.18. Nous profitons de cette désoxydation pour ajuster la hauteur d’oxyde du STI. La différence de hauteur entre le silicium et le STI est appelée hauteur de marche comme cela est illustré sur la figure 2.19. Le contrôle de la hauteur de marche est très important pour le fonctionnement électrique des transistors comme nous allons le montrer dans le chapitre III. A la fin de cette opération persiste une mince couche d’oxyde sur l’active qui sera enlevée plus tard (oxyde sacrificiel pour l’opération d’implantation).

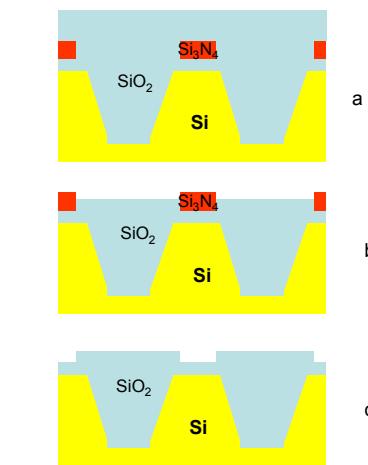


Figure 2.18- Etapes du procédé de la gravure humide : a) STI après CMP b) Désoxydation de la surface de nitrate et ajustement de la marche c) Retrait Nitrate

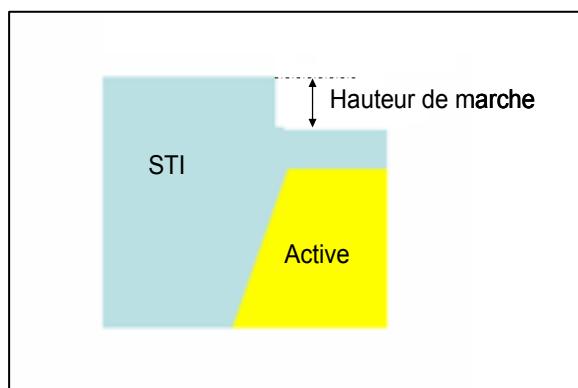


Figure 2.19- Schématisation de la hauteur de marche

2.2.4.4 Le STI amélioré

L’ensemble des opérations de fabrication utilisées pour la réalisation de la brique STI contribue à la variabilité de la hauteur de marche ainsi qu’à l’uniformité. Ces variabilités peuvent causer par la suite des problèmes électriques (des courants de fuite par exemple). Des modifications légères doivent être faites sur certains procédés pour réduire ces variabilités. Nous pouvons citer deux modes de STI amélioré.

2.2.4.4.a Le procédé de retrait SiN « SiN pullback »

Dans le but de réduire les contraintes mécaniques induites par le nitre sur les bords du STI, une autre solution a été proposée. Il s'agit du procédé de retrait SiN pullback, on parle alors du STI pullback (PB-STI). La route des opérations de fabrication du STI amélioré est la même que celle pour le STI standard, sauf qu'il y a une opération de plus. La figure 2.20 résume l'enchaînement de ces opérations. Après avoir réalisé la couche d'oxyde piédestal, une couche de poly-silicium est déposée [Li 2002] (épaisseur de 0 à 50nm). Ensuite, c'est l'enchaînement des mêmes opérations que le STI standard. Notons que cette couche est complètement enlevée lors de la gravure humide comme pour la couche de nitre. Notons que les bords du STI sont améliorés (cela est dû à l'effet de la couche de poly-silicium qui a réduit les contraintes mécaniques au niveau des bords).

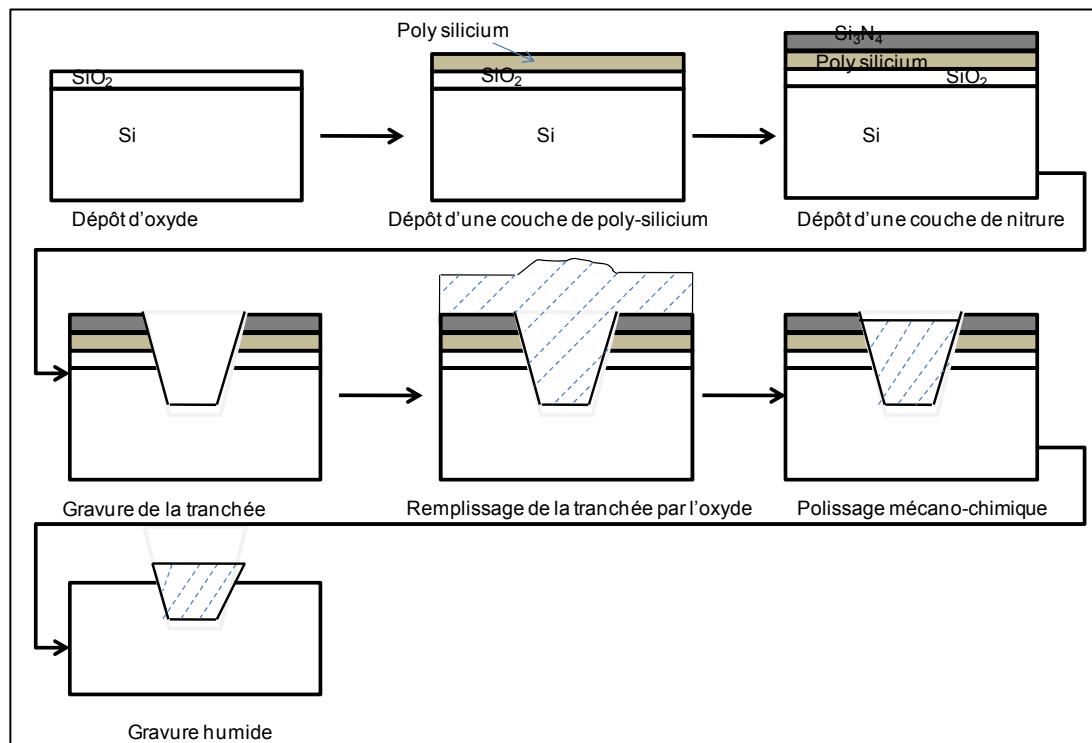


Figure 2.20- Représentation de la fabrication du PB-STI

2.2.4.4.b Le Direct STI

Le procédé CMP doit se faire assister par « la gravure inverse active ». A ceci, s'ajoute une faible sélectivité des abrasifs avec le nitre qui influence l'uniformité du polissage. De plus, le CMP détecte la fin du polissage après avoir consommé près de 500Å du nitre au lieu de s'arrêter dessus. Car le temps de polissage n'est pas fixe, le temps de polissage diffère d'une plaque à une autre. Le système régissant la fin du polissage est un système embarqué sur la machine appelé système « fin de point ou end point »³⁶.

Les technologies évoluent conjointement avec les outils de production. Cette technologie de polissage subit un effort de recherche de la part des équipementiers. Des tissus de polissage plus innovants et des abrasifs très sélectifs moins agressifs pour la plaque sont développés. Plus récemment, des abrasifs plus sélectifs ont été développés. Ceux ci sont sélectifs avec le nitre, autrement dit, le polissage s'arrête à la couche de nitre. Ceci nous permet d'avoir

³⁶ Système détectant la fin de polissage.

une bonne uniformité de l’oxyde et donc, on n’a pas besoin de l’étape de la gravure inverse. Dans le cas du Direct STI, l’épaisseur d’oxyde HDPCVD est très importante environ 8500Å (5500 Å dans le cas du STI standard). Ceci nous permet de se passer d’une autre étape de fabrication à savoir le dépôt TEOS. En conclusion, ce procédé innovant permet de réduire deux opérations du procédé standard (dépôt TEOS et gravure inverse). La réduction de deux opérations de fabrication représente un enjeu économique très important.

2.3 Impact du STI sur le rendement électrique des circuits

Les améliorations apportées à la brique STI sont de légères modifications sur certains procédés de fabrication. Cependant le concept général de l’isolation par les tranchées peu profondes reste le même. Il est important de comprendre l’impact du STI sur la suite des opérations de fabrication et sur le rendement final électrique. Nous pouvons citer l’impact de la présence des trous dans la tranchée d’oxyde. Ces trous peuvent causer des problèmes électriques par la suite pour les circuits. Les dislocations peuvent réduire la qualité d’isolation du STI. Ceci se manifeste électriquement par des court-circuits en général. La forme des bords du STI [BRYANT.1994] est également très critique. Certaines formes des bords du STI favorisent l’apparition des courants de fuite. Enfin, des variabilités dans les mesures de la hauteur de marche et de la profondeur de tranchée sont souvent constatées. Il faut noter que tous ces problèmes peuvent être corrélés aux opérations de fabrication au niveau du STI. Nous reviendrons plus en détails dans le chapitre 4 sur les deux premiers problèmes c’est à dire sur la nécessité de faire un dépôt d’oxyde avec une bonne conformité pour éviter les trous. Nous détaillerons longuement la variabilité de la hauteur de marche dans le chapitre 4 en vue de justifier le déploiement de boucles de régulation pour stabiliser la hauteur de marche. Les paragraphes précédents décrivaient les problèmes au niveau local. Sur le plan global le STI peut impacter le circuit d’un point de vue électrique et mécanique.

2.3.1 Impact du STI sur les paramètres électriques

L’impact du STI sur le rendement final est difficile à prévoir, car il y a beaucoup d’étapes de fabrication qui se succèdent après le STI et il faut compter 2 à 3 mois pour avoir les résultats des tests électriques. Ce sont les paramètres géométriques (la forme des coins du STI, la profondeur de tranchée, la hauteur de marche) qui sont analysés puis comparés avec les paramètres électriques ou paramètres mécaniques (contraintes mécaniques) pour éventuellement en déduire des corrélations. Ces corrélations sont à prendre avec beaucoup de prudence, car la variabilité de certains paramètres n’est pas due au STI au premier ordre mais à des blocs comme la grille par exemple.

Nous exposerons au chapitre 3, les différents résultats de corrélations entre les paramètres électriques et les résultats de métrologie au niveau du STI. Les paramètres de sortie qui sont étudiés sont la hauteur de marche et la profondeur de tranchée. Ces deux paramètres sont corrélés entre eux et la variabilité de la profondeur de tranchée implique une variabilité de la hauteur de marche dans le cas du STI standard. De même la variation de ces deux paramètres implique directement la variation de certains paramètres au niveau du test électrique. Dans le cas où l’impact du STI sur les paramètres électriques n’est pas du premier ordre, une faible corrélation est observée et parfois inexistante. Par exemple, si nous corrélons la tension de claquage avec les paramètres du STI, les résultats des corrélations sont très faibles car le STI a peu d’influence sur ce paramètre. Différents travaux ont été publiés sur l’impact du STI sur les paramètres électriques [SHIOZAWA 2001, HONG 2002, PAVAN 2003]. Les paramètres explorés sont souvent la résistance carrée, la tension de seuil, le

courant de fuite... Les résultats publiés présentent souvent des corrélations moyennes voir parfois faibles. Ceci est dû principalement à l'usage du STI standard.

On verra dans la suite de ce manuscrit, comment grâce à l'utilisation du Direct STI on obtient des corrélations fortes entre les paramètres de mesure physiques et les paramètres électriques. Alors qu'en utilisant le STI standard, les corrélations ne peuvent pas être déduites facilement. D'autres corrélations sont faites entre le STI et le test électrique final EWS (Electrical Wafer Sorting). Les tests EWS sont eux sélectifs³⁷. Les résultats du EWS permettent de faire une corrélation avec les opérations précédentes et les différents équipements ayant conduit au rendement final. Donc des corrélations peuvent être déduites entre les valeurs des paramètres STI et les résultats EWS [BELHARET 2007]. Par contre, si durant l'élaboration des circuits, des erreurs surviennent et sont non détectées par les outils FDC au niveau de l'équipement, on en aura une signature particulière au test final. Dans ce cas, il est très difficile de déduire des bonnes corrélations.

2.3.2 Impact du STI sur les contraintes mécaniques

La structure de bande d'un semi-conducteur découle de la symétrie de son cristal par l'intermédiaire de l'équation de Schrödinger. L'application de contraintes mécaniques sur un matériau cristallin peut modifier la périodicité de son réseau. Donc la contrainte mécanique agit directement sur la structure de bande du matériau et ainsi sur ses propriétés électroniques. Les contraintes mécaniques qui se multiplient avec la réduction des dimensions des dispositifs semi-conducteurs sont à prendre en considération car les effets sur les propriétés électriques sont considérables. L'isolation STI devient une source importante de contraintes et il est important d'évaluer aujourd'hui son impact sur les paramètres électriques des transistors MOS. Pour mieux illustrer ce concept, la figure 2.21 illustre parfaitement comment la contrainte peut être plus importante quand la distance « a » (distance entre le STI et la grille) est réduite.

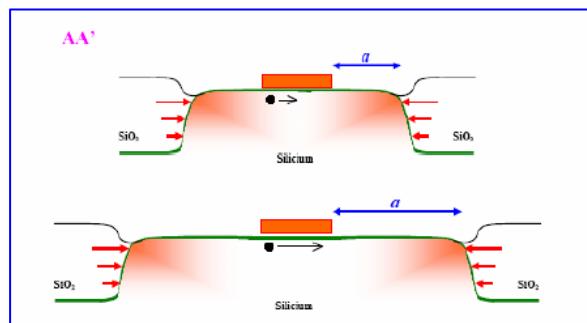


Figure 2.21- Comparaison de deux structures ayant différentes valeurs de la distance « a »
[GALLON 2004, GALLON 2004a]

Gallon et al [GALLON.2004, GALLON.2004 a] ont démontré qu'en réduisant cette distance « a », électriquement cela se traduisait par un changement de mobilité des porteurs. Su & al [SU.2003] ont démontré que le stress induit par le STI impacte non seulement la mobilité mais aussi d'autres paramètres électriques tels que la tension de seuil et d'autres paramètres de second ordre tel que le courant drain-source. L'effet du stress causé par le STI sur la zone active ne peut être expliqué physiquement que lors des traitements thermiques. La silice et le silicium se dilatent différemment laissant en fin de procédé des contraintes

³⁷ La sélection des puces qui fonctionnent et les puces erronées.

résiduelles dans le substrat. Le phénomène d’oxydation joue lui aussi un rôle dans l’apparition des contraintes au niveau de la zone active. Il est notamment possible que lors d’un traitement thermique sous atmosphère oxydante que l’oxygène diffuse à l’interface zone active/STI pour oxyder le silicium en bord de zone. La figure 2.22 montre un exemple de formation des molécules d’oxygène à l’interface Si/STI. On voit clairement à travers cet exemple comment la zone active peut être comprimée par le STI.

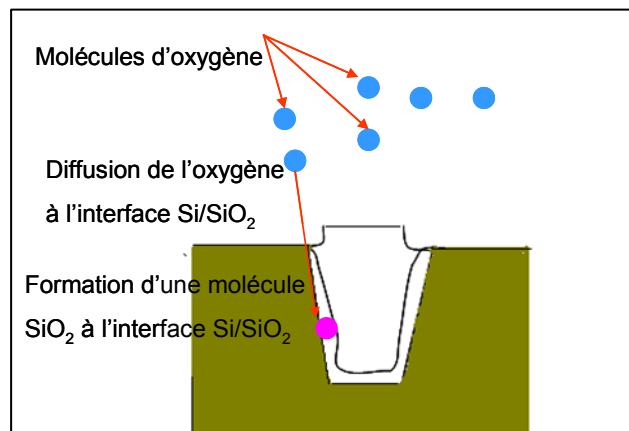


Figure 2.22- Compression de la zone active par croissance d’oxyde

Plusieurs facteurs sont responsables de la génération de ces contraintes parmi lesquelles la forme des bords du STI. Le choix des paramètres de procédés est évidemment un facteur responsable de la génération des contraintes mécaniques. Nous pouvons citer par exemple l’angle de gravure lors de la gravure sèche de la tranchée, la valeur de la température lors de l’oxydation thermique des bords de la tranchée, la valeur de la température du procédé de dépôt chimique en phase vapeur assisté par un plasma haute densité (HDP - CVD)... Les contraintes sont maximales en bords d’isolation. Elles se répartissent vers le centre de la zone comme cela est illustré sur la figure 2.23.

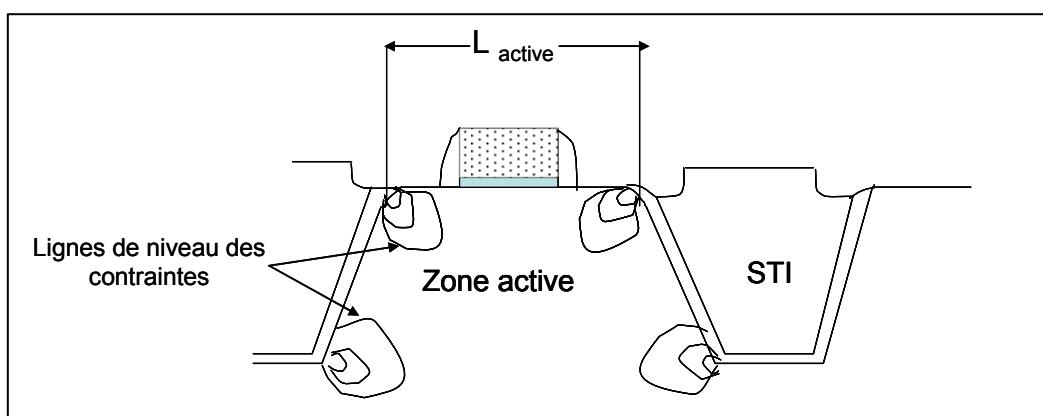


Figure 2.23- Localisation des contraintes maximales en bord du STI.

Avec la réduction de la longueur du canal et de la longueur du canal d’active (L_{active}), les contraintes maximales vont occuper une proportion croissante du volume de zone active comme cela est illustré sur la figure 2.23. Avec la réduction des dimensions des transistors, l’effet des contraintes mécaniques sera donc plus sensible au fil des générations de transistors.

Beaucoup de travaux ont été publiés sur la minimisation des contraintes mécaniques mais la plupart des travaux ont été focalisés sur des opérations bien spécifiques de la brique STI. Il s'agit généralement d'optimiser les paramètres d'un procédé donné identifié comme responsable de la génération des contraintes mécaniques. On peut citer les travaux d' Ishimaru & al [ISHIMARU1997] qui ont réussi à corrélérer le stress mécanique avec la température opératoire du dépôt TEOS. Shiozawa & al [SHIOZAWA. 2001] ont démontré l'influence du dépôt HDP-CVD sur le stress, ainsi que la contribution du stress sur le courant de fuite. Vatanabe et al [VATANABE.1996] ont travaillé sur la forme des coins supérieurs du STI. En améliorant cette forme (arrondie), le stress est réduit et le courant de fuite aussi.

Nous allons exposer dans le chapitre 3 nos travaux de recherches sur ce point et nous allons proposer notre stratégie pour le contrôle des contraintes mécaniques.

2.5 Conclusion

Dans les paragraphes précédents, nous nous sommes focalisés sur les procédés de réalisation des différents modes d'isolation tout en décrivant les améliorations qui se sont succédées au fil des procédés. Nous avons constaté que l'amélioration des procédés au niveau du bloc d'isolation répond essentiellement au fait qu'il faut éviter des problèmes électriques dans les circuits. Cette étude bibliographique est très importante pour la compréhension des effets physiques des éléments constituant la brique d'isolation, ainsi que des mécanismes de chaque procédé. Nous avons constaté que l'isolation par tranchée peu profonde (STI) est la technique d'isolation qui répond aux exigences électriques des technologies avancées. Cette technique nous permet d'avoir des surfaces d'oxydes planes en éliminant complètement le phénomène de « Bec d'oiseau », cet avantage nous permet donc d'accroître la densité d'intégration. Nous avons vu qu'on pouvait encore améliorer les performances qu'offre le STI en opérant des modifications légères sur certains procédés. Nous avons cité quelques exemples, parmi eux le « Direct STI » qui offre une bonne planarité de la surface d'oxyde. Dans les prochains chapitres, consacrés aux résultats expérimentaux, nous nous limiterons au STI standard et au Direct STI.

Nous avons également évoqué dans ce chapitre la criticité de ce module d'isolation sur le rendement final électrique. Nous avons cité un certain nombre de travaux de recherches qui ont porté sur l'impact du STI sur certains paramètres électriques ainsi que sur les contraintes mécaniques. Dans le prochain chapitre, nous allons exposer nos travaux de recherche pour expliquer et quantifier l'impact du STI sur le rendement électrique final.

Références bibliographiques

- [BASSOUS 1976] E.Bassous, H.N. Yu and V.Maniscalo, Topology of silicon structures with recessed SiO₂, J. Electrochem.Soc.: Solid-State Science And Technology, Vol. 123, No.11,p-1729-1737, 1976.
- [BELHARET 2007] Djaffar BELHARET, Agnès ROUSSY, Pascal FORNARA, Philippe COLLOT, Jacques PINATON, Yield enhancement on the wafer edge by adjusting the retaining ring pressure during the CMP process, The 10th Technical and Scientific Meeting of ARCSIS. 6-7 Décembre, Fuveau (France) 2007.
- [BRYANT 1994] Andres Bryant, Wilfried Hänsch, and Toshio Mii, “Characteristics of CMOS Device Isolation for the ULSI Age,” *International Electron Devices Meeting, Technical Digest*, pp. 671-674, 1994.
- [CHAPMAN 1987] Richard A. Chapman, Roger A. Haken, David A. Bell,C.C wei, Robert H. Havemann, Thomas E. Tang, Thomas C Holloway and Rebecca J.Gale, An 0,8 µm CMOS technology for high performance logic applications, Tech. Dig. IEDM, 1987, p. 362.
- [DEUNEVILLE 2002] Alain Deuneville, Cours DEA Microélectronique, « cours des technologies des procédés et matériaux », année 2002-2003
- [CHANG 1999] C.-P. Chang, S.F. Shive, S.C. Kuehne, Y. Ma, H. Vuong, F.H. Baumann, M. Bude, E.J. Lloyd, C.S. Pai, M.A. Abdelgadi, R.Dail, C.T. Llu, K.P. Cheung, J.I. Colonell, W.Y.C. Lai, J.F. Miner, H. Vaidya, R.C. Liu, J.T. Clemens, Enabling Shallow Trench Isolation for 0.1µm Technologies and Beyond, 1999 Symposium on VLSI Technology Digest of Technical Papers, p.161-162,1999.
- [DAVARI 1988] B. Davari, C. Koburger, T.Furukawa, Y.Taur, W. Noble, A. Megdanis, J. Warnock, and J. Mauer, A variable size shallow trench isolation (STI) technology with diffused sidewall doping for submicron CMOS, Tech.Dig. IEDM, 1988, p.92
- [DEROUX-DAUPHIN 1985] P. Deroux-Dauphin & J.P. Gonchond, Physical and electrical characterization of a SILO isolation structure, IEEE trans. Electron Dev., Nov. 1985, p. 2392.
- [FORSBERG 2004] Markus Forsberg, Ted Johansson, Wei Liu, and Manoj Vellaikald, A Shallow and Deep Trench Isolation Process Module for RF BiCMOS, *Journal of The Electrochemical Society*, **151** (12), p.839-846, 2004.
- [FUSE 1987] Genshu .Fuse, Masanori Fokumoto, Akihira Shinohara, Shinji Odanaka, Masaru sasago and Takashi Ohzone, A new isolation method with boron-implanted sidewalls for controlling narrow-width effect, IEDM Tech,Dig., 1987, p. 732
- [GALLON 2004] C. Gallon, G. Reimbold, *Member, IEEE*, Gérard Ghibaudo, *Senior Member, IEEE*, R. A. Bianchi, R. Gwoziecki,S. Orain, E. Robilliart, C. Raynaud, and H. Dansas, Electrical Analysis of Mechanical Stress Induced by STI in Short MOSFETs Using Externally Applied Stress, *IEEE TRANSACTIONS ON ELECTRON DEVICES*, VOL. 51, NO. 8, AUGUST 2004,p1254-1261
- [GALLON 2004a] C. Gallon, G. Reimbold , G. Ghibaudo , R.A. Bianchi , R. Gwoziecki, Electrical analysis of external mechanical stress effects in short channel MOSFETs on (0 0 1) silicon, *Solid-State Electronics* 48 (2004) 561–566
- [GHEZZO 1987] M. Ghezzo, M.J.Kim, J.F.Norton and R.J.Saia, “laterally sealed LOCOS isolation”, J. Electrochem.Soc.: Solid-State Science And Technology, Vol. 134, No.6, p-1475.
- [GRAN 2000] Terence Gan, Modeling of Chemical Mechanical Polishing for Shallow Trench Isolation, these de doctorat , may 2000.
- [GOODWIN 1982] C.A.Goodwin, J.W. Grossman,“Mos gate oxide defects related to treatment of silicon nitride coated wafers prior to local oxidation”, J. Electrochem.Soc.: Solid-State Science And Technology, Vol. 129, No.5,p-11066, 1982.
- [HONG 2002] S.W.Hong,G.Y. Jin,H.W. Seo,D.I.Lee, J.H. Song, J.Y.Noh,Y.C. Oh,J.Y.Kim, D.H. Kim,H.H. Kim,D.J. Won, W.W. Lee,D.H. Song, K.Y.Lee, andW.S. Lee “Evaluation of

“STI Degradation Causing DRAM Standby Current Failure in Burn-in Mode Operation Using a Carrier Injection Method”, IEEE 02CH37320.40ih Annual International Reliability Physics Symposium, Dallas, Texas, 2002.

[HORITA 2000] K.Horita, T.Kuroi, K.Shiozawa, K.Eikyu, K.Goto, Y.Inoue and M.Inuishi, Advanced shallow trench isolation to suppress the inverse narrow channel effects for 0.24 μ m pitch isolation and beyond, in VLSI Tech.Dig, p.178, 2000.

[HOSHI 1986], Naoya Hoshi, Shigeki Kayama, Toshiyuki Nishihara, Jun-ichi Aoyama, Takaaki Komatsu, Takashi Shimada, “ 1,0 μ m CMOS process for highly stable tera-ohm polysilicon load 1MB SRAM, Tech. Dig. IEDM, 1986, p. 300.

[HUI 1982] John Chi-Hung Hui, Tzu-Yin Chiu, Siu-Weng S.Wong and William G. Oldham, IEEE transactions on Electron devices, Vol.ED-29, No.4, 1982

[ISHIMARU 1997] K. Ishimaru, F . Matsuoka, M. Takahashi, M. Nishigohri, Y . Okayama, Y. Unno,M. Yabuki, K. Umezawa, N. Tsuchiya, O. Fujii, and M. Kinugawa, Mechanical Stress Induced MOSFET Punch-through and Process Optimization for Deep Submicron TEOS-03 filled STI Device, 1997 Symposium on VLSI Technology Digest of Technical Papers.

[JUANG 1999] M.H. Juang, Trade-off and process consideration for scalable poly-Si buffered LOCOS technology, Solid-State Electronics 43 (1999) 2067-2073.

[KOOI 1976] E.Kooi, J.G. Van Liverop and J.A.Appels, “Formation of Silicon Nitride at Si-SiO₂ interface during local oxidation of silicon and during heat treatment of oxidized silicon in NH₃ gas, J.Electrochem.Soc., Vol. 123, No.7,p-1117.

[LI 2002] C.H.Li, K.C.Tu, H.C.Chu, I.H.ChANG, W.R.Liaw, H.F.Lee, W.Y.Lien, M.H.Tsai, W.J.Liang, W.G.Yeh, H.M.Chou, C.Y.Chen and M.H.Chi, A robust shallow trench isolation (STI) with SiN Pull-back process for advanced DRAM Technology, 2002 IEEE/SEMI Advanced Semiconductor Manufacturing Conference.

[LIEN 2002] W.Y.Lien,W.G.Yeh, C.H.Li,C.Tu,I.H.Chang,W.R.Liaw,H.F.Lee, H.M. Chou,C.Y. Chen, and M.H. Chi, A manufacturable shallow trench isolation process for sub-0,2 μ m DRAM technologies, 2002 IEEE/SEMI Advanced Semiconductor Manufacturing Conference.

[MIKOSHIBA 1984] M. Mikoshiba, T. Homma and K. Hamano, A new trench isolation technology as a replacement of LOCOS, IEDM Tech. Dig., 1984, p. 578-581

[MIZUNO 1987] Tomohisa Mizuno, Shizuo Sawada, Satoshi Maeda and Satoshi Shindozaki, “Oxidation rate reduction in the submicrometer LOCOS process”, IEEE transactions on Electron devices, Vol.ED-34, No.11, 1987.

[NAG 1996] S. Nag, A. Chatterjee, K. Taylor, I. Ali, S. O'Brien, S. Aur, J.D.Luttmer and I-C Chen, Comparative Evaluation of Gap-Fill Dielectrics in Shallow Trench Isolation for Sub-0.25 μ m Technologies, IEEE, IEDM , p.841-844, 1996.

[NIITSU 1985] Y.Niitsu et al, latchup free CMOS structure using shallow trench isolation, Tech. Dig. IEDM, 1985, p.509

[NISHIHARA 1988] Toshiyuki Nishihara, Kazuhiko Tokunaga and Kazuyoshi Kobayashi, “ A 0,5 μ m isolation technology using advanced poly silicon pad LOCOS (APPL), Tech. Dig. IEDM, 1988, p. 100.

[NONDAKUMAR 1998] M. Nandakumar, A. Chatterjee, S. Sridhar, K. Joyner, M. Rodder and I.-C.Chen, Shallow Trench Isolation for advanced ULSI CMOS Technologies, IEDM 98-133, 1998 IEEE.

[PAN 1998] J. Tony Pan, Dennis Ouma, Ping Li, Duane Boning, Fritz Redeker, James Chung, Jason Whitby, PLANARIZATION AND INTEGRATION OF SHALLOW TRENCH ISOLATION, 1998 VMIC, Santa Clara, CA, June 1998.

[PAVAN 2003] A. Pavan, D. Brazzelli, M. Aiello, C. Capolupo, C.Clementi, C. Cremonesi, A. Ghetti, An Analysis of the Effect of the Steps for Isolation Formation on STI Process

Integration, Proceedings - Electrochemical Society, ULSI process integration III (Paris, 28 April - 2 May 2003)

[SHANKOFF 1980] T.A. Shankoff, T.T. Sheng, S.E. Haszko, R.B. Marcus and T.E. Smith, "3bird's beak configuration and elimination of gate oxide thinning produced during selective oxidation", J. Electrochem.Soc.: Solid-State Science And Technology, Vol. 127, No.1,p-216, 1980.

[SHIOZAWA 2001] Katsuomi Shiozawa, Katsuyuki Horita, Takashi Kuroi, Yuji Abe and Takahisa Eimori, Impact of STI stress on the junction characteristics, International workshop on junction technology, p.19-20, 2001

[SU.2003] Ke-Wei Su, Yi-Ming Sheu, Chung-Kai Lin, Sheng-Jier Yang, Wen-Jya Liang, Xuemei Xi, Chung-Shi Chiang, Jaw-Kang Her, Yu-Tai Chia, Carlos H. Dim, and Chenming Hu, A Scaleable Model for STI Mechanical Stress Effect on Layout Dependence of MOS Electrical Characteristics, IEEE 2003 CUSTOM INTEGRATED CIRCUITS CONFERENCE.

[TAMAKI 1988] Yoichi Tamaki, Seiichi Isomae, kazuhiko Sagara, Tokuo Kure, Masao Kawamura, "Evaluation of dislocation generation in U-Groove Isolation", J. Electrochem.Soc.: Solid-State Science And Technology, Vol. 135, No.3,p-726, 1988.

[TSAI 1988] H.H.Tsai, S.M. Chen, H-B. Chen and C-Y. Wu, An evaluation of FUROX isolation technology for VLSI/nMOSFET fabrication, IEEE transactions on Electron devices, Vol.ED-33, No.3, p.275-284,1988.

[VATANABE 1996] Hiroshi vatanabe, Kazuhiro Shimizu, Yuj i Takeuchi and Seiichi Aritome, Corner-Rounded Shallow Trench Isolation Technology to Reduce the Stress-Induced Tunnel Oxide Leakage Current for Highly Reliable Flash Memories, IEDM , IEEE, 1996.

[WOLF 1990] S.Wolf, Silicon processing for the VLSI Era, Vol2. Process Integration, LATTICE Press, California, 1990.

Chapitre 3

Corrélation entre la morphologie de la tranchée d'isolation et les propriétés électriques des transistors

Dans ce chapitre, nous allons mettre en évidence l'influence du STI sur certains paramètres électriques ainsi que sur les contraintes mécaniques. Nous allons détailler toutes les analyses des corrélations entre le STI et les paramètres électriques. Nous allons exposer également les travaux de simulations et de caractérisations électriques pour démontrer l'effet des paramètres électriques qui se corrèlent bien à la morphologie du STI. Dans la deuxième partie de ce chapitre, nous allons étudier l'effet du STI sur les contraintes mécaniques et en particulier l'effet de la distance caractérisant l'éloignement du STI par rapport à la grille. Nous allons dresser tous les résultats expérimentaux. Nous allons confirmer ces résultats à travers des simulations.

3.1 Introduction

Avant d'aborder les différentes corrélations qui existent entre la morphologie du STI et les paramètres électriques, nous allons rappeler tout d'abord le principe de fonctionnement d'un transistor MOS.

3.1.1 Rappels sur les principes de base d'un transistor MOS

Le schéma du transistor MOS est présenté sur la figure 3.1 avec les principales dimensions qui y sont associées. Le transistor est constitué de 4 électrodes : la grille (G), la source (S), le drain (D) et le substrat (B). Les grandeurs caractéristiques sont la longueur de grille notée (L), la largeur de grille notée (W) et l'épaisseur de l'oxyde de grille notée (Ep ox).

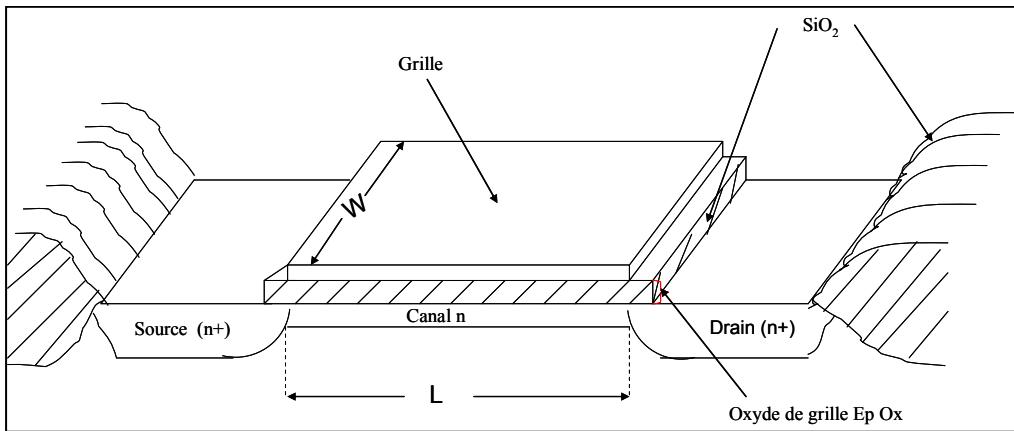


Figure 3.1- Coupe schématique d'un transistor MOS avec une vue de dessus

Le principe du transistor est de réguler le flux de porteurs qui transitent de la source au drain grâce à la tension de grille et ce, sous différents régimes commandés par la chute de potentiel entre le drain et la source. Les équations permettant de modéliser le courant dans les différentes gammes de tension sont rappelées ici.

3.1.2 Influence de la tension de grille sur la résistivité du canal

La caractéristique classique du courant drain en fonction de la tension de grille est représentée sur la figure 3.2.

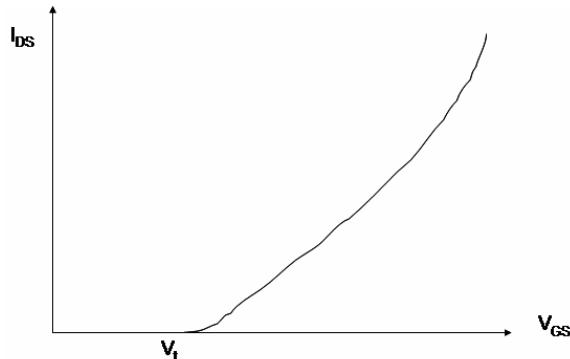


Figure 3.2- Caractéristique de transfert d'un transistor MOS avec une tension V_{DS} constante

Lorsque la tension de grille augmente, le canal se peuple et sa résistivité diminue. La charge commandée par le canal peut être modélisée simplement par:

$$Q_{canal} = W \cdot L \cdot C_{ox} \cdot (V_{GS} - V_t - \frac{1}{2}V_{DS}) \quad (3.1)$$

C_{ox} étant la capacité de l'oxyde par unité de surface : $C_{ox} = \frac{\epsilon_0 \epsilon_{ox}}{Ep_{ox}}$

Dans ces relations, ϵ_0 , ϵ_{ox} sont la permittivité diélectrique du vide et la permittivité relative de l'oxyde respectivement. Le courant afflue vers le drain dès que V_G dépasse la “tension de seuil”³⁸, V_t .

3.1.3 Régimes de fonctionnement du transistor MOS

Différents régimes de conduction se succèdent lorsque la tension au niveau du drain V_{DS} augmente. Dans la pratique, il est utile de travailler avec la transconductance (g_m), grandeur directement liée aux caractéristiques courant tension par: $g_m = \frac{\partial I_{DS}}{\partial V_{GS}} \Big|_{V_{DS} \text{ constante}}$

On distingue 3 régimes :

- Régime ohmique ou linéaire ($V_{DS} < V_{GS} - V_t$)

$$I_{DSlin} = \beta \cdot (V_{GS} - V_t - \frac{1}{2}V_{DS}) \cdot V_{DS} \quad (3.2)$$

où : $\beta = \frac{W}{L} \cdot \mu \cdot C_{ox}$ et $g_{mlin} = \beta \cdot V_{DS}$, μ désigne la mobilité des porteurs.

- Point de Pincement

Le point de pincement marque le début de la saturation du courant. Ainsi, V_P est aussi appelée tension de saturation. Le courant au point de pincement est appelé courant de saturation I_{DSat} .

$$I_{DSat} = \frac{1}{2} \beta \cdot (V_{GS} - V_t)^2 = \frac{1}{2} \beta \cdot V_{DSat}^2 \quad (3.3)$$

- Régime de saturation ($V_{DS} \geq V_{GS} - V_t$)

Lorsque la tension drain augmente au delà de la tension de saturation, le point de pincement s'éloigne du drain mais rien ne change en terme de densité de charges ni en terme de tension appliquée entre la source et le point de pincement. Ainsi le courant drain est constant. Son expression est toujours donnée par l'équation précédente. On a de plus $g_{mSat} = \beta \cdot V_{DSat}$

³⁸ La tension de seuil est définie comme étant la tension V_{GS} entre la grille et la source pour laquelle la zone d'inversion apparaît, c'est à dire la création du canal de conduction entre le drain et la source.

3.2 Corrélation entre la morphologie de la tranchée d'isolation et les propriétés électriques des transistors

Dans cette étude, on va comparer les corrélations des paramètres électriques avec les paramètres du STI standard et les paramètres du direct STI. Les corrélations obtenues sont faites soit par des analyses effectuées sur un historique des données, soit par l'exploitation des résultats de plans d'expériences spécifiques³⁹ effectués.

L'ensemble des mesures électriques potentiellement influençables par des variations de procédés de fabrication de la tranchée d'isolation seront prises en compte. Les mesures étudiées se regroupent en plusieurs classes :

- Dimensions électriques des structures
- Résistances carrées des structures
- Courants et tensions caractéristiques des transistors
- Tension de claquage et courant de fuite entre deux zones actives.

Pour les études de corrélations, nous avons analysé les données de mesure d'un produit dont la valeur cible de l'épaisseur STI est de 3000Å . Pour les analyses de simulation de la résistance carrée (figure 3.10), nous avons exploré des épaisseurs plus importantes allant de 3500 à 4700Å avec des énergies d'implantation de 440 keV et de 220keV pour les transistors PMOS et NMOS respectivement.

3.2.1 Mesures de résistance

La première famille de mesures destinées à tester le comportement électrique de l'isolation est composée de mesures de résistances effectuées sur des zones non actives de différentes dimensions.

3.2.1.1 Résistances caisson STI (Well STI)

3.2.1.1.a Rappel de la formation des caissons (well STI)

A la fin de la réalisation de la brique STI, les zones actives sont définies. La première étape post STI consiste à la réalisation des caissons N et P (N Well et P Well) suivant le type de dopage. Ces caissons sont réalisés par l'implantation ionique et traversent localement la tranchée STI. Un exemple de structure de test est illustré sur la figure 3.3. Cette structure est conçue pour surveiller la résistance (Well STI) pour les transistors PMOS. Nous distinguons

³⁹ Dans la fabrication des technologies modernes, les limites de tolérances des paramètres de mesure sont très étroites, et donc les plages des hauteurs de marches obtenues sont très limitées (généralement de 300 à 500 A°). Pour généraliser une loi issue d'une corrélation (par exemple linéaire), il faut explorer les zones en dehors de cette plage. Pour cela, des hauteurs de marches allant de 150 à 900 A° ont été réalisées. La gamme des valeurs de hauteur de marche obtenue est le résultat des expériences conçues. Le principe de ce plan d'expériences est basé sur la modification de certains paramètres au niveau du procédé CMP et au niveau de la gravure humide (WET).

sur la figure 3.3 l'épaisseur de STI qui est la somme de la hauteur de marche et la profondeur de tranchée. Cette épaisseur STI est déduite par des mesures de capacités.

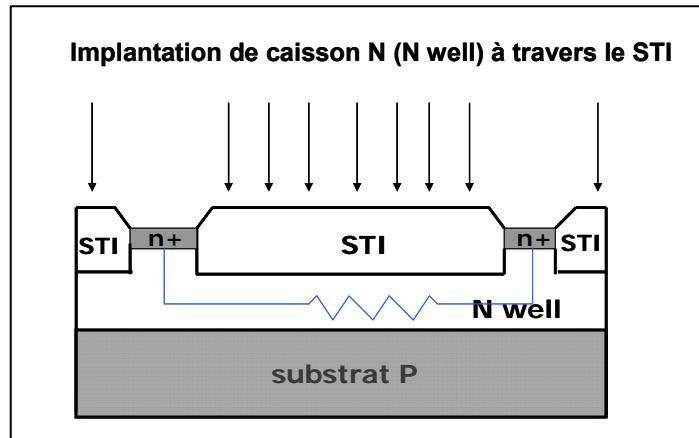


Figure 3.3- Structure de test employée pour la mesure de la résistance carrée Nwell STI dans une structure PMOS

3.2.1. b Résistances N Well STI des transistors PMOS

Corrélation épaisseur de STI – Résistance Nwell STI

Des analyses de données ont été effectuées sur un historique, comme illustré sur la figure 3.4. Il y a une très bonne corrélation linéaire entre la résistance Nwell STI et l'épaisseur STI. On verra plus loin comment cette corrélation est expliquée physiquement.

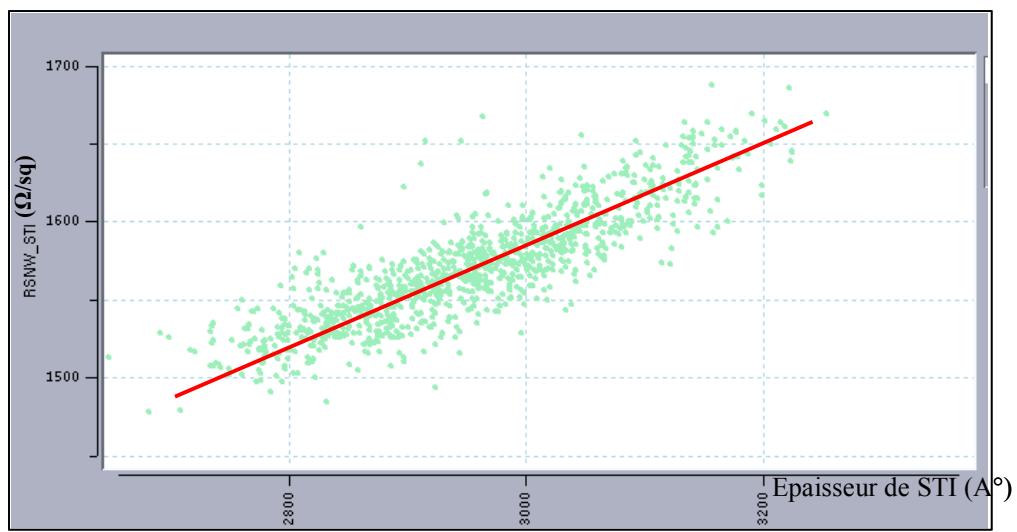


Figure 3.4- Evolution de la résistance carrée Nwell STI en fonction de l'épaisseur de STI

Corrélation hauteur de marche – Résistance Nwell STI

Afin de voir s'il y a un effet de la hauteur de marche sur la résistance Nwell STI, des expériences ont été faites pour avoir différentes hauteurs de marches. Les données des mesures de résistances ont été comparées avec les valeurs de hauteur de marche.

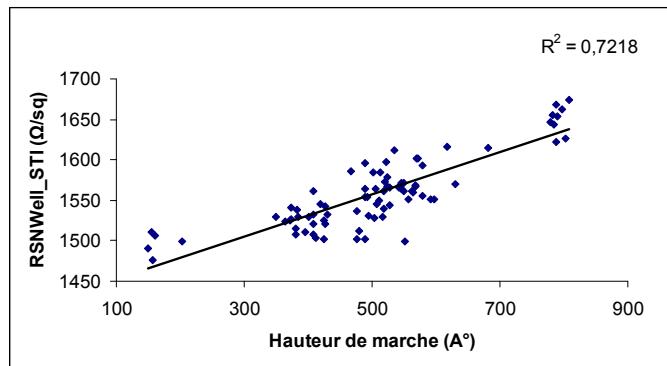


Figure 3.5- Evolution de la résistance carrée Nwell STI en fonction de la hauteur de marche

Comme cela est illustré sur la figure 3.5, il y a une très bonne corrélation linéaire entre la résistance Nwell STI et l'épaisseur STI. Nous allons expliquer les fortes corrélations observées dans les prochains paragraphes.

3.2.1.1.c Résistances PWell STI des transistors NMOS

De la même façon, des analyses ont été faites sur un historique de données pour voir l'effet de la hauteur de marche et de l'épaisseur STI sur la résistance Pwell STI. Ces résultats sont illustrés sur la figure 3.6.

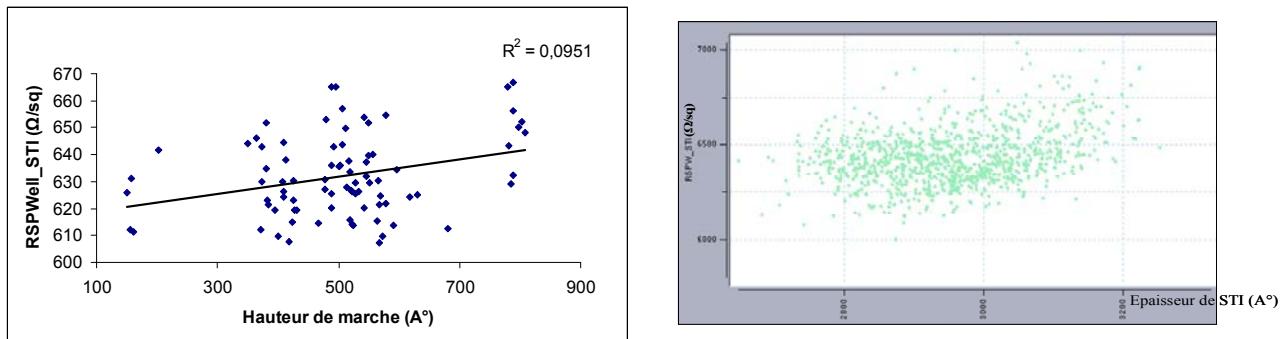


Figure 3.6- Evolution de la résistance carrée Pwell STI en fonction de la hauteur de marche et de l'épaisseur STI

Contrairement aux transistors NMOS où on a enregistré de fortes corrélations, dans le cas des PMOS et comme on peut le constater sur la figure 3.6, il y a une faible corrélation entre les paramètres STI et la résistance Pwell STI. Pour comprendre cette différence de comportement, on a procédé à des simulations décrites par la suite.

3.2.1.2 Simulations

Nous allons simuler des plaques ayant des épaisseurs STI différentes.

3.2.1.2.a Programme de Simulation

L'équipe de TDM (Technology and Device Modeling) de ST Rousset utilise un programme de simulation global. Il s'agit de Dios release 10.0. La résistance carrée est extraite de la manière suivante : En fonction de l'épaisseur de STI, le profil de dopage est simulé pour les deux types de transistors (NMOS et PMOS). Finalement les résistances carrées sont extraites suivant les relations suivantes :

$$r_n = \frac{1}{g_n} = \frac{1}{e\mu_n \int_{X_0}^X n(x)dx} \quad (3.4)$$

$$r_p = \frac{1}{g_p} = \frac{1}{e\mu_p \int_{X_0}^X p(x)dx} \quad (3.5)$$

Où : X_0 est l'épaisseur STI (Somme de la profondeur de tranchée et la hauteur de marche), X_0 se situe à l'interface Si/SiO₂, $n(x)$ est la densité des électrons et $p(x)$ est la densité des trous.

La résistance carrée mesurée est donc inversement proportionnelle à la concentration de dopants actifs confinés entre l'interface STI/Si (X_0) et la jonction électrique (X). Nous allons représenter cette zone comme hachurée sur les graphes de profils de dopants.

3.2.1.2.b Conditions de simulations

Le but de notre simulation est d'explorer par modélisation 1D les résistances carrées du caisson implanté sous la tranchée STI. La figure 3.7 montre les structures 1D pour les transistors PMOS et NMOS. Les différences de couleurs dans la structure représentent la différence de dopage dans les différentes zones de la structure (STI, caisson et substrat)

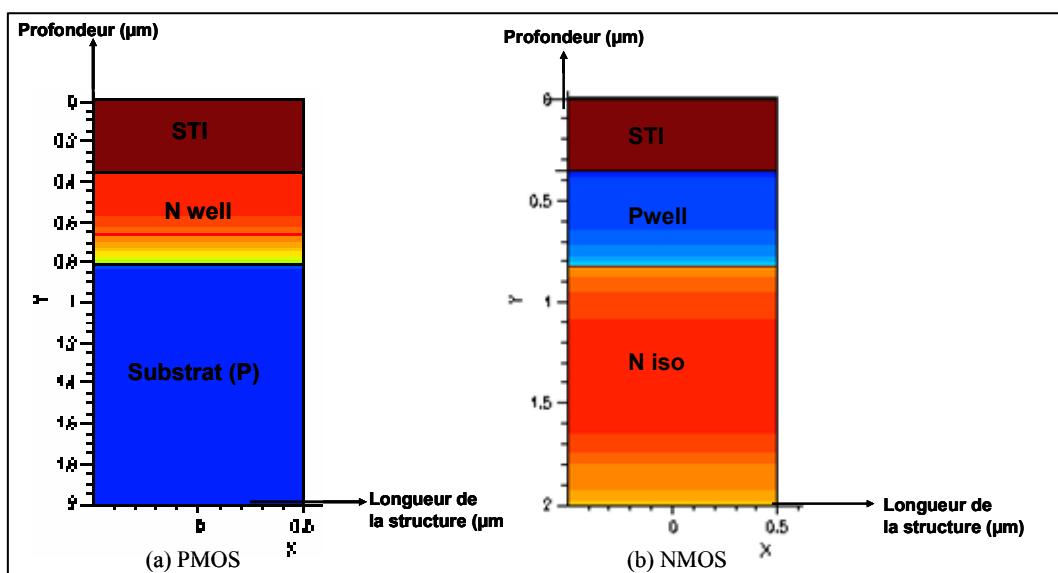


Figure 3.7- Coupe 1D de la structure pour les transistors PMOS et NMOS

Les substrats utilisés sont des substrats de type P. Pour réaliser un transistor PMOS, il suffit de créer un caisson N (Nwell) par implantation comme cela est illustré sur la figure 3.7a. Pour la réalisation d'un transistor NMOS, il faut créer un caisson de type P (P well), mais comme le substrat est de même nature que le caisson, il est obligatoire de réaliser une couche de type N avant la réalisation de caisson. Cette couche est représentée par « N iso » dans la figure 3.7b. Donc, la structure des transistors NMOS est légèrement différente de celle des PMOS.

Avec l'aide des structures 1D, nous avons simulé les profils d'implantation des dopants. L'objectif est de déduire la concentration des dopants en fonction de la profondeur. Les simulations ont été effectuées sur les deux types de structures des transistors NMOS et PMOS. L'épaisseur initiale de STI est de l'ordre de 3500Å. Toutes les simulations sont faites avec des plaques ayant les mêmes épaisseurs de STI. En faisant varier la hauteur de marche, on obtiendra 6 différentes valeurs d'épaisseurs de STI. Le procédé de gravure humide permet d'enlever de l'oxyde dans la tranchée. Avec différents temps de gravure on obtient différentes valeurs de hauteurs de marche. Les épaisseurs STI des différentes plaques sont résumées dans le tableau 3.1.

Plaque	X_0 (Å)
1	3400
2	3174
3	2978
4	2722
5	2496
6	2270

Tableau 3.1- Valeurs des épaisseurs STI pour les six plaques

3.2.1.2.c Profils de dopage

Pour la réalisation des caissons dopés N ou P, certains éléments chimiques sont sélectionnés suivant des critères. Les dopages N usuels utilisent le phosphore (P) plus favorable pour la solubilité limite et le coefficient de diffusion. Toutefois l'arsenic (As) s'utilise aussi et ce, pour sa bonne solubilité limite et son faible coefficient de diffusion pour mieux contrôler les profondeurs des jonctions peu profondes. Les dopages P utilisent exclusivement le bore (B). Lui seul, présente une solubilité limite suffisante pour atteindre les concentrations requises [Deuneville 2002]. Dans cette étude, les éléments employés pour le dopage des caissons N well et P well sont le phosphore et le bore respectivement. Toutefois pour l'ajustement de la tension de seuil des transistors NMOS et PMOS, l'indium et l'arsenic sont employés. Nous présenterons les profils de dopage pour les deux types de transistors étudiés.

Caisson de type n (Nwell)

L'évolution des concentrations des espèces dopantes en fonction de la profondeur pour les transistors PMOS est illustrée sur la figure 3.8.

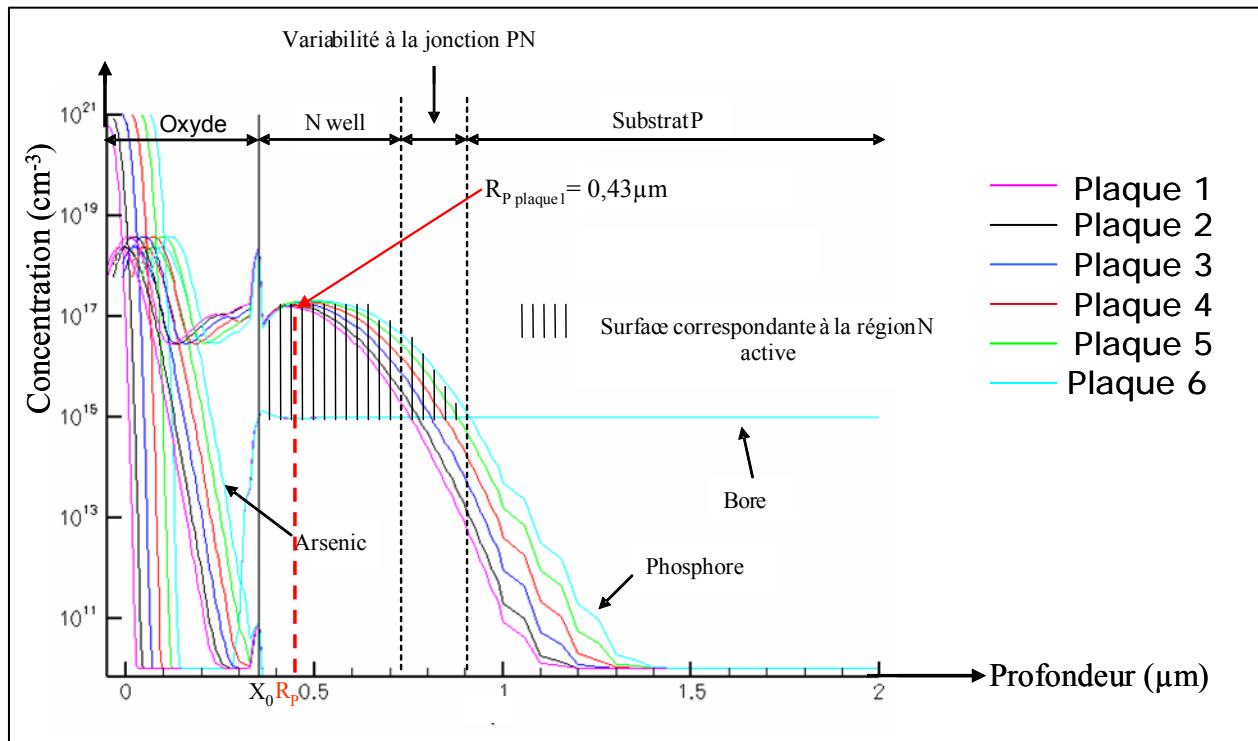


Figure 3.8- Profil des dopants dans le Nwell STI

On distingue sur la figure 3.8, les six profils de dopage. Nous pouvons distinguer sur ce graphe la zone hachurée qui correspond à la couche de type n enterrée sous le STI. Cette zone s'étend de l'interface Si/STI jusqu'au point d'intersection entre le profil de bore et celui de phosphore (jonction pn). D'après les six profils, cette zone est plus importante quand X_0 (épaisseur STI) décroît. Cet effet influence directement la valeur de la résistance carrée, car la résistance carrée est inversement proportionnelle à la concentration des dopants. Nous pouvons conclure que l'augmentation de l'épaisseur du STI induit directement une augmentation de la résistance Nwell STI. Ainsi, nous pouvons expliquer les fortes corrélations observées entre la résistance N well STI et la profondeur du STI.

Caisson de type p (Pwell)

La figure 3.9 illustre le profil des concentrations des espèces en fonction de l'épaisseur du STI dans le cas des transistors PMOS. Comme nous pouvons le constater sur la figure 3.9, la zone active de type P est représentée par la zone hachurée. Cette zone s'étend de l'interface Si/STI jusqu'au point d'intersection entre les profils de concentration du bore et du phosphore (jonction pn).

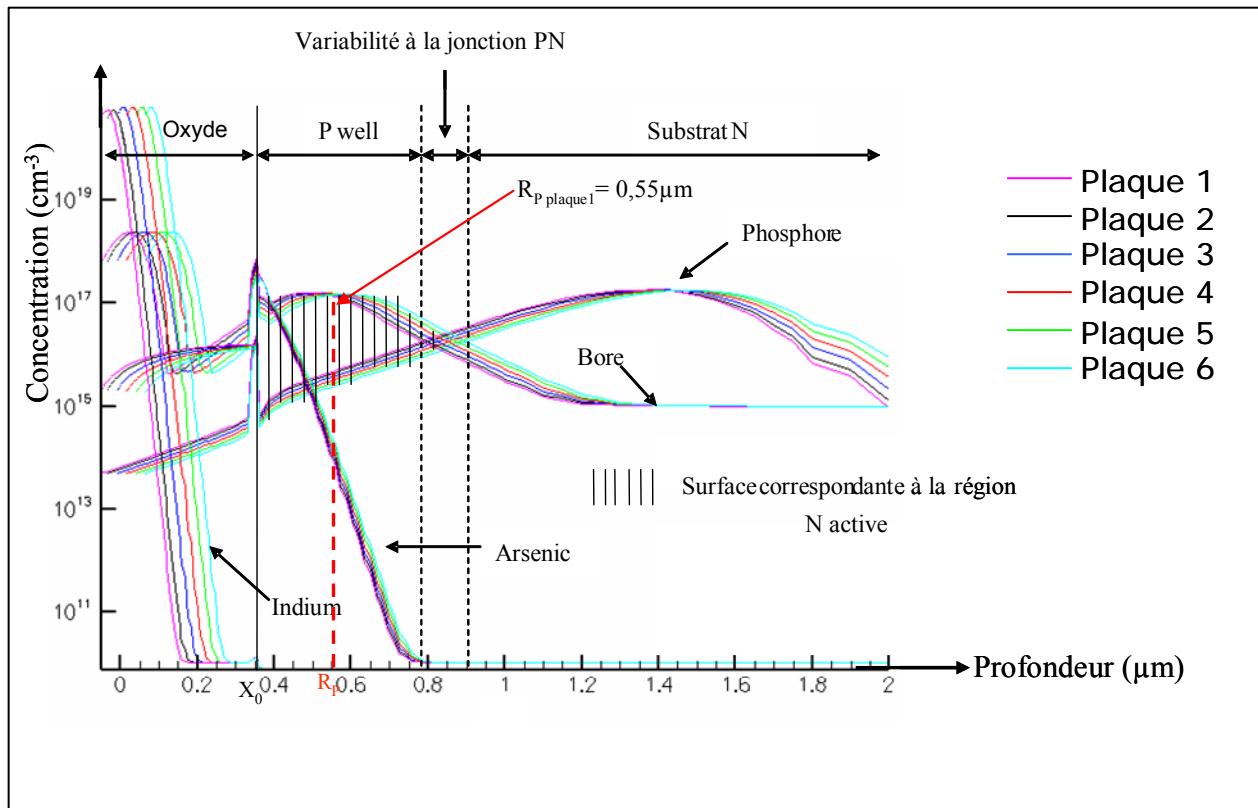


Figure 3.9- Profil des dopants dans le Pwell STI

Effet de l'énergie d'implantation

En comparant les figures 3.8 et 3.9, l'allure des profils de phosphore et de bore est presque la même: cette allure est marquée par un pic dans la région active, ce pic correspond à la concentration maximale des dopants. Cette forme de distribution des dopants est de type gaussienne. Cela veut dire que le profil de dopage est centré à une profondeur R_p avec une largeur ΔR_p et une amplitude maximum C_p en R_p . La position de R_p dans les figures 3.8 et 3.9 est différente. Nous observons que dans le Nwell, le pic d'implantation est près de l'interface Si/STI alors que dans le cas des Pwell, ce pic est loin de l'interface Si/STI. Par exemple, la plaque ayant une épaisseur STI de 3400Å, le pic d'implantation est à 0,09μm de l'interface Si/STI dans le cas de Nwell alors que dans le cas de Pwell, le pic d'implantation est situé à 0,21μm de l'interface Si/STI.

La figure 3.10 illustre les résultats de simulations de la résistance carrée de la couche implantée sous le STI.

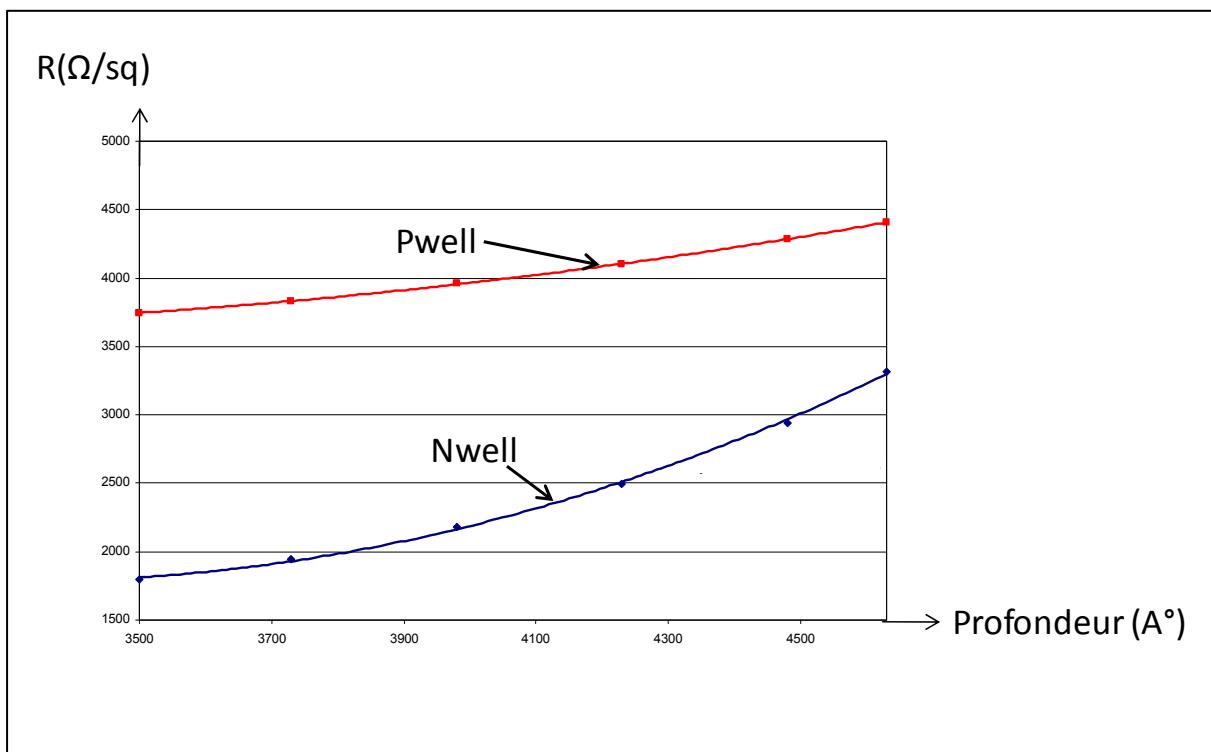


Figure 3.10- Simulations de la résistance carrée de la couche implantée sous le STI.

Nous observons sur la figure 3.10 que la pente des deux équations est différente pour les deux types de transistors. Cette différence de pente explique la différence de l'évolution des résistances N well STI et P well STI avec l'épaisseur STI. Nous confirmons donc la bonne linéarité de Nwell.

Pour l'implantation du Nwell et du P well, Les énergies d'implantation utilisées sont de 440 et 220 keV respectivement. En vue de bien comprendre l'effet de l'énergie d'implantation sur la position du pic d'implantation, nous avons effectué des simulations du procédé d'implantation du N well sur les six plaques ayant différentes épaisseurs STI, avec deux énergies d'implantation de 440 et 600 keV. Les résultats des simulations sont illustrés sur la figure 3.11.

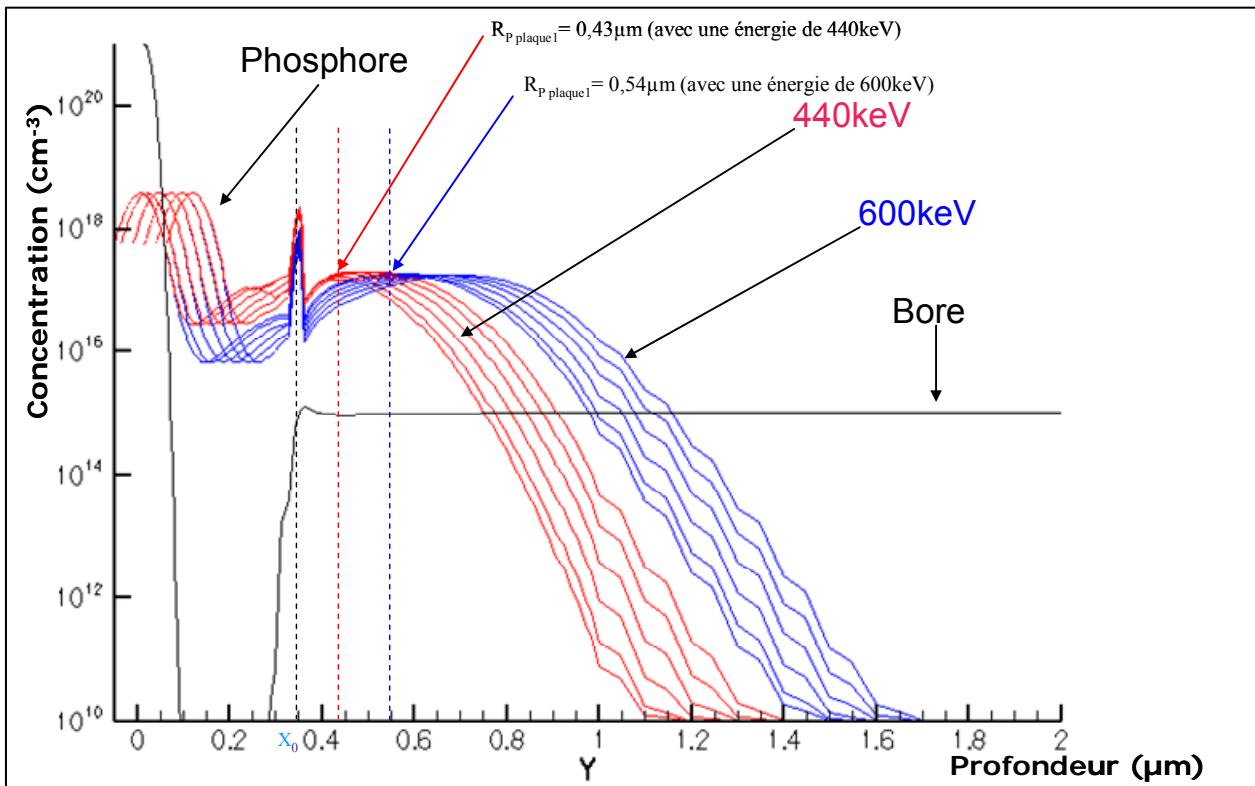


Figure 3.11- Effet de translation du pic d'implantation avec des énergies différentes lors d'implantation de Nwell

Les résultats des simulations montrent en effet que le pic d'implantation est translaté quand on passe d'une énergie d'implantation de 440 à 600 keV. Par exemple, pour la plaque ayant comme épaisseur STI de 3400 Å, le pic d'implantation se trouve à 0,09 µm de l'interface Si/STI quand on implante avec une énergie d'implantation de 440 keV. Quand cette dernière passe à 600 keV, le pic se trouve à 0,2 µm de l'interface Si/STI.

Nous avons procédé à des simulations pour calculer la dose implantée sous le STI. Nous avons étudié 3 cas dont deux cas pour le Nwell (deux énergies d'implantations 440 et 600keV) et un cas pour le Pwell (l'énergie d'implantation est de 220keV). La figure 3.12 illustre les résultats de calcul.

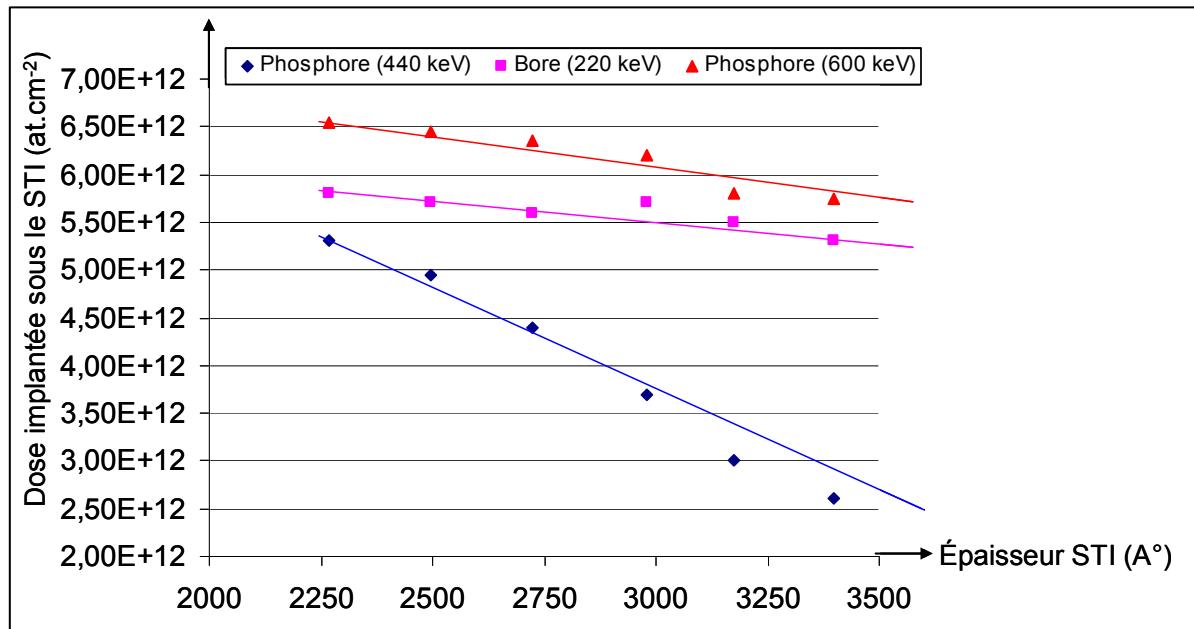


Figure 3.12- Calcul de l'intégrale des concentrations des atomes après implantation d'un Nwell et d'un Pwell avec différentes énergies d'implantation

Nous observons sur le graphe que le fait de passer d'une énergie d'implantation de 440 keV à 600 keV modifie la pente de l'équation donnant la concentration des dopants en fonction de la profondeur. La pente de l'équation de Pwell est presque la même que celle de N well (à 600 keV). De ce fait, Nous pouvons observer une bonne linéarité sur la résistance P well, si on modifie l'énergie d'implantation.

Nous venons de démontrer que la linéarité de la Résistance Nwell STI en fonction de l'épaisseur STI est expliquée par le fait que le pic d'implantation est près de l'interface Si/SiO₂. Plus on s'éloigne de cette interface, moins la forme de la réponse est linéaire. Les très bonnes corrélations trouvées permettent de surveiller la résistance NWELL STI. Il est possible de surveiller aussi la résistance Pwell STI en variant l'énergie d'implantation de sorte à avoir une linéarité.

3.2.2 Principales caractéristiques des transistors

Nous allons étudier à présent les principales caractéristiques de transistors test et en particulier le courant I_{DS} et la tension de seuil des transistors parasites.

3.2.2.1 Le courant I_{DS}

Plusieurs structures de test ont été employées pour l'analyse du paramètre Ids. Des transistors avec différentes dimensions, notamment des largeurs différentes, ont été utilisés dans les deux types de transistors suivis à savoir les NMOS et PMOS ($W=0,25\mu m$ $L=10\mu m$). La figure 3.13 illustre l'évolution de Ids en fonction de la hauteur de marche.

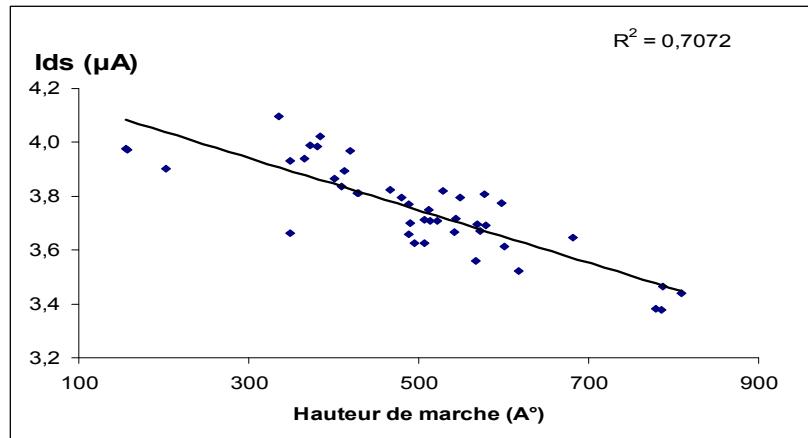


Figure 3.13- Evolution du courant Ids en fonction de la hauteur de marche pour un V_{GS} de 0.5V.

Nous constatons une forte corrélation. Cette forte corrélation peut être expliquée d'un point de vue physique. Nous pouvons expliquer cela par l'effet de la largeur du transistor (W) comme nous pouvons le vérifier avec la relation suivante:

$$I_{DSlin} = \frac{W}{L} \cdot \mu \cdot C_{ox} \cdot (V_{GS} - V_{th} - \frac{1}{2}V_{DS}) \cdot V_{DS} \quad (3.6)$$

La figure 3.14 illustre la vue de profil d'un transistor. Nous distinguons les dimensions du transistor et en particulier la largeur dessinée du transistor symbolisée par (W_0). Pour une valeur fixe (dessinée) de STI. Si la valeur de la hauteur de marche dépasse de quelques nm, ceci impactera directement la largeur comme cela est illustré sur la figure 3.14. La valeur finale de la largeur est (W_1).

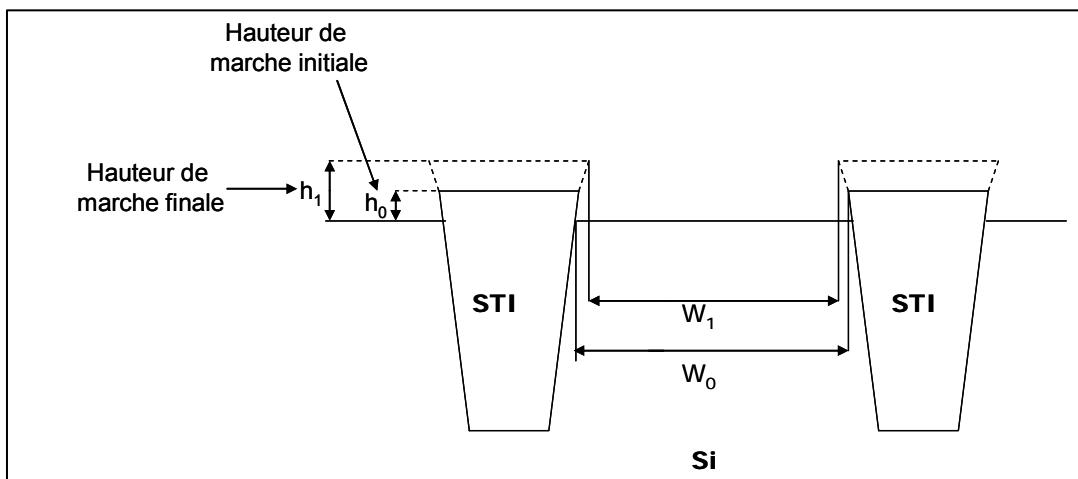


Figure 3.14- Schématisation du transistor

Comme nous pouvons le constater, les variations induites par la hauteur de marche influent directement sur la largeur effective (W_1) des transistors (effet géométrique).

$$W_1 = W_0 + \Delta W \Rightarrow \frac{\Delta W_1}{W_1} = \frac{\Delta W}{W_1} \quad (3.7)$$

3.2.2.2 Tension de seuil des transistors parasites

La deuxième caractéristique des transistors étudiés est la tension de seuil V_t qui est définie comme étant la tension de grille qui initie l'apparition de la couche d'inversion à l'interface (précisément du côté de la source).

3.2.2.2.a Structure de test

Pour l'étude de la tension de seuil V_t , nous avons employé une structure de test. Cette structure est illustrée sur la figure 3.15. Nous distinguons sur la figure un transistor MOS dont l'épaisseur d'oxyde de grille est l'épaisseur d'oxyde STI. Cette dernière sera variée pour voir son effet sur la tension de seuil.

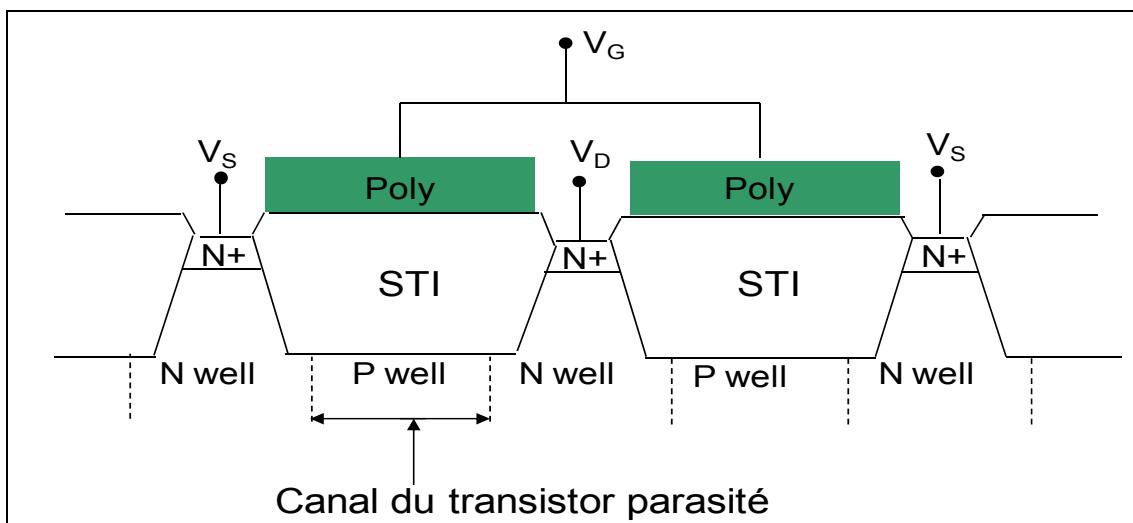


Figure 3.15- Structure de test conçue pour le monitoring de la tension de seuil V_t

3.2.2.2.b Analyses des corrélations

La figure 3.16 représente l'évolution de la tension de seuil en fonction de l'épaisseur STI et de la hauteur de marche STI. Il apparaît clairement dans les deux cas de très bonnes corrélations.

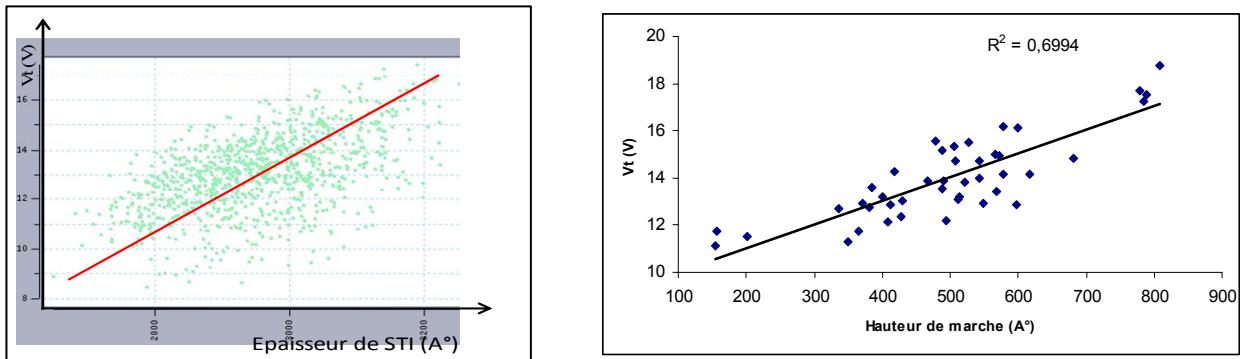


Figure 3.16- Evolution de la tension de seuil en fonction : a) de l'épaisseur STI b) de la hauteur de marche STI

C'est le seul paramètre qui présente une grande variabilité par rapport au STI. Comme nous pouvons le constater sur la figure 3.16-b, quand la hauteur de marche varie de 150 Å à 850 Å, on observe une variation de 65% de la tension de seuil. Pour comprendre cette forte variation ou sensibilité par rapport au STI, on a effectué des caractérisations électriques sur certaines plaques.

3.2.2. c Caractérisation

Des plaques avec différentes hauteurs de marche ont été analysées au niveau des tests électriques. Les plaques ont été sélectionnées sur la base de différentes hauteurs de marche (216 Å, 500Å et 791 Å). Le but des caractérisations est d'en ressortir une caractéristique classique de fonctionnement des transistors à savoir le courant de drain en fonction de la tension de grille. La tension de grille varie de 0V à 30 V, les tensions du substrat et de la source sont nulles et la tension de drain varie de 0,1 à 20V. la tension de seuil est déduite à

partir de la relation suivante :

$$I_{DSlin} = \beta \cdot (V_{GS} - V_t - \frac{1}{2}V_{DS}) \cdot V_{DS}$$

La figure 3.17 donne les résultats de la caractéristique $Id(V_G)$ pour les 3 plaques. On observe que la tension de seuil augmente avec l'augmentation de la hauteur de marche.

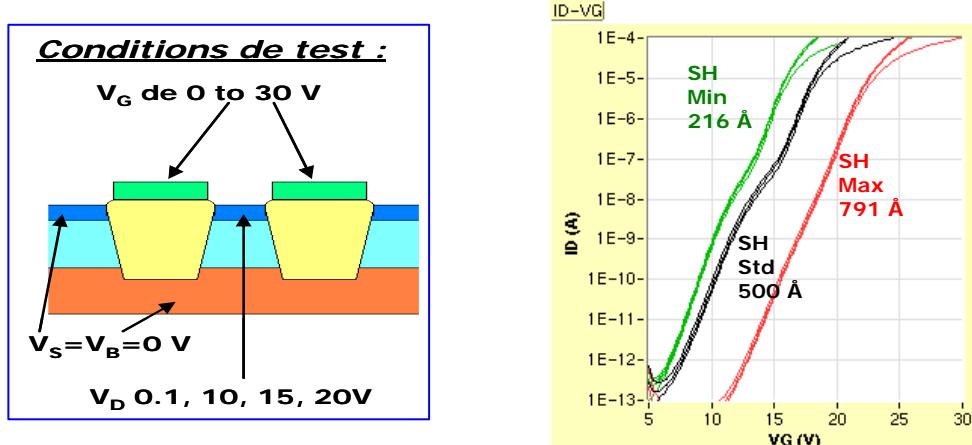


Figure 3.17- a)Conditions de test (transistor parasite Nwell) b) La caractéristique $Id(V_g)$ pour 3 hauteurs de marche.

Il est évident qu'en appliquant une tension positive sur la grille (dans notre cas V_{GS} varie de 0 à 30 V), on attire les électrons à l'interface isolant- semiconducteur (SiO_2/Si) et on

repousse les trous. A partir d'une certaine valeur de la tension de seuil V_t , une couche d'inversion apparaît et le transistor devient de plus en plus passant. Donc la valeur de la tension de seuil (V_t) dépend fortement de la densité de dopage à l'interface Si/SiO₂. Le changement de la tension de seuil en fonction de la hauteur de marche s'explique donc par la variation de dopage à l'interface Si/SiO₂. On conclut que l'augmentation de l'épaisseur d'oxyde STI induit une augmentation de la densité de dopage au niveau de l'interface Si/SiO₂ ainsi qu'une augmentation de la tension de seuil. Pour confirmer ce résultat, nous avons simulé la densité de dopage du substrat à l'interface Si/SiO₂ en fonction de l'épaisseur STI pour 3 plaques ayant des hauteurs de marche différentes. Les résultats de la densité de dopage sont illustrés dans le tableau 3.2.

plaques	Hauteur de marche (Å)	Épaisseur du STI (Å)	Dopage du substrat (cm ⁻³) à l'interface Si/SiO ₂
11	500	4132	9.34E+16
12	791	4519	1.01E+17
13	216	3949	9.11E+16

Tableau 3.2- Densité de dopage du substrat en fonction de l'épaisseur ST

3.2.3 Indicateurs électriques pour le suivi des boucles de régulations R2R du module STI

Les corrélations exposées dans les paragraphes précédents ont été confirmées par l'analyse d'autres lots de production. Il apparaît bien que la tension de seuil des transistors parasites est le paramètre le plus important à prendre en considération, car sur la plage des valeurs des paramètres STI étudiés, c'est le seul paramètre qui présente une grande variabilité sur cette plage (environ 65%). Il sera donc le seul indicateur de suivi des boucles de régulation R2R qu'on va décrire dans le chapitre 4.

3.3 Corrélation entre le STI et les contraintes mécaniques

Dans le deuxième chapitre, nous avons cité certains travaux de recherche qui ont été effectués pour démontrer la criticité des contraintes mécaniques générées au niveau du module STI sur le rendement électrique des circuits. Dans les prochains paragraphes, nous allons détailler notre approche pour quantifier les contraintes mécaniques générées par le STI quand la distance entre ce dernier et la grille est réduite. Nous allons enrichir cette approche par différents résultats expérimentaux et des simulations.

3.3.1 Rappel sur l'effet des contraintes STI

Tout d'abord, nous allons rappeler les notions fondamentales concernant les contraintes mécaniques. Lorsque nous appliquons une contrainte mécanique sur un matériau, celui-ci se déforme. C'est la loi de Hooke qui décrit le mieux la relation élastique entre la contrainte mécanique et la déformation que le matériau va subir. Les lois de l'élasticité décrivent les variations des composantes de la déformation [ϵ_{ij}] lorsqu'une contrainte [σ_{ij}] est

appliquée. Les contraintes appliquées sur le matériau sont notées σ_{ij} où i indique la direction de la force et j la normale au plan sur lequel elle s'applique comme cela est illustré sur la figure 3.18.

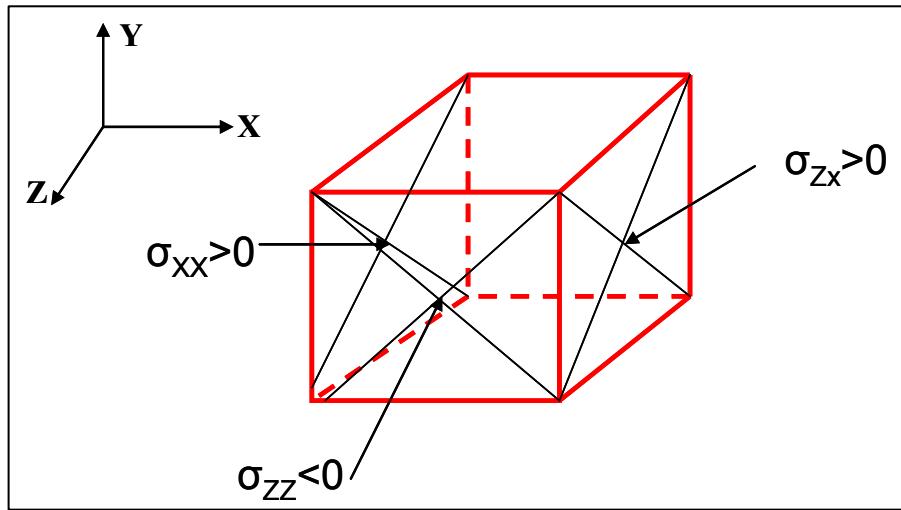


Figure 3.18- Direction et plan associé à une contrainte

Les coefficients de déformation sont reliés aux contraintes par la matrice d'élasticité (loi de Hooke valable dans le domaine linéaire des petites déformations):

$$\begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{vmatrix} = \begin{vmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{vmatrix} \begin{vmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{yz} \\ \epsilon_{zx} \\ \epsilon_{xy} \end{vmatrix} \quad (3.8)$$

Les C_{ij} sont les modules d'élasticité propres au matériau considéré. Cette matrice représente une propriété physique du matériau, elle se simplifiera donc en fonction de la symétrie de son cristal. Pour un cristal cubique, elle se réduit à:

$$\begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{vmatrix} = \begin{vmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{14} \end{vmatrix} \begin{vmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{yz} \\ \epsilon_{zx} \\ \epsilon_{xy} \end{vmatrix} \quad (3.9)$$

Il est dès lors possible, connaissant la contrainte appliquée, d'inverser cette matrice pour calculer la déformation subie par le cristal.

3.3.2 Cristal soumis aux contraintes de l'isolement STI

Nous avons mentionné dans le chapitre II que les contraintes mécaniques au niveau de la brique STI sont dues principalement aux phénomènes d'oxydation et de dilatation. D'après

la géométrie de la plaquette et la forme des cuvettes d'isolement, ces contraintes sont principalement normales à l'interface zone active/STI.

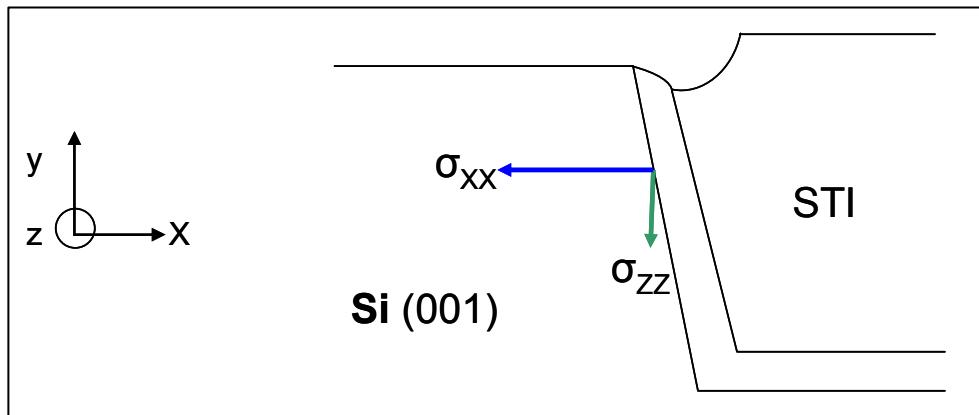


Figure 3.19- Contraintes d'isolement normales aux interfaces zone active/STI

Ortolland [ORTOLLAND 2006] a montré que les contraintes augmentent dans la direction X et Z lorsque nous avons des dispositifs ayant une longueur de grille plus petite avec une largeur constante. Les contraintes mécaniques sont donc très critiques pour les technologies avancées.

3.3.3 Effet sur la mobilité des porteurs

Lors de la réalisation des dispositifs MOS, des contraintes mécaniques non intentionnelles sont générées tout au long des étapes de fabrication [Ferreira 1996, Senez 2001]. Le STI qui est le premier module réalisé lors de la réalisation des circuits est la première source de génération de contraintes [HUANG 1996]. Smith découvre en 1954 que l'application d'une contrainte mécanique à un semi-conducteur en modifie sa résistivité électrique [Smith1954]. Cet effet, connu sous le nom de piezorésistivité trouve son origine dans la modification de la structure de bande du semi-conducteur lors de l'application d'une contrainte. En effet, l'application d'un champ de déformation induit dans le silicium une variation des distances inter atomiques et par la même une modification de ses propriétés physiques intrinsèques qui sont représentées souvent par les diagrammes de bandes d'énergie. La figure 3.20 montre dans le cas du silicium deux bandes de valences et deux bandes de conduction en l'absence de contrainte puis sous effort mécanique. La figure 3.20 est une représentation réduite de la structure de bande complète. Elle montre pour le silicium deux des six bandes de conduction et de valence dans le cas d'un réseau non contraint puis lorsqu'il est contraint par un champ de contrainte quelconque.

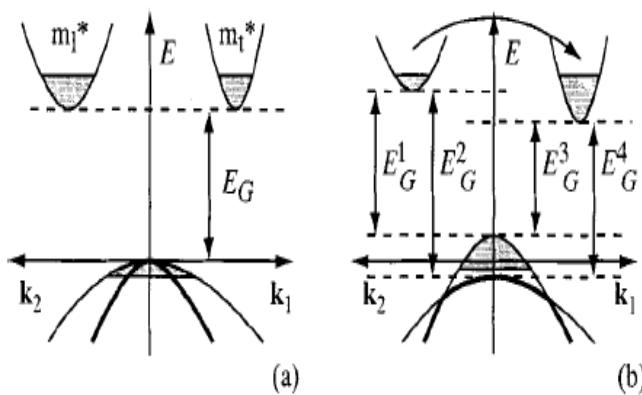


Figure 3.20- Diagramme des bandes d'énergie a) sans contraintes b) sous contraintes mécaniques

Il apparaît que lorsqu'une contrainte mécanique est appliquée au cristal, la structure de bande est modifiée. Trois modifications principales sont à relever :

- la largeur de la bande interdite⁴⁰ (E_g) est modifiée.
- les minima de bande de conduction sont modifiés, ce qui induit une répartition différente des électrons dans les deux vallées⁴¹.
- il apparaît une levée de dégénérescence dans les bandes de valence. En d'autres termes, leurs maxima se séparent. Par ailleurs on note aussi un changement dans la courbure de ces bandes.

Ces modifications conduisent à une variation de la mobilité apparente des porteurs de charge (électrons et trous), en particulier à travers une variation de leur masse effective [REGGIANI 1985].

Nous allons maintenant nous pencher sur les conséquences de la déformation des bandes sur le transport des porteurs⁴². L'impact principal d'un tel effet est la variation de la mobilité des porteurs de charge qui modifie directement les propriétés macroscopiques des transistors. En effet, lorsque l'on trace la caractéristique I_d (V_g) d'un transistor, on peut extraire la mobilité en étudiant la zone du régime linéaire selon l'expression suivante :

$$M_o = 0,5 * \left(\frac{W}{L} \right) * C_{ox} * \mu_0 \quad (3.10)$$

Où : C_{ox} est la capacité d'oxyde

⁴⁰ La largeur de bande interdite E_g (le gap) d'un semi-conducteur est définie comme l'écart entre le minimum absolu des bandes de conduction et le maximum absolu des bandes de valences.

⁴¹ Le silicium étant un cristal cubique dans les directions $\langle 100 \rangle$, $\langle 010 \rangle$, $\langle 001 \rangle$, $\langle \bar{1}00 \rangle$, $\langle 0\bar{1}0 \rangle$ et $\langle 00\bar{1} \rangle$ sont équivalentes. Ce qui nous donne 6 minima équivalents, également appelé vallées.

⁴² Pour les électrons, le modèle proposé par Herring et Voigt pour expliquer les propriétés anisotropiques des cristaux à structure électronique multivallées est basé sur le mécanisme de transfert d'électrons entre les différentes vallées. La mobilité d'un électron μ_e sous un faible champ électrique et appartenant à une vallée particulière de la bande de conduction peut s'exprimer : $\mu_e = \frac{q\tau}{m^*}$. Où m^* est la masse effective des électrons dans la direction du transport et τ le temps de relaxation des porteurs et q la charge de l'électron.

μ_0 est la mobilité des porteurs

W et L sont les dimensions du transistor

M_0 est la pente de la région linéaire de la caractéristique $I_d(V_G)$.

Dans le cas où la variation de la tension de seuil est négligeable, la variation de la mobilité entraîne directement la variation du courant de drain selon l'équation :

$$\frac{I_d(\sigma) - I_d(0)}{I_d(0)} = \frac{\Delta I_d}{I_d} = \frac{\Delta \mu}{\mu} \quad (3.11)$$

Où $I_d(\sigma)$ est la valeur du courant I_d pour une plaque contrainte

$I_d(0)$ est la valeur du courant I_d pour une plaque non contrainte.

3. 3.4 Influence de la distance entre la grille et le STI

L'effet des contraintes mécaniques est illustré sur la figure 3.21. Nous remarquons sur la figure comment la distance « a » (distance STI- grille) joue un grand rôle pour la détermination des niveaux de contraintes.

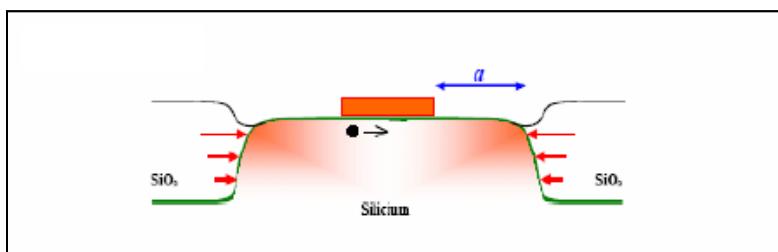


Figure 3.21- Effet des contraintes du STI sur les parties actives

Nous allons voir dans les prochains paragraphes l'approche proposée pour modéliser l'effet de cette distance.

3.3.4.1 Mesures électriques

Nous avons vu l'importance des deux dimensions principales, L et W dans la partie précédente consacrée aux corrélations STI - paramètres électriques. Cependant, un transistor et encore plus un circuit ne se résume pas à ces deux dimensions. C'est pour cela qu'il est nécessaire de se pencher sur les autres paramètres capables de moduler cet effet de transmission des contraintes mécaniques [ORTOLLAND 2006]. Plusieurs paramètres sont susceptibles de modifier le niveau des contraintes mécaniques parmi lesquels:

- la hauteur de grille
- la distance entre deux grilles
- la distance grille - STI
- les contacts

D'une génération à une autre, toutes ces dimensions vont être diminuées. Ce qui va se traduire d'un point de vue mécanique par une modification du niveau de contraintes dans le canal. C'est l'effet de la distance entre la grille et le STI que nous allons montrer. Pour accomplir cette étude, une structure de test a été utilisée pour surveiller les contraintes. La structure de test est un réseau de transistors permettant de placer le transistor à mesurer dans

une structure dense représentative d'un environnement produit comme cela est illustré dans la figure 3.22.

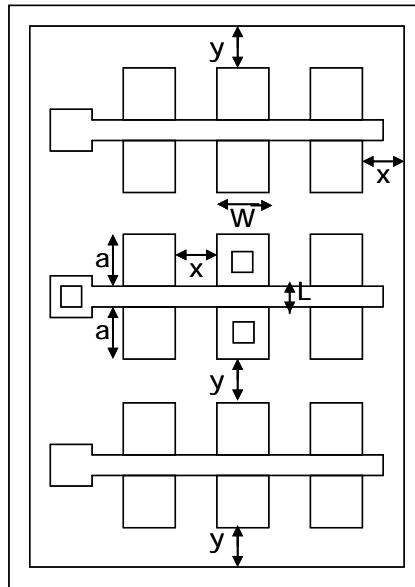


Figure 3.22- Structure de test pour surveiller le stress

Comme illustré sur la figure 3.22, les valeurs des différentes dimensions sont:

- $X=Y$ (espacement entre les parties actives) = $0,32\mu\text{m}$.
- L (longueur des transistors) = $0,18\mu\text{m}$
- W (largeur des transistors) = $0,28\mu\text{m}$ pour les transistors étroits, et pour les transistors larges $W=10\mu\text{m}$
- « a » varie de $0,4$ à $4,9\mu\text{m}$.

Toutes les plaques ont été mesurées au test paramétrique (PT) final avec l'aide d'une cartographie contenant 32 points de mesures.

- Le moyen le plus simple de vérifier l'impact des contraintes du STI est de considérer des transistors ayant tous, hors contraintes, les mêmes caractéristiques électriques et de faire varier l'intensité de la déformation dans le canal de conduction en modifiant la distance entre la grille et la tranchée tout en gardant une distance constante entre les contacts et la grille (les lignes de champ électrique restent alors inchangées).

La figure 3.23 représente la variation de la mobilité en fonction de la distance « a » pour les transistors NMOS & PMOS (épaisseur de l'oxyde de grille est de 32\AA). Les plaques mesurées ont des hauteurs de marche standards de l'ordre de 450\AA .

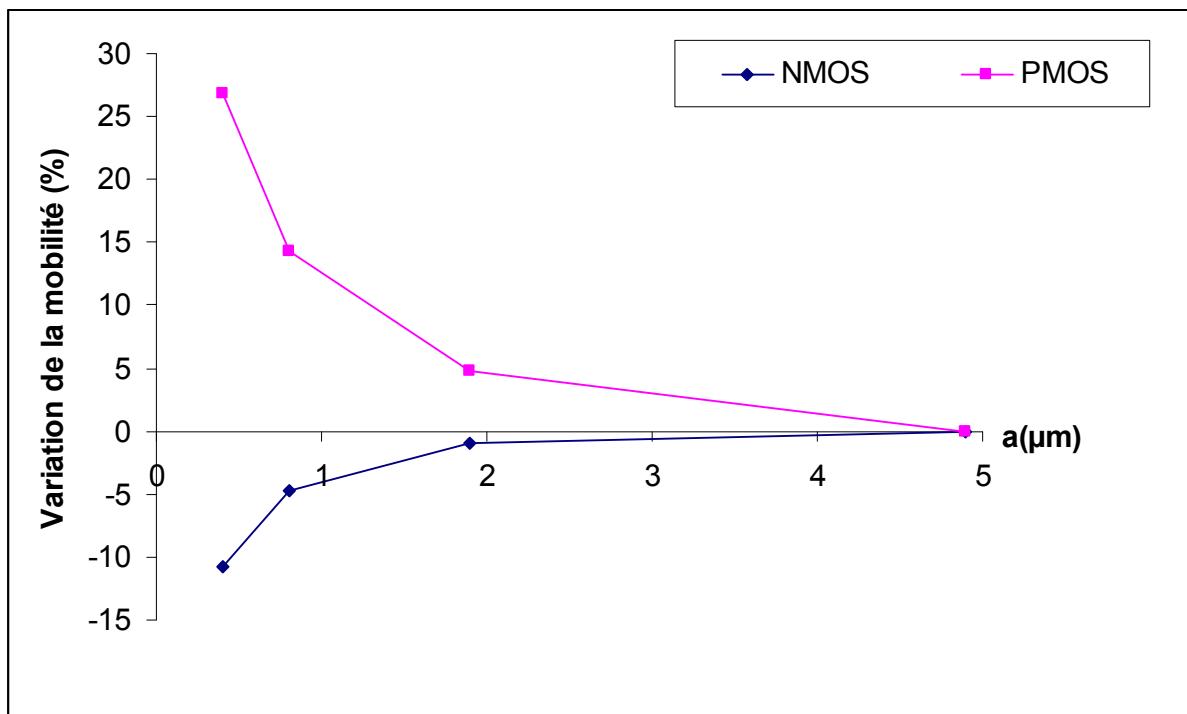


Figure 3.23- Variation de la mobilité en fonction de la distance « a » pour les transistors NMOS et PMOS dont les dimensions ($L=0,18\mu\text{m}$ et $W=0,28\mu\text{m}$) par rapport à une référence non stressée mécaniquement.

Pour une diminution de la distance « a », la mobilité des porteurs se dégrade dans le cas du NMOS alors qu'elle augmente dans le cas du PMOS. Pour une même valeur de la distance « a », la variation absolue de la mobilité est supérieure dans le cas des trous (PMOS). Pour les transistors NMOS, quand nous diminuons la distance entre la grille et la tranchée d'isolation, nous augmentons les contraintes en compression du canal de conduction dans la direction X, la valeur du courant de conduction diminue. Les performances de ce type de transistor semblent donc dégradées. Pour les transistors PMOS, plus la distance est petite, plus la compression due au STI est forte dans le canal de conduction et meilleur est le courant de conduction. Les contraintes du STI permettent donc de moduler les performances des différents types de dispositifs. Elles améliorent les performances des PMOS et réduisent celles des NMOS.

3.3.4.2 Mesures sous flexion

Pour expliquer cette différence de comportement des transistors, nous avons procédé à des expériences pour quantifier la contrainte et donc les valeurs des coefficients piezorésistifs dans le silicium dopé. On peut illustrer cet effet en imposant à un transistor une contrainte mécanique connue. Une machine de flexion 4 points [Kasbari 2007] permet de fléchir des lamelles de silicium en compression comme en tension. Cette machine est illustrée sur la figure 3.24.

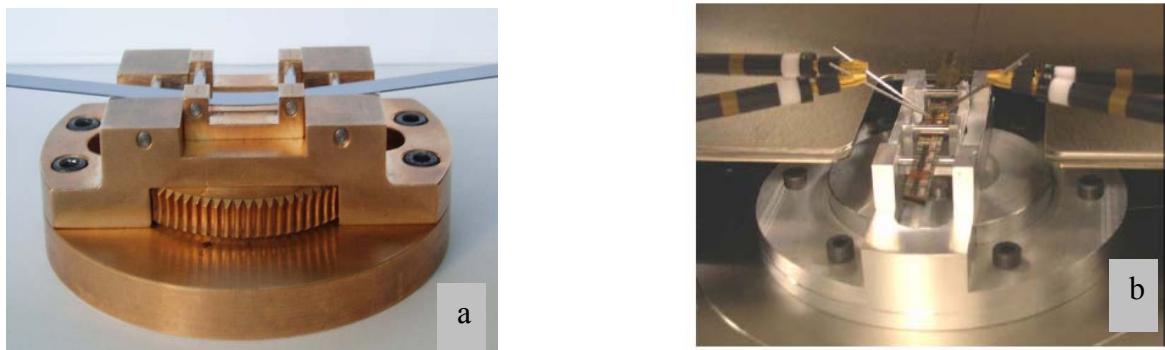


Figure 3.24- Dispositif expérimental : a) Machine de flexion b) mesures de flexion 4 points [KASBARI 2007]

La lamelle de silicium est glissée entre deux supports externes et deux supports internes mobiles. Le déplacement du chariot (cylindre interne) vers le bas courbe la lamelle induisant une contrainte en compression à la surface du silicium (fig. 3.24a). En inversant le déplacement du chariot on peut aussi mettre la lamelle de tension uni-axiale (courbure vers le haut). La flexion 4 points permet d'appliquer un moment de flexion constant entre les deux supports internes. Le transistor situé à cet endroit subit donc une contrainte (tension ou compression) uni-axiale selon la plus grande dimension de la lamelle. Dans les deux cas, la surface du silicium reste libre pour accueillir des pointes dédiées à la mesure électrique⁴³ (fig 3.24b)

3.3.4.3 Résultats des mesures sous flexion – Extraction des coefficients piezorésistifs

Les résultats des mesures sous flexion sont représentés sur la figure 3.25. Pour chaque contrainte imposée, une valeur de $I_d(\sigma)$ est extraite et comparée avec une valeur de $I_d(0)$. Ensuite, le rapport $\Delta I_d/I_d(0)$ est calculé. D'après la relation (3.11) et à partir de ce rapport, nous pouvons représenter la variation de la mobilité $[\Delta\mu/\mu(0)]$ en fonction des contraintes mécaniques imposées pour les deux types des transistors NMOS et PMOS.

⁴³ L'étude du comportement électrique des composants lors d'une analyse couplée : électrique/mécanique, impose de pouvoir solliciter mécaniquement le composant pendant la mesure électrique. La mesure doit respecter deux conditions clés :

- La mesure électrique ne doit pas être perturbée par le dispositif d'application des contraintes.
- La contrainte imposée à l'échantillon doit être calibrée et constante durant toute la phase de mesure.

Le développement, la calibration du banc de mesure et un banc de flexion 4 points est détaillé dans la thèse de M. Kasbari [KASBARI 2008].

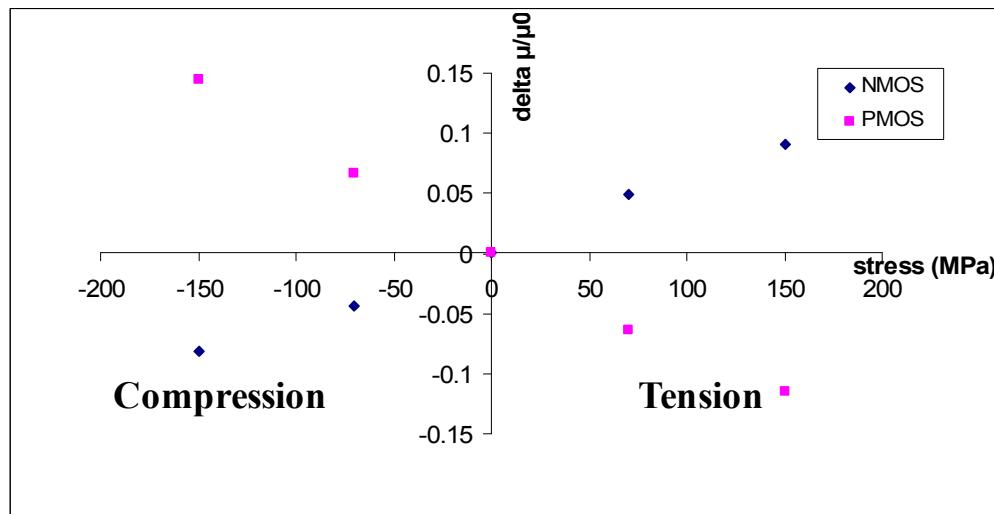


Figure 3.25- Variation de la mobilité en fonction de la contrainte appliquée

En ordonnée est représentée la variation de mobilité relative $\frac{\Delta\mu}{\mu_0}$ alors que l'axe des abscisses porte les valeurs de contraintes mécaniques. Dans la convention habituelle en microélectronique et en particulier pour l'étude des contraintes mécaniques, elles sont exprimées en mégapascal (MPa) avec un signe moins pour les contraintes en compression et un signe plus lorsqu'elles sont en tension.

L'objectif du tracé de la mobilité en fonction des contraintes est d'extraire les coefficients piezorésistifs.

Extraction des coefficients piezorésistifs

Les surfaces des plaquettes de silicium utilisées actuellement dans la filière CMOS est dans le plan réticulaire (001). L'axe principal utilisé pour la conduction du courant est l'axe [110] comme cela est illustré sur la figure 3.26.

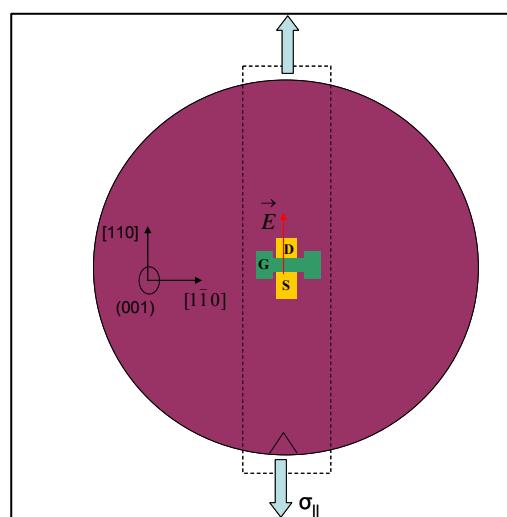


Figure 3.26- Représentation de l'orientation de champ électrique (\vec{E}) et la contrainte appliquée

$$\boxed{\text{La contrainte est calculée à partir de la relation suivante : } \sigma = -\frac{eE}{2R}} \quad (3.12)$$

Où : E est le module de Young
 e est l'épaisseur du film
 R est le rayon de courbure.

Les coefficients piezorésistifs sont exprimés le long de l'axe principal ou perpendiculairement à cet axe. On parle alors de coefficients longitudinaux et transverses notés π_L et π_T respectivement. Les coefficients piezorésistifs en général dépendent de la direction cristallographique⁴⁴, du dopage du silicium et de sa température [Kanda1982]. Le résultat présenté en figure 3.25 ainsi que la procédure d'extraction et les conditions expérimentales sont présentés plus en détails dans la thèse de M. Kasbari [KASBARI 2007]. On constate que, la variation de la mobilité des porteurs suit une fonction linéaire de la contrainte. La pente des droites donne la sensibilité du silicium dopé à la contrainte mécanique. La figure 3.25 permet d'extraire pour un type N et un type P les valeurs de ces coefficients longitudinaux suivant la relation suivante :

$$\boxed{\frac{\Delta\mu}{\mu_0} = \Pi_L * \sigma} \quad (3.13)$$

Pour les NMOS, le coefficient de piézorésistance longitudinal est: $\pi_L(N) = 600.10^{-12} Pa^{-1}$, pour les PMOS, le coefficient de piézorésistance est: $\pi_L(P) = -900.10^{-12} Pa^{-1}$.

Nous pouvons donc interpréter à présent la différence de la variation de mobilité en fonction de la distance « a » pour les transistors NMOS et PMOS par le fait de la différence des valeurs et signes de leurs coefficients piezorésistifs.

3.3.4.4 Résultats électriques

A partir des résultats donnés par la figure 3.25, reflétant la sensibilité mécanique du silicium et les tests dédiés à la surveillance de la distance « a », on peut représenter la variation de contrainte dans le canal du transistor en fonction de la distance « a ».

⁴⁴ Sur les plaques de silicium de 200mm, l'orientation cristalline d'un substrat se définit en pratique à l'aide de l'entaille, ou « notch » de la plaque. Elle va donner alors la référence pour le sens de la plaque et donc pour l'orientation des puces : la direction tangentielle au notch correspond à l'orientation de la plaque.

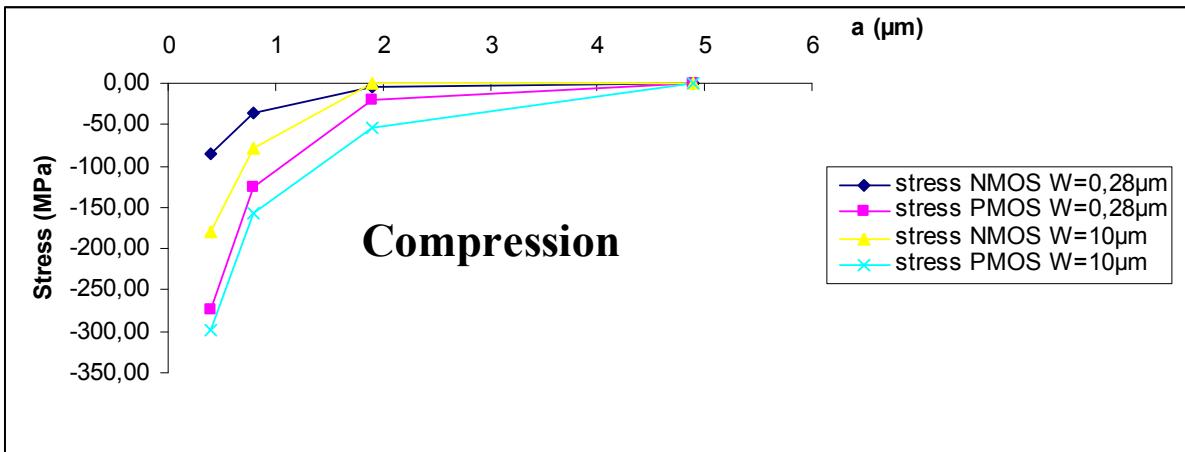


Figure 3.27- Extraction de niveaux de contraintes en fonction de la distance « a » pour les transistors NMOS & PMOS

La figure 3.27 illustre la variation des contraintes en fonction de « a ». Nous observons sur ce graphe que les contraintes sont en compression et très importantes pour de petites valeurs de « a » pour les transistors NMOS et PMOS. De plus, les contraintes dans la direction X sont importantes pour les transistors larges ($W=10 \mu\text{m}$) comparés aux transistors étroits ($W=0,28\mu\text{m}$). Pour ce type de contraintes (dans la direction X), nous nous focaliserons que sur les transistors ayant un $W=10\mu\text{m}$. Nous pouvons expliquer la différence des niveaux de contraintes observées sur la figure 3.27 par le fait que les conditions des procédés de fabrication ne sont pas les mêmes pour la réalisation des transistors (NMOS et PMOS) et principalement au niveau de l'étape de fabrication de la grille des transistors. Nous pouvons citer certains procédés comme l'implantation ionique (nous avons pu constater dans ce chapitre-partie 3.2.1.2 les différentes valeurs des énergies d'implantation ; 440 keV pour les NMOS et 220 keV pour les PMOS). Nous pouvons citer également d'autres procédés comme l'amorphisation et la diffusion. Lors de ces procédés, le comportement thermique influence la génération des contraintes mécaniques. Ces dernières dépendent donc de la température.

L'ensemble de ces procédés contribue à l'apparition des contraintes mécaniques. La différence des conditions des procédés pour la réalisation des deux types des transistors peut bien expliquer les résultats de la figure 3.27.

3.3.5 Modélisation numérique

Pour mieux modéliser les variations des contraintes en fonction de la distance « a », nous avons conçu un plan d'expériences. Ce plan d'expériences a été réalisé en simulation⁴⁵.

3.3.5.1 Programme et conditions des simulations

Nous avons utilisé le module Dios du logiciel Sentaurus proposé par la société Synopsys.

⁴⁵ On entend par simulation la simulation de procédés. L'utilisation d'un simulateur de procédés permet de reproduire fidèlement toutes les étapes de la fabrication de la brique STI (dépôt nitrure, gravure de la tranchée, dépôt CVD, CMP, gravure humide). Même si les calculs sont plus longs ils permettent d'obtenir un résultat très proche de la structure réelle.

La figure 3.28 illustre le transistor contraint. Nous allons calculer les contraintes mécaniques dans les directions (X,Y,Z).

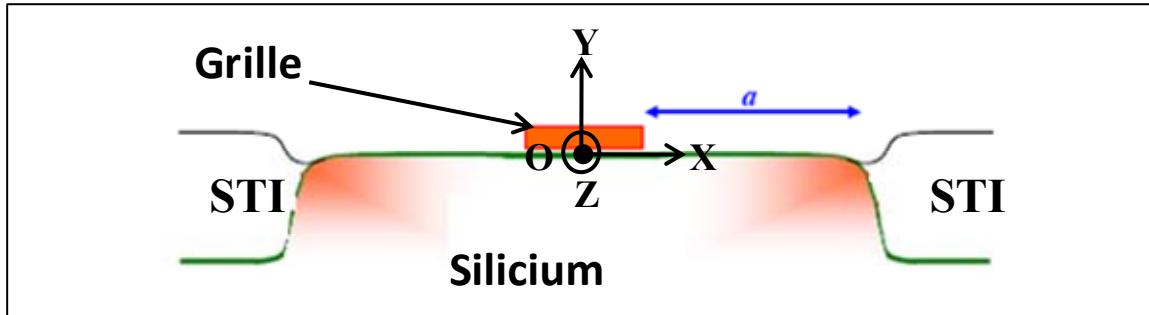


Figure 3.28- Représentation d'un transistor contraint dans les 3 directions (X,Y,Z).

Le calcul des contraintes mécaniques est mené dans le cadre de l'hypothèse des déformations planes. En effet, en supposant un barreau de silicium infiniment long (dans la direction Z), on peut restreindre le problème tridimensionnel à un calcul des déformations limité au plan. Selon l'axe (OY), la structure reste flottante (pas de limitation de coté vertical), les contraintes dans ce cas sont en tension.

Les contraintes quant à elles peuvent se développer et se calculer dans les trois directions de l'espace. Pour autant, on négligera les composantes de la contrainte selon l'axe (OY). Cette approximation est possible puisque la composante σ_{yy} est petite devant les composantes σ_{xx} et σ_{zz} et parce que l'on ne considère, dans cette étude, que l'effet piezorésitif dans le plan de la résistance (X,Z). Pour chaque simulation, la valeur des composantes σ_{xx} et σ_{zz} est extraite. Dans toutes les simulations numériques, la contrainte mécanique se développe librement dans les différents matériaux en accord avec les conditions physiques du procédé de fabrication. Si la valeur de la contrainte intrinsèque des couches est arbitrairement fixée (contrainte bi axiale isotrope dans les films), les coefficients de dilatation thermique des différents matériaux sont activés induisant des contraintes thermiques calculées en fonction des températures de dépôt.

Ces simulations ont en outre été faites en supposant un comportement purement élastique et linéaire⁴⁶ des matériaux. Pour des raisons de symétrie de structures étudiées et aussi pour le gain de temps, les simulations ont été faites sur des demi-structures comme cela est illustré sur la figure 3.29. Nous avons également choisi d'arrêter la simulation après le retrait du nitrate car nous ferons une analyse en relatif. Ce choix est justifié par le fait que ça nous permettra de s'affranchir de la contribution d'autres variables au niveau des opérations post-STI.

⁴⁶ La linéarité permet de limiter l'étude à un seul type de contrainte, en tension par exemple.

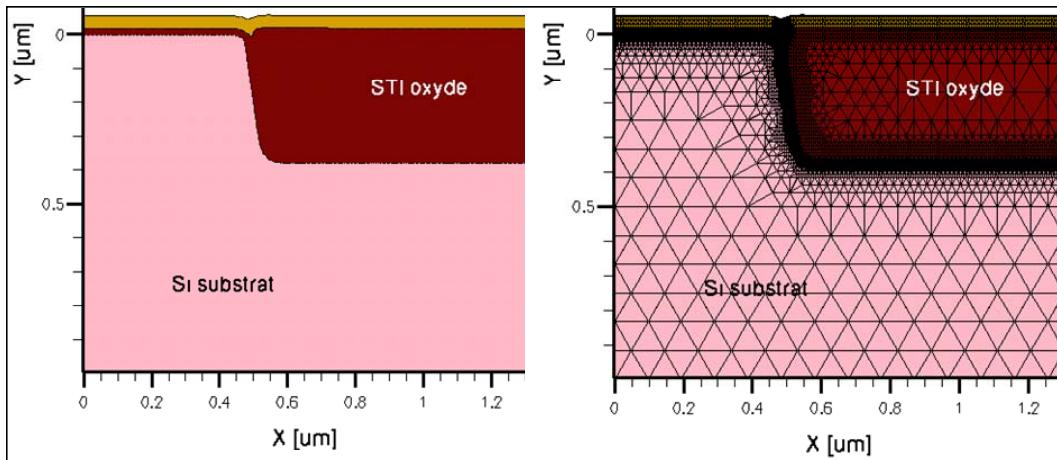


Figure 3.29- a) Structure de base pour la simulation de la contrainte mécanique.
b) utilisation du maillage adaptatif.

Les simulations permettront d'exprimer les contraintes en fonction de la distance « a ». Nous pouvons parler donc d'un modèle prédictif. Pour les technologies ultimes, ce modèle peut servir les concepteurs des circuits pour définir la dimension critique (minimale) de « a » lors de dessin des circuits.

3.3.5.2 Plan d'expériences

Nous avons simulé des demi-structures ayant des longueurs de grille de l'ordre de 0,18μm. La distance « a » varie de 0,05 à 20μm. Le tableau 3.3 résume les différentes grandeurs.

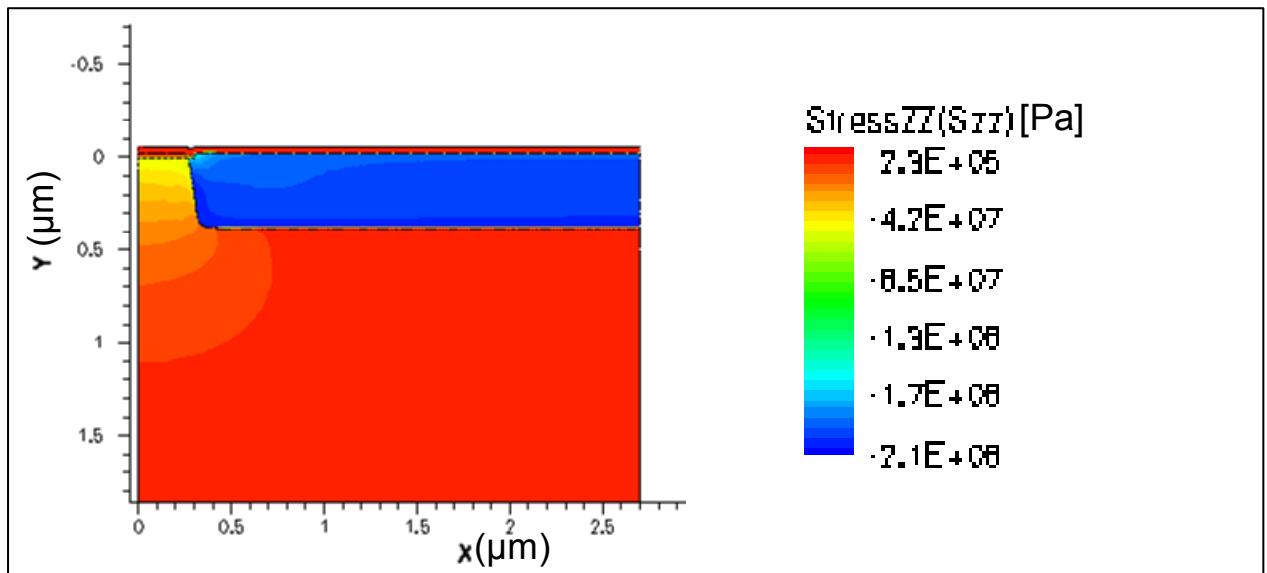
Plaques	L_G (μm)	a (μm)
1	0,18	0,05
2		0,2
3		0,4
4		0,8
5		1
6		1,5
7		2

Tableau 3.3- Données de la distance « a »

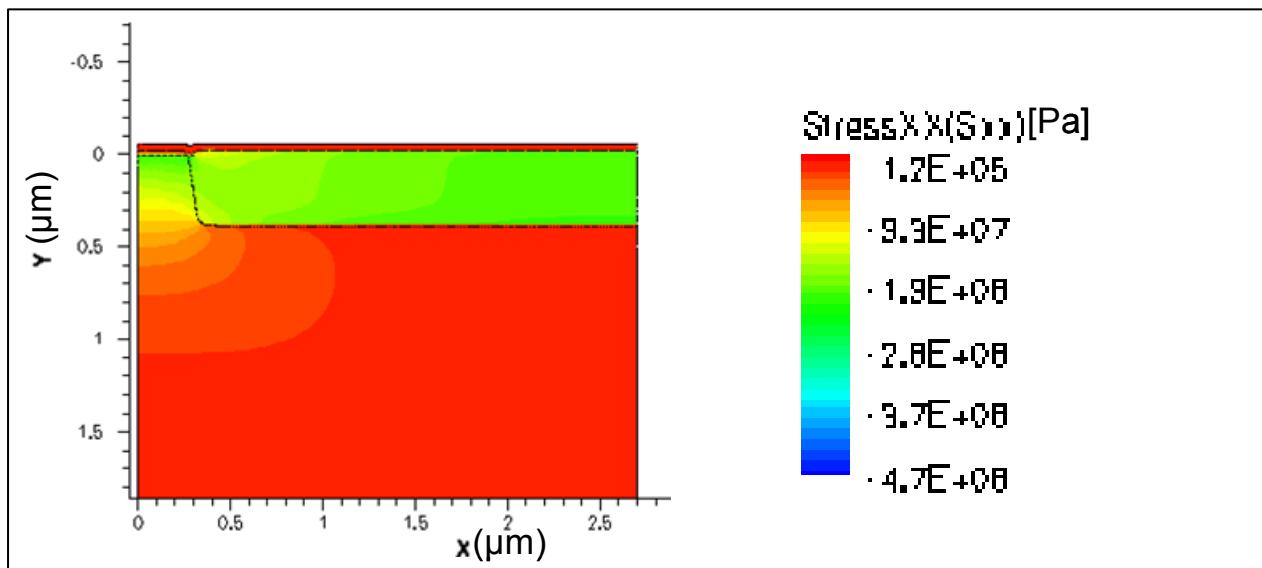
3.3.5.3 Résultats des simulations

3.3.5.3.a Champs de contraintes

Les simulations permettent d'avoir des cartographies des champs de contraintes sur tous les points des demi-structures utilisées. La figure 3.30 montre les résultats des simulations du champ de contraintes σ_{XX} et σ_{ZZ} pour une valeur de $a=0.2\mu m$ dans un plan (XOY). Les couleurs au niveau de la zone active représentent les différents niveaux de contraintes. Ce sont des lignes iso- contraintes. L'axe (X) est l'axe horizontal dans le sens de la largeur du canal, L'axe (Y) est l'axe vertical. Le point $(x,y)=(0,0)$ est le centre du canal.



a) σ_{ZZ}



b) σ_{XX}

Figure 3.30- Répartition du stress (σ_{XX} et σ_{ZZ}) dans la zone active pour une valeur de « $a=0.2\mu\text{m}$ »

Les contraintes sont maximales sur les bords du STI [RIVERO 2007]. Ces contraintes se déplacent ensuite vers le centre de la zone active comme nous pouvons le constater sur la figure 3.30. Nous confirmons à travers cet exemple que l'oxyde de la tranchée STI est en compression dans le plan (x, z).

3.3.5.3.b Profil des contraintes au niveau de la surface du canal

Nous allons faire une coupe à un niveau qui correspond à $Y=-0.04\mu\text{m}$. Le choix de ce niveau est justifié par le fait qu'il est très proche de la surface du canal.

Nous allons étudier le profil des contraintes dans la direction (X) et (Z) en fonction de la longueur (direction X du canal). La figure 3.31 montre le profil de contrainte pour différentes valeurs de « a » à une profondeur ($y=-0.04\mu m$).

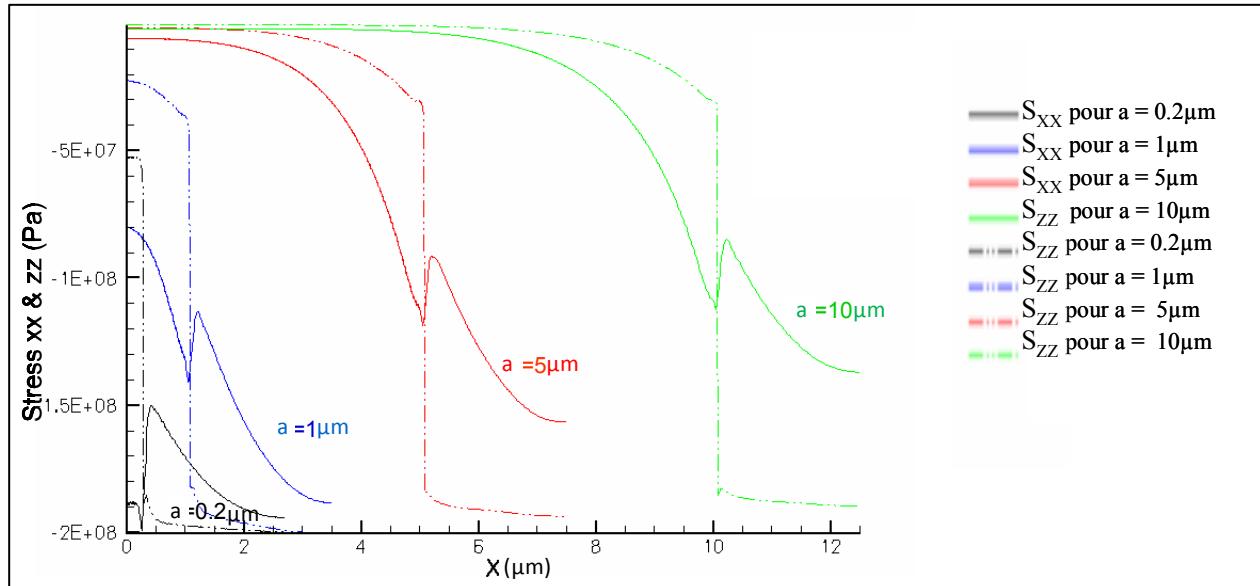


Figure 3.31- Représentation de stress en fonction de la longueur (suivant la direction x du canal) pour différentes valeurs de « a » au milieu du canal.

Comme cela est illustré sur la figure 3.31. Pour les grandes valeurs de « a », les contraintes sont très faibles dans les deux directions tandis que pour les petites valeurs de « a », le stress est important.

3.3.5.3.c Profil des contraintes au milieu du canal ($X=0$)

De la même manière que précédemment, nous allons étudier le profil des contraintes dans la direction (X) et (Z) en fonction de la profondeur de la tranchée. Nous allons faire une coupe au niveau du milieu du canal. La figure 3.32 montre le profil de contrainte selon une coupe transversale dans le sens de la largeur pour un transistor.

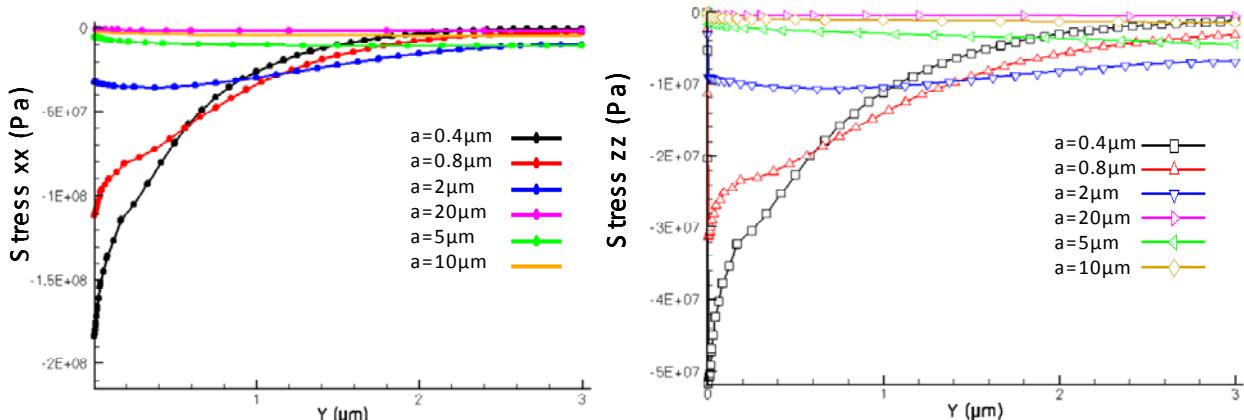


Figure 3.32- Représentation de stress en fonction de la profondeur de tranchée pour différentes valeurs de « a » au milieu du canal.

La figure 3.32 représente le niveau de contrainte pour la composante x et z en fonction de la profondeur pris en $x=0$ (centre de l'active). Comme cela est illustré sur la figure 3.32. Pour les grandes valeurs de « a », les contraintes sont très faibles dans les deux directions tandis que pour les petites valeurs de « a », le stress est important.

3.3.5.3.d Profil des contraintes au niveau de la surface ($X=0$, $y=0$)

Pour voir cette tendance sur un niveau local de la zone active, on a pris un point de référence qui est le point au milieu du centre de canal ($x=0$, $y=0$) pour analyser les niveaux de contraintes S_{XX} et S_{ZZ} .

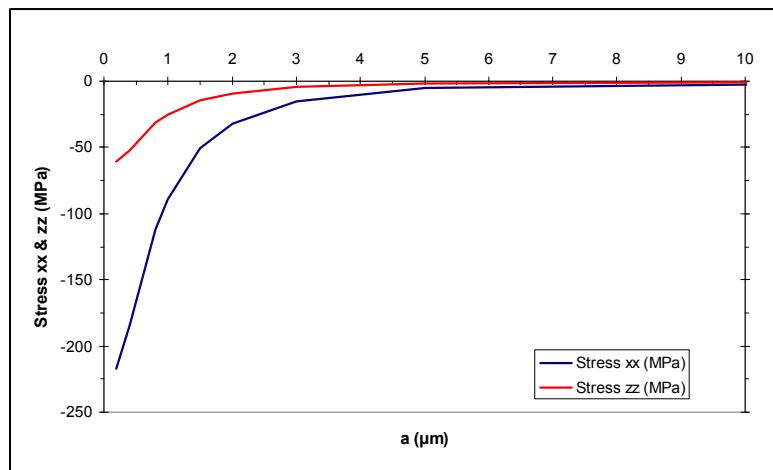


Figure 3.33- Variation du niveau des contraintes en fonction de la distance « a » pour les transistors NMOS (pour les transistors $W=10\mu m$)

Comme cela est illustré sur la figure 3.33, les contraintes mécaniques sont très importantes quand on rapproche le STI de la partie active. Nous observons également que les contraintes dans la direction X sont plus importantes que les contraintes dans la direction Z. Nous pouvons expliquer cette différence de niveaux des contraintes dans les deux directions (X) et (Z) par la taille des transistors étudiés ($W=10\mu m$).

3.3.5.3.e Modèle des contraintes en fonction de la distance « a »

Ces résultats de simulations sont en accord avec ce qui a été trouvé en expérimental (mesures sous flexion). De ce fait, à partir de la figure 3.33, nous pouvons établir deux modèles simples donnant les niveaux de contraintes S_{XX} et S_{ZZ} .

$$\text{Stress (XX)} = 248,47 \cdot \exp(-0,977a) \quad (3.14)$$

$$\text{Stress (ZZ)} = 69,74 \cdot \exp(-0,977a) \quad (3.15)$$

Nous constatons que l'argument (0,977) est le même dans les deux relations, cela veut dire que les contraintes se développent de la même façon dans les directions (X) et (Z).

La longueur de grille (L_G) des transistors employés lors des simulations est de l'ordre

de $0,18\mu\text{m}$. Si nous réduisons les dimensions des transistors et donc L_G , les contraintes mécaniques seront plus importantes [ORTOLLAND 2006].

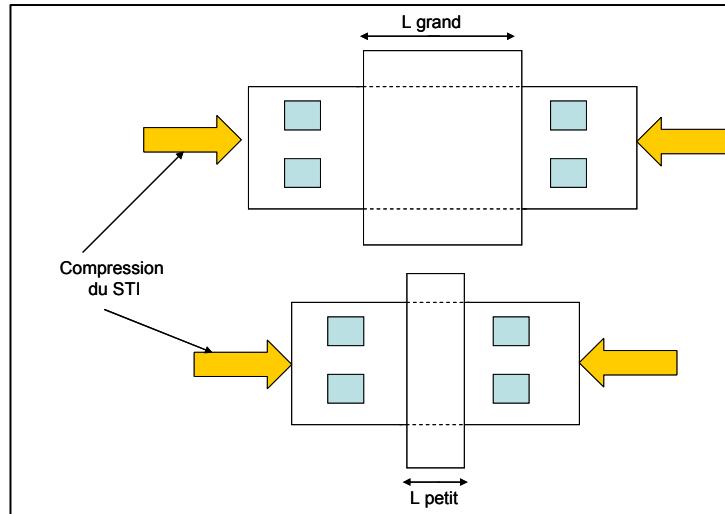


Figure 3.34- Représentation schématique de l’effet des contraintes du STI lorsque nous réduisons la longueur de grille.

Lorsque la longueur de grille est plus petite, une plus grande proportion du canal de conduction est alors déformée. Les transistors à faible L devraient donc être plus sensibles aux contraintes en compression du STI comme cela est illustré sur la figure 3.34. Pour les nouvelles technologies, les contraintes seront plus importantes et donc un modèle prédictif sera nécessaire pour prédire les performances électriques en fonction des contraintes mécaniques.

L’objectif recherché, à savoir la quantification du niveau des contraintes en fonction de « a » et l’établissement des modèles prédictifs étant atteint, on exposera dans les prochains paragraphes l’effet de la hauteur de marche sur le niveau des contraintes.

3.4 Conclusion

Nous avons montré dans la première partie de ce chapitre les corrélations observées entre la résistance carrée des caissons implantés à travers le STI et l’épaisseur de STI. La résistance carrée des caissons (Nwell STI) croît linéairement avec l’épaisseur STI pour les transistors PMOS alors que pour les transistors NMOS, l’évolution de la résistance carrée des caissons (Pwell STI) est quasi constante (faible corrélation). Nous avons pu démontrer la différence de comportements entre les deux caissons grâce à des simulations d’implantation. Ces analyses ont révélé des positions de pic d’implantation différentes. Nous avons démontré qu’on pouvait améliorer la linéarité de la résistance des caissons P en fonction de l’épaisseur STI et cela en modifiant l’énergie d’implantation. L’autre paramètre électrique qui est fortement corrélé avec la morphologie du STI est la tension de seuil des transistors parasites. Elle présente une forte variabilité (environ 65% quand on fait varier la hauteur de marche de 150 \AA à 850 \AA). Cette forte variabilité a été un argument suffisant pour choisir ce paramètre comme un indicateur du suivi des boucles de régulation. Les caractérisations électriques du transistor $I_D(V_G)$ ont confirmé le résultat des corrélations à savoir que l’augmentation de l’épaisseur d’oxyde STI induit une augmentation de la tension de seuil.

Dans la partie consacrée aux contraintes mécaniques, nous avons confirmé l'effet de la distance entre le STI et la grille (distance caractérisant l'éloignement de la tranchée par rapport à la grille (distance «a»)). Pour une diminution de cette distance, la mobilité des porteurs se dégrade dans le cas du NMOS alors qu'elle augmente dans le cas du PMOS. Nous avons montré que pour une même valeur de cette distance, la variation absolue de la mobilité est supérieure dans le cas des PMOS (environ 35 % quand on fait varier la distance de 0,4 à 4,9 μ m alors que pour les NMOS cette variation est estimée à 10%). Nous avons expliqué ces résultats par le fait que quand le STI se rapproche de l'active, il y a augmentation des contraintes mécaniques (stress mécanique) et en particulier en compression dans les directions X et Z. Les mesures sous flexion nous ont permis d'extraire les coefficients piezorésistifs pour les deux types de transistors. Les mesures sous flexion ainsi que les résultats électriques nous ont permis de quantifier le stress mécanique en fonction de la distance « a ». Les résultats expérimentaux sont en accord avec les résultats de la modélisation numérique effectuée. Nous avons établi un modèle prédictif donnant le niveau des contraintes en fonction de la distance « a ».

Dans le chapitre suivant, nous allons exposer toutes les stratégies de contrôle pour réduire la variabilité au niveau de la brique STI.

Références bibliographiques

- [FERREIRA 1996] P. Ferreira, V. Senez and B. Baccus, Mechanical stress analysis of an LDD MOSFET structure" IEEE. Trans. Electron Devices, vol.43, n°5, pp.1525-1532,sept. 1996.
- [GALLON 2004] C. Gallon, G. Reimbold, *Member, IEEE*, Gérard Ghibaudo, *Senior Member, IEEE*, R. A. Bianchi, R. Gwoziecki,S. Orain, E. Robilliart, C. Raynaud, and H. Dansas, Electrical Analysis of Mechanical Stress Induced by STI in Short MOSFETs Using Externally Applied Stress, *IEEE TRANSACTIONS ON ELECTRON DEVICES*, VOL. 51, NO. 8, AUGUST 2004,p1254-1261
- [GALLON 2004a] C. Gallon, G. Reimbold , G. Ghibaudo , R.A. Bianchi , R. Gwoziecki, Electrical analysis of external mechanical stress effects in short channel MOSFETs on (0 0 1) silicon, *Solid-State Electronics* 48 (2004) 561–566
- [HUANG 1996] C. Huang, H. Soleimani, G.Grula,N.D. Arora and D. Antoniadis, Isolation process dependence of channel mobility in thin-film SOI devices, *IEEE Electron Device letter*, Vol. n°6,pp.291-293, 1996.
- [KASBARI 2007] M. Kasbari, thèse de doctorat, CMPGC-EMSE, 2007
- [ORTOLLAND 2006] Claude Ortolland, Etude des effets des contraintes mécaniques induites par les procédés de fabrication sur le comportement électrique des transistors CMOS des nœuds technologiques 65 nm et en deça. Thèse de doctorat, INSA Lyon, 2006.
- [RIVERO 2007] Christian Rivero, rapport technique interne, ST Rousset, 2007
- [SENEZ 2001] V. Senez, T. Hoffmann, E. Robilliart et al, Investigations of stress sensitivity of 0,12 CMOS technology using process modelling, IEDM Tech. Digest, pp. 831-834, Electron Devices Meeting 2-5 Dec.2001.
- [REGGIANI 1985] L. Reggiani, hot electron transport in semiconductors, New-york, springer-verlag, 1985, ISBN: 0387133216

Chapitre 4

Boucles de régulation R2R sur le module STI

Dans ce chapitre, nous allons détailler toutes les solutions proposées pour réduire la variabilité au niveau de certains procédés considérés comme critiques que nous allons déterminer par une analyse de variabilité. Ce chapitre est composé de trois grandes parties. Dans chaque partie, nous allons décrire les variabilités constatées, les paramètres de procédés critiques ainsi que les plans d'expériences effectués pour trouver un modèle du procédé qui sera le cœur de la boucle de régulation R2R.

4.1 Analyses de variabilités

Les variabilités sont souvent constatées au niveau des mesures de la hauteur de marche et de son uniformité. Elles sont la résultante des différentes variations au niveau des procédés de fabrication. Nous allons identifier les procédés qui contribuent à cette variabilité et nous allons proposer des solutions dans les parties suivantes de ce chapitre pour réduire la variabilité au niveau de chaque procédé identifié comme un procédé critique.

4.1.1 Introduction

Dans les paragraphes précédents, nous avons discuté des corrélations entre les paramètres STI (hauteur de marche et épaisseur STI) avec certains paramètres électriques. Ces analyses nous ont permis de déduire un indicateur électrique (V_t des transistors test parasites) pour le suivi du STI.

Une simple analyse sur un historique de données (environ 100 lots) montre qu'il y a une dispersion de lot à lot comme cela est illustré sur la figure 4.1.

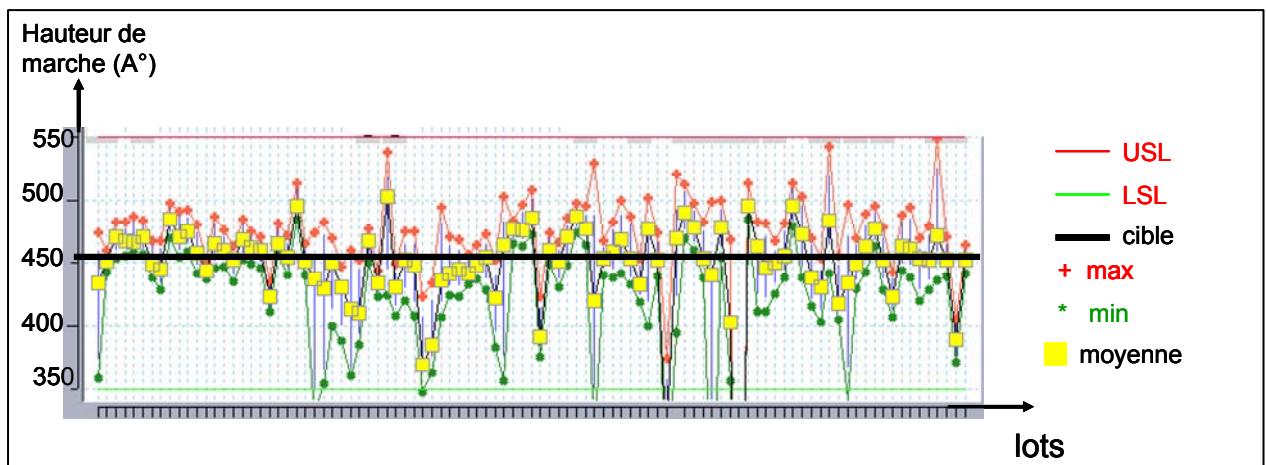


Figure 4.1- Dispersion de la hauteur de marche de lot à lot

Comme on a une forte corrélation entre la hauteur de marche et l'indicateur (V_t), une dispersion sur ce dernier semble évidente. Une analyse sur un historique de données confirme cette hypothèse comme cela est montré sur la figure 4.2. Nous observons sur la figure 4.2 les données de mesures de la hauteur de marche des différents lots. Certaines mesures sur le graphe indiquent que les valeurs de hauteur de marche sont hors contrôle puisqu'elles ne sont pas comprises dans les limites de contrôle basses (LCL) et hautes (UCL).

Nous avons évoqué dans le chapitre précédent que la tension (V_t) est fortement corrélée avec la hauteur de marche et donc les variabilités de la hauteur de marche de lot à lot vont se traduire par une variabilité de V_t . La figure 4.2 confirme cette situation où chaque Box plot représente les mesures des valeurs de V_t sur une plaque d'un lot. Sur une même plaque, nous observons également la non uniformité.

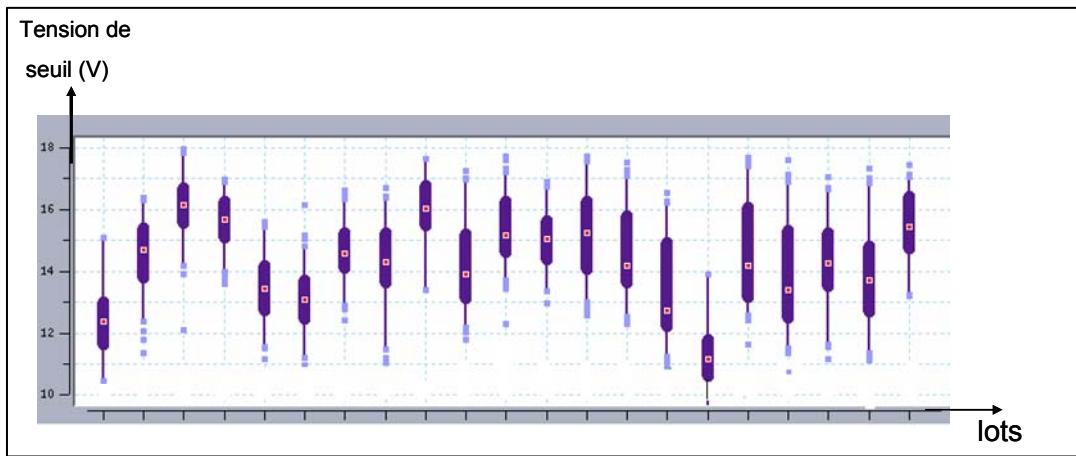


Figure 4.2- Dispersion de la tension de seuil des transistors test parasites.

On détaillera dans les paragraphes qui suivent les actions proposées pour réduire cette dispersion de lot à lot et la réduction de la non uniformité sur la plaque.

4.1.2 Rappel du procédé de la réalisation de la brique STI

Nous avons décrit au chapitre 2 les opérations élémentaires de la réalisation de la brique STI. Nous allons récapituler les principales opérations de fabrication mais aussi toutes les étapes de mesure comme cela est illustré sur la figure 4.3. Comme nous pouvons le constater nous mesurons une ou plusieurs grandeurs physique après une opération de fabrication.

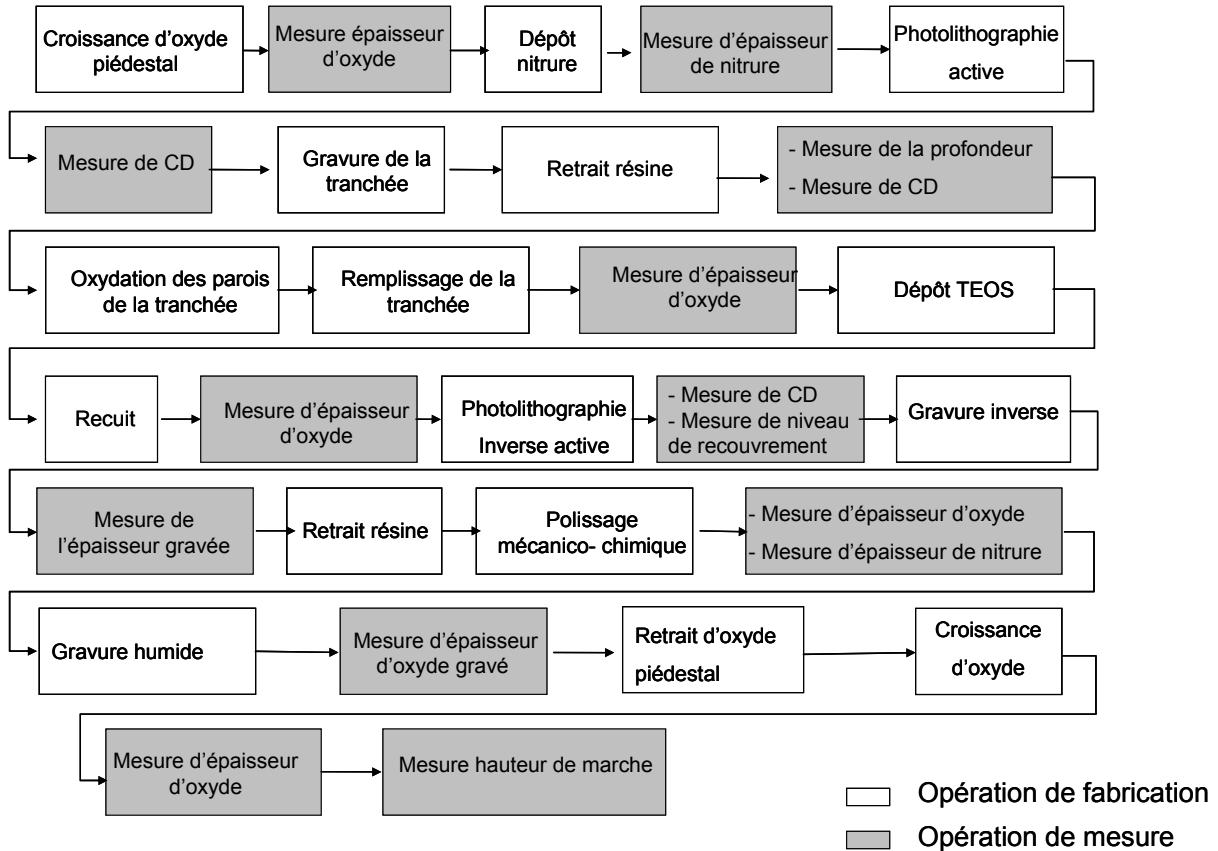


Figure 4.3- Récapitulatif des opérations de fabrication de la brique STI (pour les technologies dont la longueur de grille des transistors <0.18µm)

Nous avons effectué des analyses sur un historique de données regroupant toutes les données des étapes de mesures.

4.1.3 Etude de la variabilité

Les premières investigations visent à quantifier la dispersion et la nature de la dispersion. Il est recommandé de disposer d'un maximum d'observations sur une durée de temps importante incluant les différentes interventions sur la machine (comme les maintenances préventives). Sur le premier point, la dispersion est très importante à quantifier pour démontrer son impact sur le rendement final mais aussi pour justifier le déploiement des boucles de régulation. La nature de la dispersion peut être de:

- batch à batch
- lot à lot
- plaque à plaque
- site à site

Nous avons collecté un historique de données (6mois), nous avons ensuite traité les données pour n'analyser que les plaques (les lots) qui ont suivi le même enchainement des étapes de fabrication et des mesures. Il y a peu de plaques (ou des lots) qui sont mesurées au niveau de toutes les étapes de mesures à cause de l'échantillonnage imposé, c'est ce qui explique le choix de l'historique de 6 mois. Nous avons comparé la moyenne des paramètres

de mesures avec leur valeur cible pour quantifier la dispersion. La figure 4.4 montre les résultats des variabilités des différents procédés de fabrication contribuant à la réalisation de la brique STI dans les deux cas étudiés (STI standard et Direct STI).

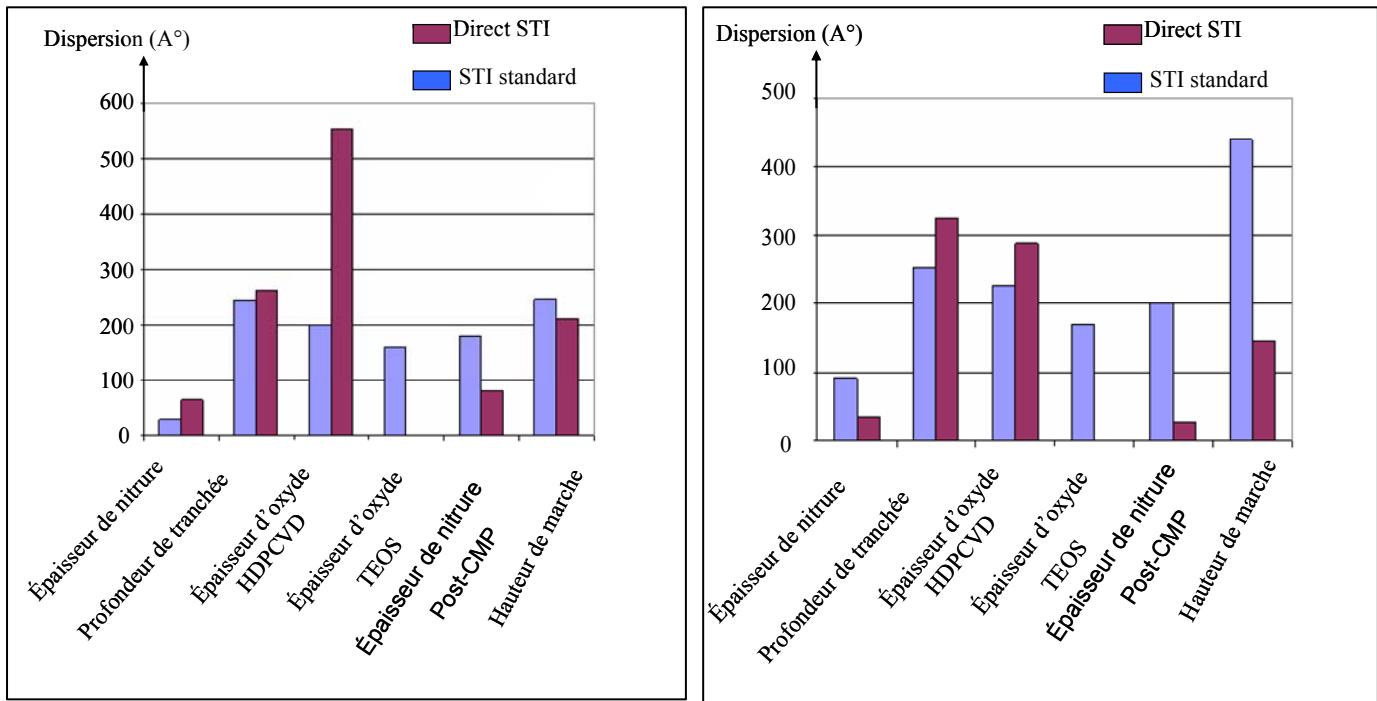


Figure 4.4- Un exemple d'étude de variabilité des paramètres de mesures du module STI
a) de lot à lot b) de plaque à plaque

Comme Cela est illustré sur la figure 4.4, les données de deux technologies ayant deux modes de STI différents (STI standard et Direct STI) sont comparées. Il apparaît très clairement que les procédés de gravure de la tranchée, le dépôt d'oxyde HDPCVD et le CMP sont les procédés les plus contributeurs à la dispersion de la hauteur de marche. Nous observons sur la figure 4.4 que la dispersion de lot à lot de l'épaisseur d'oxyde HDPCVD déposée est plus importante dans le cas du Direct STI (550 Å au lieu de 200 Å dans le cas du STI standard). Ceci peut s'expliquer par la différence des valeurs cible de l'épaisseur d'oxyde déposée (6800 Å pour le Direct STI et 5500 Å pour le STI standard)

Notons que la variation du dépôt nitrate est négligeable devant la variation des autres procédés. Pour cette raison, la variation de l'oxyde piédestal « padox » n'est pas citée dans cette analyse.

4.1.4 Boucles de régulation sur le module STI

4.1.4.1 Direct STI

Suite aux analyses de variabilité, 3 boucles de régulations R2R ont été proposées pour réduire la variabilité au niveau de ces procédés :

- Une boucle de type « feedback » au niveau du procédé de gravure sèche
- Une boucle de type « feedback » au niveau du procédé de dépôt HDPCVD
- Une boucle de type « feed forward » entre les procédés de CMP et la gravure humide (gravure WET)

L'ensemble des boucles proposées sont illustrées sur la figure 4.5

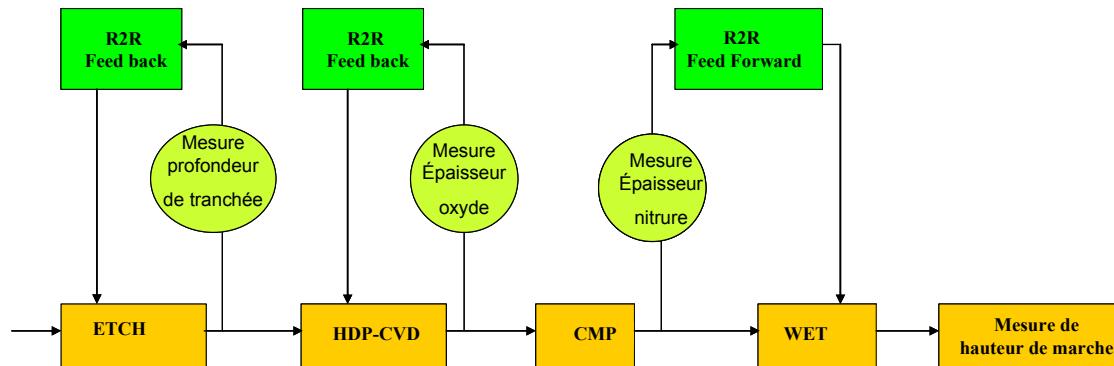


Figure 4.5- Boucles de régulations proposées sur le module STI (Direct STI)

4.1.4.2 STI standard

De la même façon, on va implémenter les mêmes boucles proposées précédemment dans le cas du STI standard. Une boucle de type « feed back » a été rajoutée car le STI standard est marqué par une non uniformité de l'épaisseur d'oxyde après CMP mais aussi de la hauteur de marche. Il est impératif d'améliorer l'uniformité de l'oxyde à l'aide d'une boucle « feed back » comme cela est illustré sur la figure 4.6.

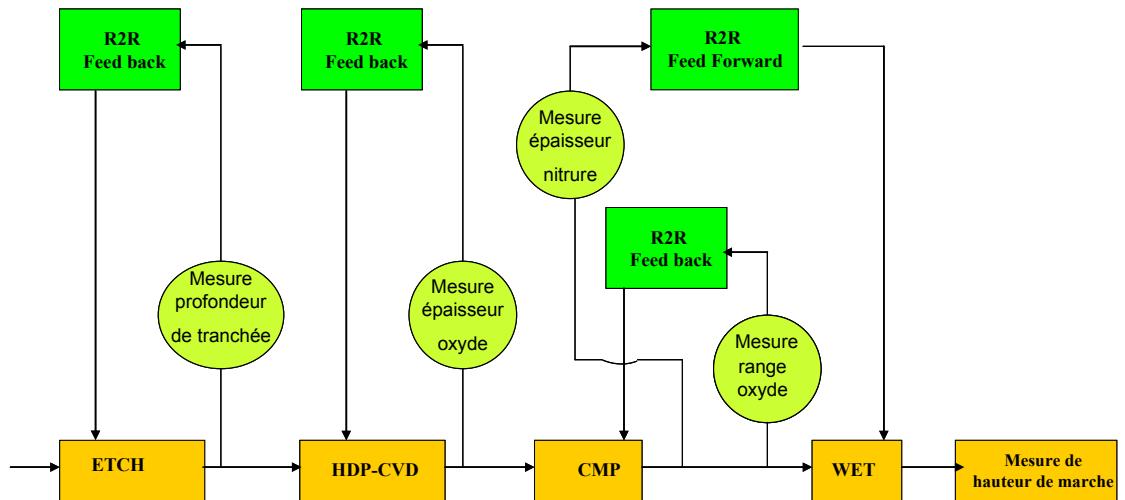


Figure 4.6- Boucles de régulations proposées sur le module STI (STI standard)

Le déploiement des boucles de régulation doit être justifié d'un point de vue économique. Pour cela, des simulations de gain en dispersion de la hauteur de marche ont été

effectuées. L'ensemble des actions permet de réduire la dispersion de la hauteur de marche exprimée par l'écart type (standard déviation).

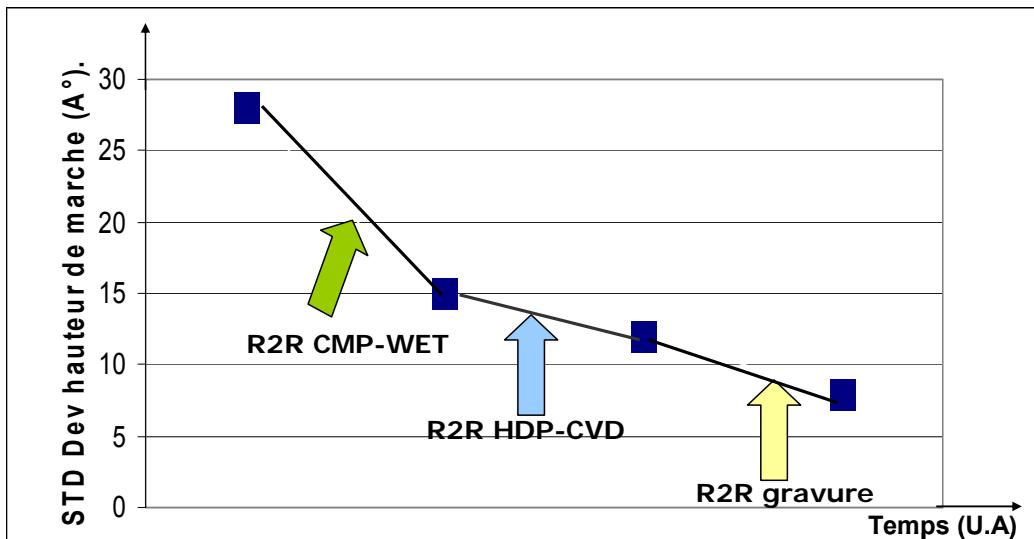


Figure 4.7- Exemple de simulation sur le gain en dispersion de la hauteur de marche pour le Direct STI suite aux actions proposées.

La figure 4.7 montre un exemple de simulation de gain sur un produit utilisant le direct STI. L'écart type de la hauteur de marche sans la présence d'une régulation automatique sur les équipements est de 28Å, avec le déploiement de la boucle « feed forward » sur les procédés du CMP et la gravure humide, l'écart type de la hauteur de marche est réduit à 15 Å. En implémentant la boucle « feed back » sur le procédé HDP-CVD, l'écart type de la hauteur de marche peut être réduit à 12Å. L'écart type de la hauteur de marche peut être réduit à 7Å en réalisant la boucle de type « feedback » sur le procédé de la gravure sèche.

4.1.5 Organisation du travail

Dans les parties suivantes nous décrirons les actions entreprises pour la réalisation des boucles illustrées sur les figures 4.5 et 4.6.

- Boucle feed back HDP-CVD (ajustement de lot à lot pour le STI standard et le Direct STI) ; on détaillera toutes les étapes essentielles pour l'obtention du modèle de ce procédé qui est le cœur de la boucle R2R.
- Boucle feed back CMP (uniquement pour le STI standard). Dans le souci d'améliorer l'uniformité sur le STI standard, la boucle R2R feed back permettra un réajustement des paramètres CMP de lot à lot. Nous détaillerons les travaux menés pour modéliser l'effet des pressions de tête de la machine de polissage sur l'uniformité de l'épaisseur d'oxyde post-CMP.
- Boucle feed forward CMP-WET (pour le Direct STI), en vue de compenser les dispersions de l'épaisseur de nitrure et d'oxyde après CMP qui impactent directement la hauteur de marche. Un réajustement des paramètres du procédé de la gravure humide peut réduire cette dispersion et donc réduire la dispersion de la hauteur de marche. Il faut noter que cette boucle peut être appliquée pour le STI standard à condition que la non uniformité soit réduite

- La boucle feed back (gravure) ne sera pas décrite dans ce manuscrit car les outils de mesure de la profondeur de tranchée ne sont pas assez robustes et reproductibles pour les employer dans une telle boucle. Par exemple, la mesure par la microscopie à force atomique (AFM : Atomic Force Microscopy) est une technique très lente et qui présente des erreurs de mesure allant jusqu'à 50 Å. Il faut donc disposer des techniques robustes de mesure telle que la scattérométrie (la technique d'éllipsométrie améliorée) [LENSING 2001] pour avoir des mesures robustes.

Pour chacune des boucles R2R, un modèle de procédé sera recherché. Suivant les paramètres de procédé ajustables et les mesures à contrôler, les modèles peuvent être de type :

- SISO (single input single output) où nous disposons d'une seule variable d'entrée (paramètre du procédé) et d'un seul paramètre de mesure
- MIMO (multiple input multiple output) où nous disposons de plusieurs variables d'entrée (paramètres du procédé) et de plusieurs paramètres de mesure
- MISO (multiple input single output) où nous disposons de plusieurs variables d'entrée (paramètres du procédé) et d'un seul paramètre de mesure
- SIMO (single input multiple output) où nous disposons d'une seule variable d'entrée (paramètres de procédé) et de plusieurs paramètres de mesures

Partie 1

Procédé de dépôt chimique en phase vapeur assisté par plasma haute densité (HDP-CVD)

Dans cette partie, nous exposerons la modélisation du remplissage d'oxyde des tranchées HDP-CVD. Ce modèle constituera le cœur de la boucle de régulation feedback R2R du procédé HDP-CVD. Nous détaillerons le principe du procédé HDP-CVD et ses principales caractéristiques ainsi que l'équipement de dépôt. Ensuite, on exposera les types de variabilités constatées lors du remplissage des tranchées. Les résultats des analyses multivariées seront illustrés et les paramètres identifiés comme critiques seront listés. Le modèle du procédé sera également détaillé. On commenterà chacun des modèles établis: celui de l'épaisseur et de son étendue. Le modèle d'épaisseur est très satisfaisant, par contre le modèle de l'étendue de l'épaisseur dépend fortement du type de la cartographie de mesures. Nous avons amélioré ce modèle en utilisant une technique d'apprentissage statistique.

4.2.1 Introduction

Dans le but de limiter les contaminations et d'améliorer le taux de couverture de marche⁴⁷, la technologie de dépôt chimique en phase vapeur CVD (Chemical Vapor Deposition) à basse pression a été introduite dans les années 80. Cette méthode est aujourd'hui l'une des plus utilisées en microélectronique pour la réalisation de dépôts d'isolants sur semi-conducteurs. Elle permet d'obtenir des films uniformes et reproductibles en épaisseur, possédant une bonne conformité⁴⁸ et un bon taux de couverture [COTE 1995]. Le procédé CVD permet de déposer des couches épaisses d'oxyde de silicium (SiO_2) [PIERSON 1999], nécessaires à la conception des dispositifs. Le procédé CVD est employé sur différents niveaux lors de la fabrication des circuits intégrés:

- niveau de l'isolation : isolations profondes (STI). Ces couches d'oxydes permettent d'isoler électriquement les dispositifs.
- niveau pré-métal : les couches d'oxyde pré-métalliques (PMD : Pre-Metal Dielectric) empêchent la migration des ions (cas du PSG : Phospho-Silicate Glass).
- niveau inter métal (IMD : InterMetallic Dielectric), les couches diélectriques assurent une isolation des lignes de métal.
- prépassivation et passivation.

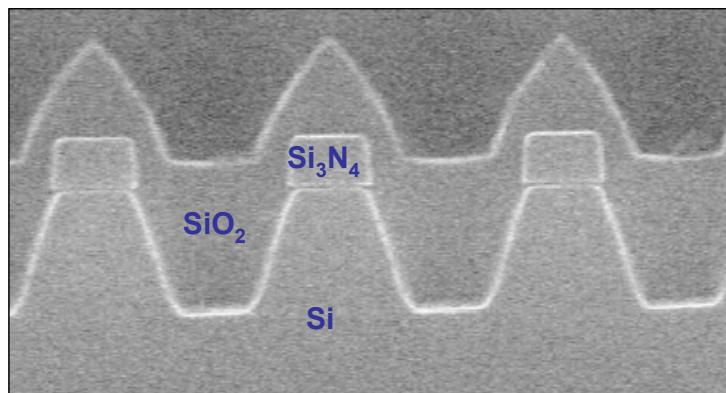


Figure 4.8- Remplissage des tranchées de STI (dépôt HDP-CVD + TEOS)

La figure 4.8 montre un exemple de dépôt d'oxyde dans les tranchées STI. Pour une bonne qualité d'isolation, il faut déposer un oxyde de bonne stoechiométrie et de bonne qualité. La moyenne de l'épaisseur déposée ainsi que l'étendue de l'épaisseur sont les paramètres de mesure qu'il faut surveiller d'un point de vue de contrôle des procédés. Dans les paragraphes suivants, nous allons décrire notre proposition pour réduire la dispersion de l'épaisseur et de son étendue. Nous allons tout d'abord décrire le principe et les mécanismes de dépôt HDP-CVD.

4.2.2 Principe du dépôt HDP-CVD

4.2.2.1 Rappel des principales notions de dépôt CVD

⁴⁷ La mesure de la reproductibilité de la forme d'une partie de substrat par un film déposé.

⁴⁸ La capacité à reproduire de manière exacte la topographie des couches sous-jacentes.

Les films obtenus par la technique de dépôt CVD sont caractérisés souvent par certains indices géométriques. Ces indices géométriques sont illustrés sur la figure 4.9. On désigne souvent la couverture de marche et la conformité comme principaux indicateurs pour caractériser la qualité d'un dépôt.

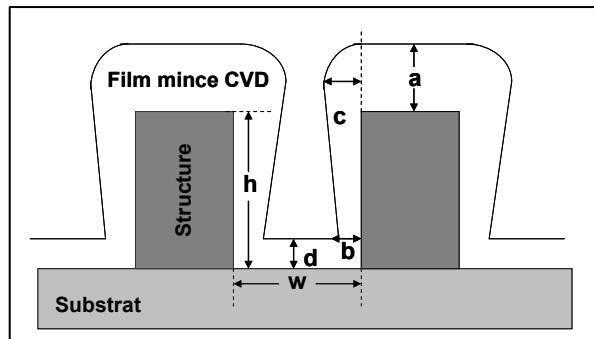


Figure 4.9- Schématisation des principaux indices géométriques caractérisant le dépôt

On définit :

- la couverture de marche des parois exprimée par le rapport (b/a)
- la couverture de marche du fond exprimée par le rapport d/a
- la conformité exprimée par le rapport (b/c)
- le rapport $(c-b)/b$, ce rapport est appelé « overhang »
- le facteur de forme est le rapport h/w (rapport de la profondeur sur la largeur).

4.2.2.2 Mécanisme de dépôt pour le remplissage d'oxyde dans les tranchées

La brique STI ayant pour rôle d'isoler électriquement les transistors, il faut qu'elle soit intègre. Cela veut dire que l'oxyde déposé doit correctement remplir la tranchée sans créer directement de trou [NAKAMURA 2001] ou que les étapes ultérieures au dépôt d'oxyde dans la tranchée ne viennent attaquer l'oxyde. La figure 4.10 illustre l'apparition de trous dans la tranchée d'oxyde dans le cas d'un procédé PECVD classique.

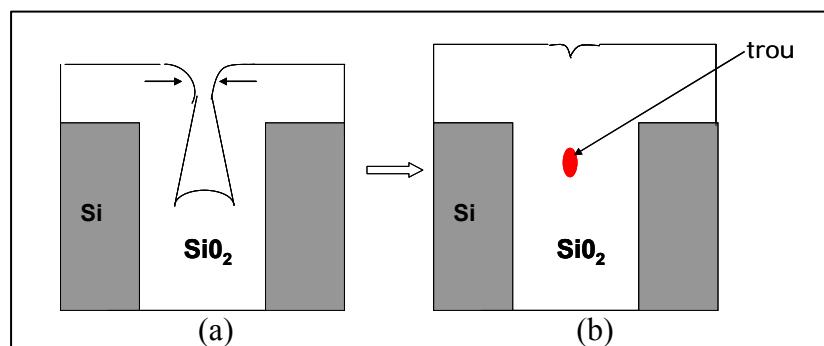


Figure 4.10- (a)Vue en coupe d'un dépôt PECVD⁴⁹ classique (b) apparition d'un trou dans la tranchée

⁴⁹ PECVD : Plasma Enhanced CVD (CVD assistée par plasma)

Pour éviter ce phénomène, il convient de graver spécifiquement les parties « surdéposées » et dans l'idéal, d'effectuer cette gravure simultanément au dépôt de manière à limiter la durée du procédé. C'est le principe du dépôt CVD assisté par plasma (HDP-CVD). Cette gravure est de type ionique et est réalisée grâce à des ions d'argon (Ar^+) dont la neutralité chimique permet de ne pas dénaturer les propriétés du dépôt. Il est évident que cette technique facilite le remplissage des tranchées et permet d'empêcher l'apparition de trous. Le dépôt HDP-CVD est la meilleure technique de dépôt d'oxydes dans les tranchées STI [ABDELGADIR 2006]. La technique est rendue possible du fait que le dépôt et le facettage par gravure ionique ont des origines physiques et chimiques différentes. Ainsi, les deux procédés sont indépendants et complémentaires. Ils peuvent être ajustés de manière à déposer des couches d'oxyde uniformes et d'excellente qualité, capables de remplir d'étroites tranchées. Le mécanisme de dépôt HDP-CVD est illustré sur la figure 4.11.

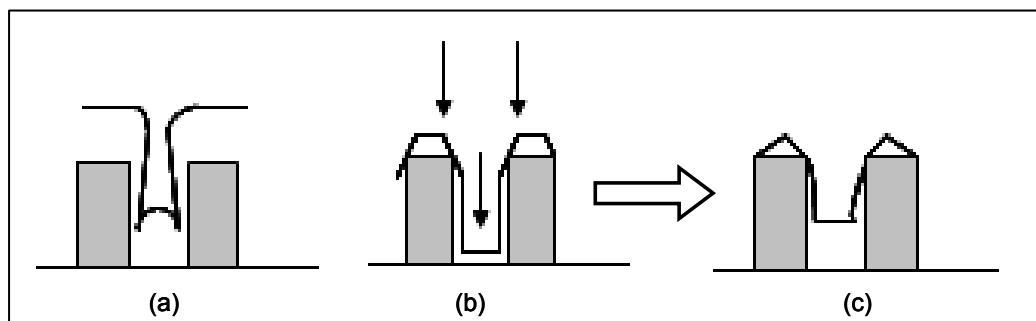
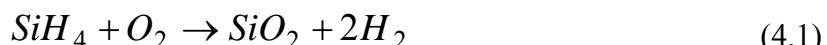


Figure 4.11- Mécanisme du procédé (a) dépôt (b) gravure (c) dépôt HDP-CVD.

La technique du dépôt HDPCVD allie simultanément un dépôt d'oxyde assez isotrope (figure 4.11a) à une gravure assez anisotrope comme cela est illustré sur la figure 4.11b. Trois gaz interviennent au cours de ce procédé [SUH 2003 b] : l'oxygène (O_2), le silane (SiH_4) et l'argon (Ar). Les espèces pulvérisées se redéposent dans la tranchée et contribuent au remplissage homogène (sans trous) de la tranchée. La réaction chimique entre les précurseurs conduit à la formation d'oxyde de silicium par la réaction suivante :



La gravure est réalisée grâce à un plasma faisant intervenir un procédé mécanique à base d'ions (Ar^+ et O^+). C'est un plasma haute densité qui présente un taux d'ionisation d'environ 0.01% et une concentration d'ions de $10^{11}/10^{12}\text{cm}^{-3}$ (contre 0.0001% et 10^{10}cm^{-3} habituellement). Le plasma contient les espèces injectées (Ar^+ , O^+), mais aussi les composés issus de la dissociation du silane. La pression est régulée à une valeur relativement basse [SUH 2003 a] de manière à augmenter le libre parcours moyen des ions. Ces derniers acquièrent ainsi suffisamment d'énergie pour ioniser les molécules lors des collisions.

4.2.2.3 Le rapport Dépôt / Pulvérisation⁵⁰ (D/S)

⁵⁰ La pulvérisation est une technique qui repose sur le bombardement d'une cible par un gaz pour arracher mécaniquement des atomes de la cible.

Le procédé HDP-CVD est quantifié grâce au rapport des vitesses de dépôt et de gravure (D/S). Ce paramètre est directement relié à la capacité du procédé à remplir des tranchées plus ou moins étroites [MUNGEKAR 2006]. La vitesse de gravure (S) est déterminée par gravure d'une plaquette test d'oxyde avec les conditions du procédé réelles mais sans silane. Une mesure de la vitesse de dépôt du remplissage de la tranchée permet une détermination du taux de dépôt sur plaquette test.

$$\boxed{\text{Vitesse de dépôt nette} = \text{vitesse de dépôt (plaqué test)} - \text{vitesse de gravure (plaqué test)}} \quad (4.2)$$

Le rapport D/S suit la relation :

$$\boxed{\frac{D}{S} = \frac{\text{vitesse de dépôt (plaqué test)}}{\text{vitesse de gravure (plaqué test)}} = \frac{\text{vitesse de dépôt (nette)} + \text{vitesse de gravure (plaqué test)}}{\text{vitesse de gravure (plaqué test)}}} \quad (4.3)$$

Le remplissage est réalisé en deux étapes. Par exemple, pour avoir une couche d'oxyde avec une épaisseur de 5500 Å, on dépose dans un premier temps, une première couche de 1500 Å avec un rapport D/S de 4 puis une deuxième couche de 4000 Å avec un rapport D/S de 7,2. La faible valeur du rapport (D/S=4) choisie lors de remplissage de la première couche est justifiée par le fait que nous voulons éviter la création de trous dans la tranchée lors du procédé HDP-CVD. Ensuite nous pouvons enchaîner le dépôt avec des grandes vitesses de dépôt d'où le choix de rapport (D/S=7,2) pour la deuxième couche. Comme le dépôt se fait en deux étapes successives, celles-ci sont précédées par des étapes de conditionnement puis suivies d'étapes de purge de la chambre une fois le dépôt terminé. Comme cela est montré sur la figure 4.12, l'épaisseur totale du dépôt est la somme de l'épaisseur déposée pendant la première étape (Y1) et celle déposée pendant la deuxième étape (Y2). Ce type de dépôt en deux étapes est appelé également double HDP-CVD (DHDP-CVD).

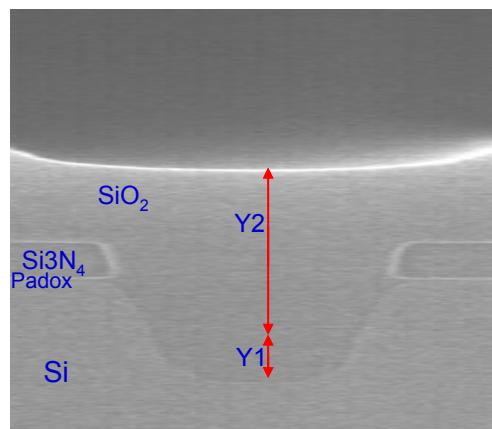


Figure 4.12- Dépôt d'oxyde en deux étapes

4.2.2.4 Applications des dépôts HDP-CVD

Hormis le remplissage d'oxyde dans les tranchées STI. Le procédé HDP-CVD permet également de faire des dépôts d'oxydes dopés ou non dopés destinés au remplissage des espaces les plus critiques à différents niveaux :

- * Le verre de silice non dopé (USG : Undoped Silicate Glass) est un oxyde utilisé pour l'isolation des lignes de métal (IMD) depuis la technologie $0.35\mu\text{m}$. Ses propriétés sont les suivantes :
 - Constante diélectrique faible : $\epsilon_r = 4.3$
 - Indice de réfraction : $n = 1.46$;
 - Faible proportion de SiH et SiOH (inférieure à 4%).
- * Le verre de silice fluoré (FSG : Fluorinated Silicate Glass) ou oxyde de silicium dopé au fluor (F) est classé dans la catégorie des oxydes à basse permittivité diélectrique. Le couplage capacitif entre les lignes de métal est réduit et permet ainsi d'abaisser le temps de réponse global du circuit. Ses propriétés sont les suivantes :
 - Constante diélectrique faible : $\epsilon_r = 3.7$.
 - Indice de réfraction : $n = 1.445$.
- * Le verre de silice phosphoré (PSG: Phospho-Silicate Glass) est utilisé en tant que couche de passivation. Il possède la propriété de piéger les contaminants ioniques et protège ainsi le circuit.

4.2.2.5 Description de l'équipement

L'équipement utilisé pour le dépôt HDP-CVD est composé de 3 chambres de dépôt comme cela est illustré sur la figure 4.13. Le chargement des plaques s'effectue par l'un des deux sas de chargement. Les plaques sont ensuite déposées sur une cassette avant d'être transportées au carrefour commun aux trois chambres. La chambre de transfert est équipée d'un bras qui permet de transférer les plaques d'un élément à l'autre de l'équipement. Nous distinguons également sur la figure 4.13 la chambre de refroidissement dans laquelle les plaques transitent après le dépôt, avant de revenir dans leur cassette.

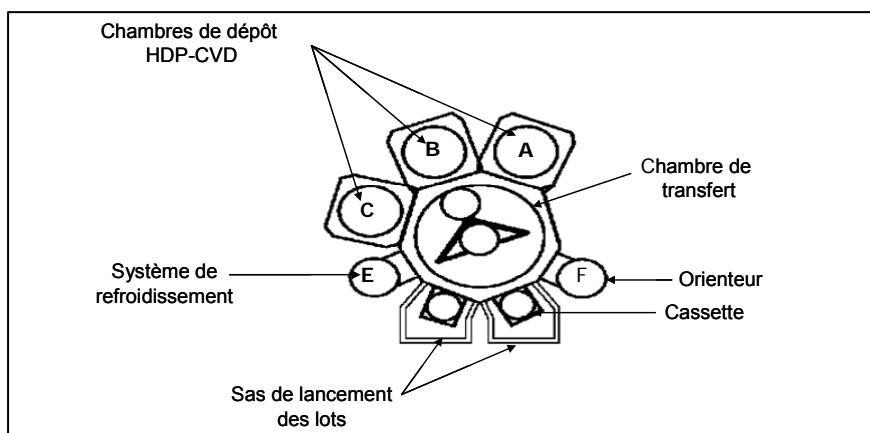


Figure 4.13- Configuration de l'équipement HDP-CVD.

Les plaques entrent dans les chambres de dépôt grâce à une fente ouverte de manière automatique pour cet effet. Elles sont ensuite placées sur un porte échantillon (chuck) qui les fixe de manière électrostatique. Les différents éléments d'une chambre de dépôt HDP-CVD sont représentés sur la figure 4.14.

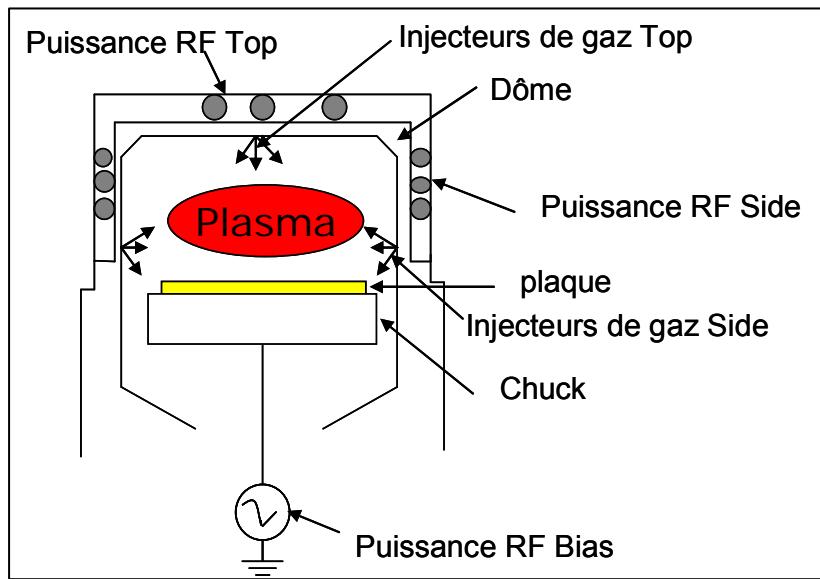


Figure 4.14- Schéma d'une chambre de dépôt

Nous pouvons classer la configuration d'une chambre par deux parties distinctes :

- Le dôme⁵¹ en céramique autour duquel est situé deux électrodes. La puissance RF⁵² source est injectée dans ces électrodes ce qui produit un plasma haute densité par couplage inductif. Un double injecteur silane- argon/oxygène se situe au centre de dôme.
- La partie basse où se trouve le chuck polarisé par la RF Bias⁵³ et 24 injecteurs de gaz (16 pour SiH₄ ou Ar et 8 pour O₂). La plaque est maintenue sur le bâti par effet électrostatique. Le circuit de pompage est constitué d'une pompe primaire et d'une pompe turbo- moléculaire.

Nettoyage de la chambre

Un système de nettoyage permet de réaliser périodiquement des nettoyages de la chambre en attaquant les particules présentes sur les parois de la chambre par du fluor. Celui ci est obtenu par dissociation de NF₃ hors de la chambre grâce à un magnétron (2,45 GHz) qui génère des micro-ondes. Ces dernières sont véhiculées au moyen d'un guide d'ondes jusqu'à l'applicateur⁵⁴. Les mico-ondes excitent les électrons libres du gaz NF₃, une collision avec les molécules de NF₃ permet de les dissocier. On obtient ainsi des atomes de fluor F^{*} dissociés et réactifs (* symbolise une surface activée).

⁵¹ Il est en céramique afin de pouvoir résister aux « agressions » du plasma. Il peut subir de fortes contraintes thermiques et mécaniques. Il est isolant et transparent aux ondes Radio Fréquence (RF).

⁵² Deux générateurs RF séparés fournissent la puissance nécessaire pour former le plasma haute densité (10^{12} ions/cm³). On distingue deux sources : la source top permettant d'allumer le plasma et la source side qui sert à canaliser les lignes de champ et à confiner le plasma.

⁵³ Le plasma est allumé par les deux bobinages side et top (alimentés par des générateurs RF). La tension Bias contrôle le flux d'ions vers la plaque.

⁵⁴ Il est constitué de deux tubes concentriques : l'un en saphir dans lequel est injecté du NF₃, l'autre en quartz entourant le premier. Entre les deux, l'eau circule pour le refroidissement.

4.2.3 Les axes principaux de recherches consacrés à l'implémentation des boucles de régulation R2R pour le procédé CVD

Pour positionner notre contribution sur la modélisation du procédé HDP-CVD, nous allons décrire les différents travaux académiques et industriels effectués sur la modélisation du procédé CVD et aussi l'implémentation de boucles de régulation de type R2R.

4.2.3.1 Dépôt par épitaxie

Les travaux de thèse de A. GOWER [GOWER 2002] à l'institut MIT portaient sur la réalisation d'un modèle de l'uniformité de dépôt épitaxial de silicium pour une boucle de régulation R2R de type feed back. Le modèle proposé est de type SIMO où une seule variable d'entrée est contrôlée : le temps de dépôt. Les paramètres de mesure à contrôler sont les mesures d'épaisseurs et la résistivité du film déposé (la résistivité est employée pour contrôler le dopage du film).

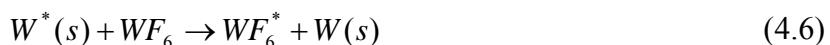
4.2.3.2 Dépôt du Tungstène par CVD

Le modèle de la boucle de régulation R2R du procédé du dépôt CVD de tungstène rapporté par S.SREENIVASAN [SREENIVASAN 2001] de l'université de Maryland est une suite des travaux déjà initiés par GOUGOUSI [GOUGOUSI 2000] et aussi des travaux de thèse de Y.XU [XU 2000]. GOUGOUSSI a constaté une dérive sur l'équipement, la température décroît de 5°C après avoir traité une plaque sachant que la température opératoire du procédé varie de 350 à 400°C. Le modèle élaboré est de type SISO, une seule variable de procédé à ajuster (le temps de dépôt) et le paramètre de mesure à contrôler est l'épaisseur d'oxyde déposé. Le suivi en temps réel de la concentration de HF dans la réaction par spectrométrie de masse permet de mesurer la vitesse de dépôt et indirectement la température de nucléation. Cette réaction de nucléation du W est thermiquement activée et conduit à la formation de HF [XU 2000].

Les réactions de nucléation à l'interface W/Si sont données ci-dessous :



Une fois la couche de nucléation à la surface de silicium formée, la suite des opérations donne lieu à une réaction de nucléation sur la couche de W en formation. La formation de la couche W provoque une réduction de WF₆ en générant une surface activée comme nous pouvons le constater dans la relation 4.7.



Pour chaque atome W déposé en surface, 3 molécules de H₂ et une molécule de WF₆ sont consommées et six molécules de HF sont produites d'après la relation 4.7. Le paramètre à contrôler est l'épaisseur de tungstène déposée. Le paramètre qui relie directement la

variation de l'épaisseur du tungstène déposé est le signal de spectrométrie de masse du HF, directement proportionnel au flux de W condensé.

4.2.3.3 Modèles simples pour le Dépôt CVD

J.STEFANI [STEFANI 2006] de Texas Instruments a développé une technique de contrôle sur les procédés de pulvérisation cathodique (dépôt métal), les procédés de dépôt électrochimique et de dépôt CVD et parmi eux le dépôt HDP-CVD. Cette technique est appelée Run by Run (RbR). Cette technique peut être qualifiée de R2R discontinue, car pour cette technique, le réajustement intervient après les tâches de qualifications⁵⁵ et il s'arrête avant la prochaine tâche de qualification. Tous les modèles proposés par J.Stefani sont très simples et ils ne consistent qu'à réajuster le temps de dépôt. Ces modèles sont de type SISO où un seul paramètre de procédé est contrôlé. Ce type de modèles ne permet pas de prendre en compte les effets des autres variables du procédé. Par contre, il présente un avantage d'un point de vue industriel. Les modèles proposés peuvent être généralisés à n'importe quel procédé de dépôt. Donc, le temps de déploiement de ces modèles en FAB est rapide.

4.2.3.4 Dépôt LPCVD

U. POTTHOFF de la société INFINEON a publié les premiers travaux d'industrialisation de la boucle feed-forward du procédé de dépôt LPCVD (ajustement de lot à lot). Le modèle proposé est de type MIMO. Les variables du procédé à ajuster sont le temps du procédé et la température. Les variables à contrôler sont les épaisseurs sur plusieurs points de la plaque. Le choix de ces paramètres s'explique par la particularité du procédé LPCVD. Ce dernier est caractérisé par un gradient de température dans les différentes zones de four.

4.2.3.5 Modélisation des mécanismes du dépôt HDP-CVD

Les premiers travaux sur les mécanismes de dépôt d'oxyde par le procédé HDP-CVD ont été étudiés par S-M. SUH de l'université du Minnesota [SUH 2003]. Il s'agissait de modéliser le phénomène de nucléation en phase gazeuse dans un réacteur plasma à couplage inductif. Le mécanisme de cinétique chimique des particules a été décrit en détail pour enfin développer le modèle de la croissance de l'oxyde à partir de silane et d'oxygène. Le réacteur a été modélisé en 1 D et la distribution spatiale des concentrations des espèces employées a été calculée. Une analyse chimique a été faite ensuite pour déterminer les réactions dominantes. La réaction chimique à la surface entre les espèces neutres du plasma est la réaction⁵⁶ la plus dominante de la croissance d'oxyde. Les conclusions de cette étude montrent que la température des gaz et la pression opératoire (très basse ; de l'ordre de 5 à 20 mTorr) [SUH 2002] sont les paramètres les plus contributeurs au phénomène de croissance d'oxyde.

⁵⁵ Opérations pour qualifier un procédé

⁵⁶ Le procédé de dépôt chimique en phase vapeur excité par plasma se décompose en 5 mécanismes indépendants :

- La réaction chimique à la surface entre les espèces neutres du plasma
- Le dépôt induit par les espèces ioniques. Ce mécanisme est principalement gouverné par la densité des espèces ionisées.
- Le mécanisme de pulvérisation (gravure des espèces déposées) principalement corrélé à la densité des espèces ionisées ainsi que de leur énergie (contrôlée par la puissance Bias)
- La re-déposition directe des espèces engendrée par la pulvérisation
- La re-déposition indirecte des espèces engendrée par la pulvérisation.

4.2.3.6 Positionnement de nos travaux de recherches

Nous constatons qu'à travers l'état de l'art présenté dans les paragraphes précédents sur l'implémentation des boucles de régulation sur le procédé CVD, à part les travaux proposés par J. STEFANI et S-M.SUH, il n'y a pas eu beaucoup de travaux ayant traité le procédé HDP-CVD STI. Nous estimons que les modèles proposés avec une seule variable à réajuster sont insuffisants pour contrôler les variabilités du procédé HDP-CVD car, c'est un ensemble de paramètres qui influencent la variabilité de l'épaisseur déposée. Nous proposons donc un modèle multivarié qui tient compte de l'effet de chaque paramètre identifié comme critique. Pour cela, il faut bien identifier la criticité de chaque paramètre.

Dans les paragraphes qui suivent, on exposera les différents travaux qui nous ont permis d'obtenir le modèle du procédé HDP-CVD au niveau de la brique STI. Le modèle recherché sera le cœur de la boucle de régulation R2R de type « feed-back ». Le modèle doit donc être prédictif. La figure 4.15 montre le schéma synoptique de la boucle de régulation R2R. Le principe de cette boucle de type feed back est de réajuster le(s) paramètre(s) critique(s) du lot (n+1) après avoir pris connaissance des résultats de métrologie du lot (n).

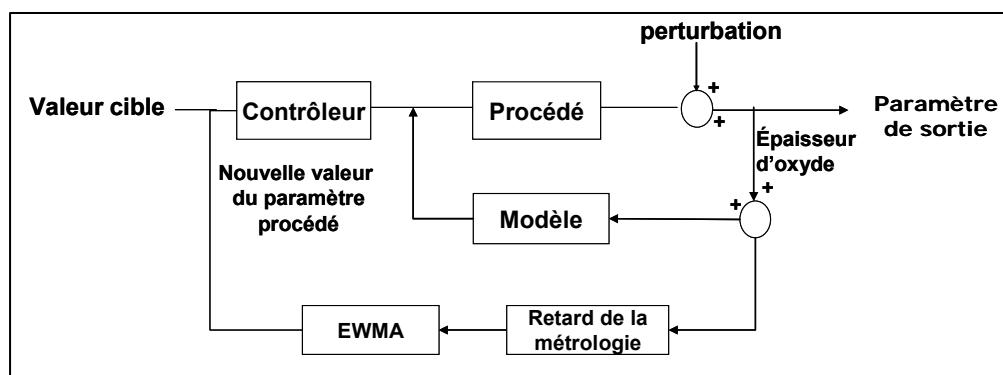


Figure 4.15- Schématisation de la boucle de régulation R2R

Donc il faut créer des stratégies de contrôle lot à lot en tenant compte des mesures disponibles. L'objectif final est de réduire la dispersion du dépôt, de s'approcher de la valeur cible recherchée et de compenser les dérives.

4.2.4 Variabilité du procédé HDP-CVD

4.2.4.1 Types de variabilité

Des analyses ont été faites sur un historique de données en vue de classifier le type de variabilité rencontrée.

4.2.4.1.a Variabilité lot à lot

Les variabilités du procédé HDP-CVD lot à lot traduisent les différences entre les lots. Ces différences sont constatées au niveau des paramètres de mesures (mesures de l'épaisseur d'oxyde ainsi que l'uniformité d'oxyde déposé généralement exprimée par son étendue). La figure 4.16 montre un exemple des mesures de l'épaisseur d'oxyde. Les points correspondent

à des points de mesures individuelles (sites de mesure). Ceci montre la présence d'une variabilité de type lot à lot. Comme nous pouvons le constater sur la figure 4.16, l'étendue de l'épaisseur varie de lot à lot. Pour le lot 4, l'étendue de l'épaisseur est de 130Å, alors que pour le lot 21, l'étendue de l'épaisseur est de 140Å.

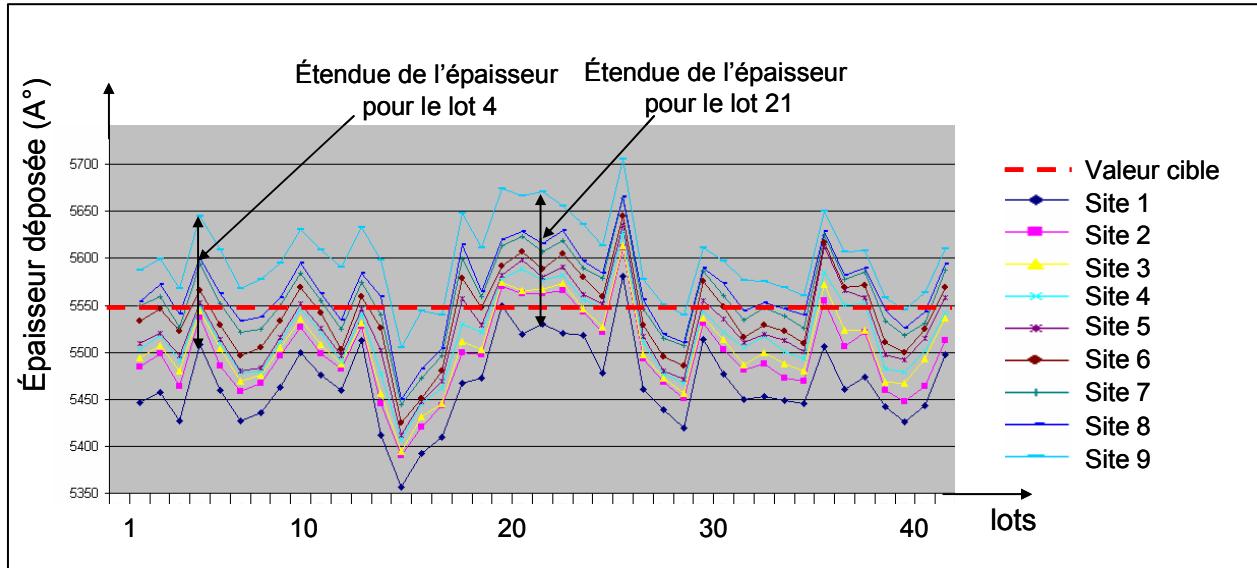


Figure 4.16- Evolution de l'épaisseur déposée lot à lot pour différents sites de mesures.

4.2.4.1.b Variabilité plaque à plaque

On observe également une variabilité de type plaque à plaque comme cela est illustré sur la figure 4.17. Des mesures d'oxyde ont été effectuées sur 4 lots processés consécutivement sur un équipement de dépôt HDP-CVD. Comme nous pouvons le constater sur la figure 4.17, l'écart entre la valeur cible et la valeur de mesure varie de plaque à plaque. Cet écart est de 40 Å pour la première plaque, alors que pour la plaque 40, cet écart est de 105Å.

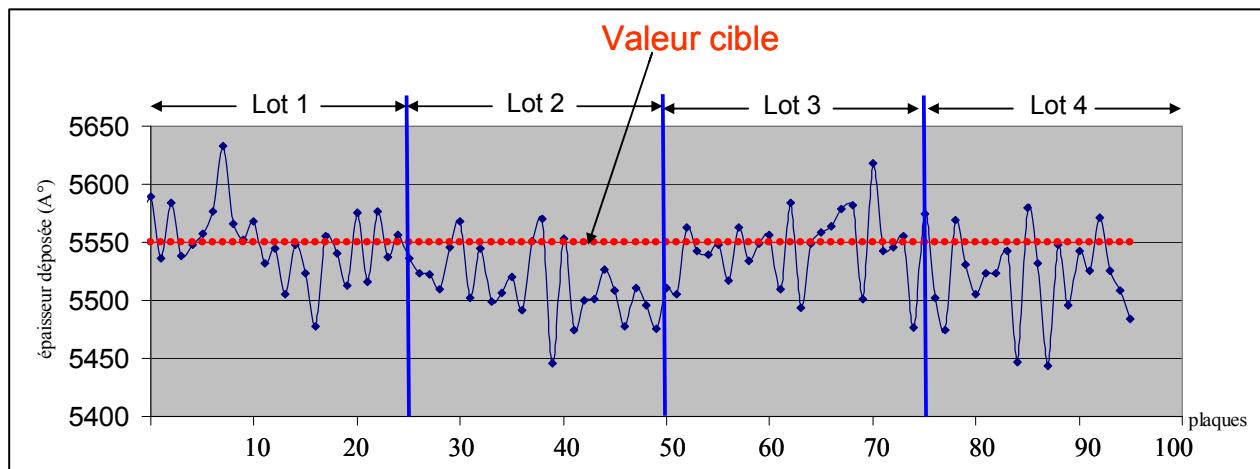


Figure 4.17- Evolution de l'épaisseur totale pendant les 2 steps.

Effet de la première plaque

Une analyse fine a été réalisée sur un historique de données (4 lots processés consécutivement sur un équipement de dépôt HDP-CVD) pour déterminer les causes de la variabilité plaque à plaque. Les résultats de cette analyse ont montré un écart entre les mesures de la première et la dernière plaque comme nous pouvons le constater sur la figure 4.18. Le nettoyage de la chambre est déclenché après avoir traité 5 plaques, les barres rouges sur le graphe indiquent la fin du nettoyage. Nous observons sur la figure 4.18 qu'il y a un écart entre l'épaisseur de la première plaque (processée juste après la fin de nettoyage) et la dernière plaque traitée (cinquième plaque processée juste après la fin de nettoyage). Cet écart est de 100 Å.

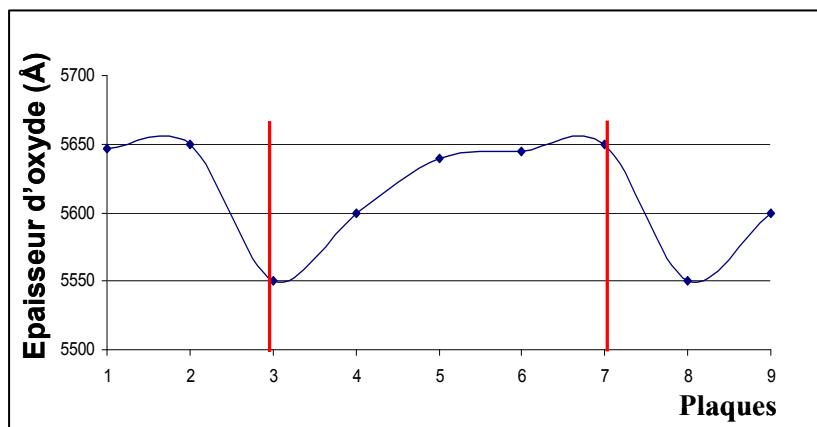


Figure 4.18- Evolution de l'épaisseur d'oxyde avant et après le nettoyage de la chambre

Nous proposons une méthode simple pour éliminer l'effet de la première plaque. Il s'agit de rajouter des temps additionnels de dépôt Δt_1 et Δt_2 pour les deux premières plaques processées respectivement après le nettoyage de la chambre. Connaissant la vitesse de dépôt moyenne et l'écart entre les deux premières plaques processées et la dernière plaque, on peut calculer avec précision les valeurs des temps Δt_1 et Δt_2 . Ainsi, on peut réduire la variabilité de plaque à plaque de 20 Å au lieu de 100 Å. Les temps du procédé de dépôt optimisés par plaques sont donnés dans le tableau 4.1.

Ordre des plaques après le nettoyage	Temps du procédé (secondes)
1	$t_0 + \Delta t_1$
2	$t_0 + \Delta t_2$
3	t_0
4	t_0
5	t_0

Tableau 4.1- Temps du procédé par plaques

4.2.4.1.c Variabilité intra- plaque

La figure 4.19 représente la reconstruction de la surface pour une plaque à faible dispersion d'épaisseur d'oxyde. Nous observons sur la figure 4.19 plusieurs régions au niveau

de la plaque. Nous distinguons la région centrale avec une épaisseur moyenne de 5240Å ainsi que la région proche des bords de la plaque avec une épaisseur moyenne de 5200Å. On parle donc d'une variabilité centre- bords.

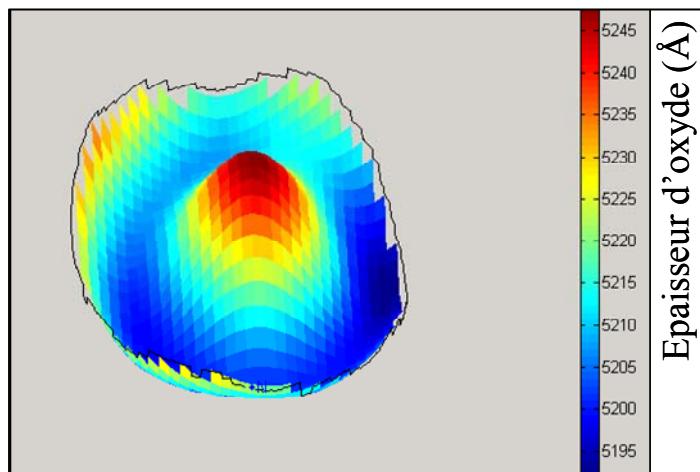


Figure 4.19- Exemple illustrant la variabilité intra-plaque de l'épaisseur d'oxyde déposé.

Ces variabilités ne peuvent pas être observées en disposant de mesures faites sur une cartographie contenant seulement 9 points de mesures. L'analyse plus fine du profil de dépôt ne peut donc être réalisée qu'en disposant d'une cartographie contenant un grand nombre de points sur la plaque. Le profil de dépôt illustré sur la figure 4.19 est la reconstruction de la surface à partir de 49 points de mesure.

En comparant les différentes dispersions constatées (lot à lot, plaque à plaque et intra-plaque) avec les limites des cartes de contrôle (épaisseur d'oxyde, étendue), les dispersions sont très importantes et peuvent provoquer des dérives. Donc une régulation s'impose pour réduire les dispersions. Pour la variabilité intra- plaque, on verra par la suite qu'on pourra la réduire par l'intermédiaire du modèle de l'étendue. Donc, seules les variabilités lot à lot et intra- plaques sont traitées dans ce manuscrit.

Dans la dernière partie du chapitre 1, nous avons détaillé la méthodologie établie pour le déploiement des boucles de régulation R2R. Nous allons l'appliquer à tous les procédés. Après avoir identifié la nature de la variabilité, il faut identifier les paramètres de procédé critiques.

4.2.4.2 Analyse multivariée

Des analyses multivariées ont été effectuées sur un historique de données (6 mois) de mesures et de paramètres équipements

4.2.4.2.a Analyse en composantes principales (ACP) de l'épaisseur déposée

Les résultats des analyses en composantes principales sont illustrés sur la figure 4.20 ou chaque point correspond à une plaque processée. La figure 4.20 montre la position de chaque point par rapport à la construction des composantes principales (axe 1 et axe 2) et par rapport à l'ensemble des données. Comme nous l'avons bien expliqué au chapitre 1, l'axe 1 correspond à l'axe qui contient la plus grande quantité de dispersion et l'axe 2 correspond à la

seconde grande dispersion. Nous observons sur la figure, 3 populations. Après vérification des résultats, il apparaît que les 3 populations correspondent aux données des 3 chambres de l'équipement de dépôt.

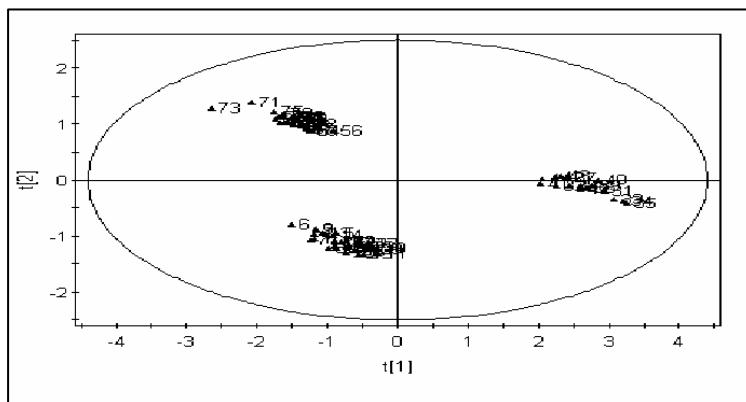


Figure 4.20- Résultats de l'analyse des composantes principales

Ce phénomène est assez connu dans l'industrie des semi-conducteurs, c'est l'effet chambre. Il se traduit généralement par une différence de comportements des chambres du même équipement [HOYER 2004]. Cette différence de comportement correspond à différentes cinétiques de dépôt. Les boucles de régulation vont servir à ne plus avoir d'effet chambre. L'emploi de la méthode ACP peut être un bon moyen pour piloter la maintenance préventive. Avec un effet de chambre supprimé, on a sur le graphe qu'une seule population regroupant toutes les données des 3 chambres.

4.2.4.2.b Méthodes des moindres carrés (PLS)

D'autres analyses multivariées ont été effectuées sur un ensemble de données des paramètres en vue de déterminer les paramètres qui influencent la variation de l'épaisseur déposée. Les résultats des analyses de la méthode des moindres carrés (PLS) sont donnés dans la figure 4.21.

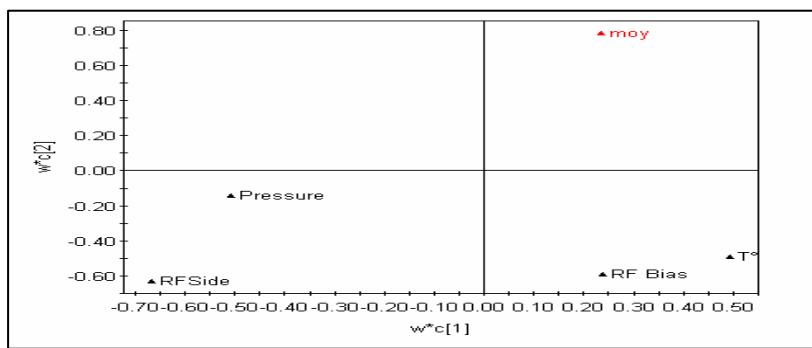


Figure 4.21- Résultats de l'analyse PLS

Comme nous l'avons déjà expliqué dans le chapitre 1 (partie 1.3.2.2), la méthode des moindres carrés (PLS) est une méthode qui permet d'estimer les coefficients d'un polynôme d'ajustement entre les variables à expliquer Y_k (dans ce cas l'épaisseur d'oxyde déposée) et un ensemble de variables explicatives X_i (paramètres du procédé de dépôt HDPCVD). Dans

notre cas, Y est symbolisée sur la figure par « moy », cette dernière représente la moyenne de l'épaisseur déposée. Les variables process sont la puissance RF bias, la pression, la puissance RF side et la température. Les variables présentes dans le graphe sont pondérées par des coefficients. La valeur des coefficients est comprise entre 0 et 1. Si la valeur des coefficients est proche de 0, ces variables sont peu influentes sur l'épaisseur d'oxyde déposée. Par contre, la pression, les puissances RF Side et Bias ainsi que la température sont les variables influentes sur l'épaisseur d'oxyde déposée.

4.2.4.3 Identification des paramètres critiques

4.2.4.3.a Paramètres issus de l'analyse PLS

Les résultats des analyses multivariées montrent que les paramètres critiques sont : la pression de la chambre, la puissance RF side, la puissance RF Bias et la température. En vue de construire un modèle simple, il est souhaitable de réduire d'avantage le nombre de paramètres. Pour cela, il est nécessaire de comprendre d'un point de vue physique le degré de criticité de chaque paramètre sur l'épaisseur déposée.

La pression

La pression influence la gravure de manière significative en modifiant le libre parcours moyen des ions d'argon. La pression agit directement sur l'énergie des atomes et donc sur l'efficacité de la gravure [MUNGEKAR 2005]. Ceux ci seront moins accélérés dans la gaine du plasma si la pression augmente. La variation de la gravure impacte directement le rapport D/S. Pour que ce dernier soit dans les limites de spécifications, la baisse de la vitesse de gravure requiert donc une baisse de la vitesse de dépôt D et donc la diminution des débits de gaz des procédés par exemple.

La pression opératoire sous laquelle s'effectue le procédé de dépôt se dégrade avec le temps. Cette variation est observée après un temps d'utilisation important de l'équipement de dépôt. Ceci peut s'expliquer par le fait que, pendant le procédé HDP-CVD, un film polymère se forme sur les parois du réacteur (par la diffusion et la réaction de certaines particules avec les parois de l'équipement). Ce phénomène s'accentue avec le temps d'utilisation de l'équipement [Joubert 2004]. Lors des maintenances préventives⁵⁷ (PM), on procède à un nettoyage des parois des chambres de l'équipement pour enlever les résiduels chimiques et toutes les traces de films polymères. La figure 4.22 compare la pression d'une chambre avant la maintenance préventive (pendant ce temps on a traité 30 lots) avec la pression juste après la maintenance préventive (pendant ce temps on a traité 10 lots).

⁵⁷ Maintenances périodiques faites sur l'équipement pour assurer son bon fonctionnement (remplacement des pièces usées par exemple)

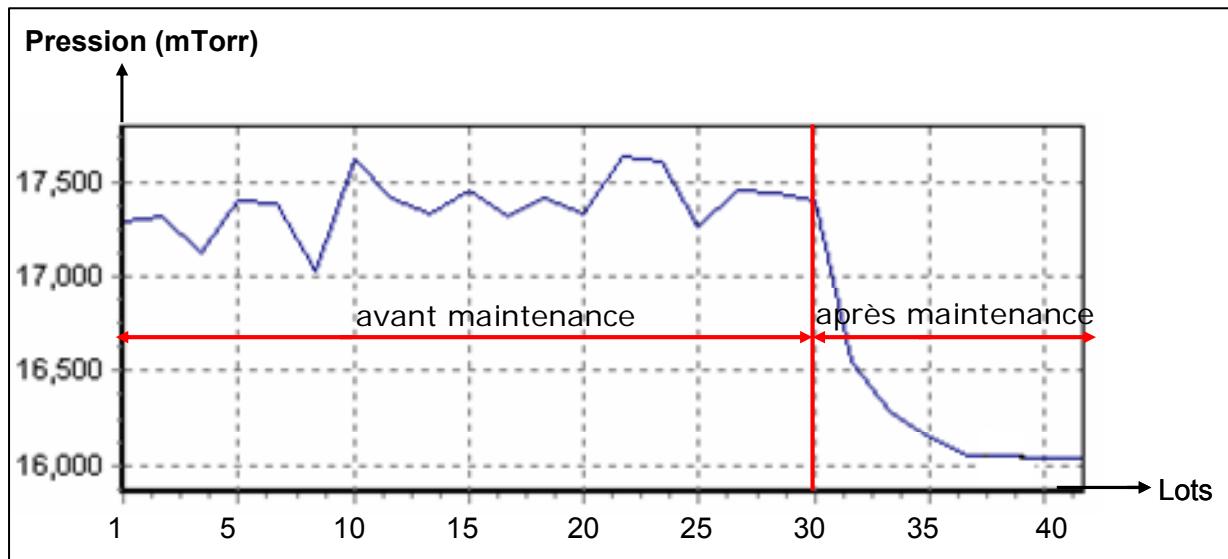


Figure 4.22- Evolution de la pression avant et après une maintenance préventive.

Comme cela est illustré sur la figure 4.22, la valeur de la pression opératoire est très proche de la valeur cible (16 mTorr), alors qu'avant la maintenance préventive, la valeur de la pression opératoire s'écarte de la valeur cible avec un écart moyen de 1,4 m Torr. Pour réduire l'effet des parois et donc réduire la variation de la pression de la chambre avec le temps, on procède à un nettoyage périodique des chambres de l'équipement après avoir traité n plaques. Le déclenchement du nettoyage de la chambre a lieu après avoir traité 5 plaques comme cela est illustré sur la figure 4.23. Nous avons suivi l'évolution de la pression d'une chambre après le nettoyage. La figure 4.23 illustre cette variation de plaque à plaque. Il apparaît que la pression atteint une certaine valeur stable après un temps correspondant au temps de traitement des deux premières plaques.

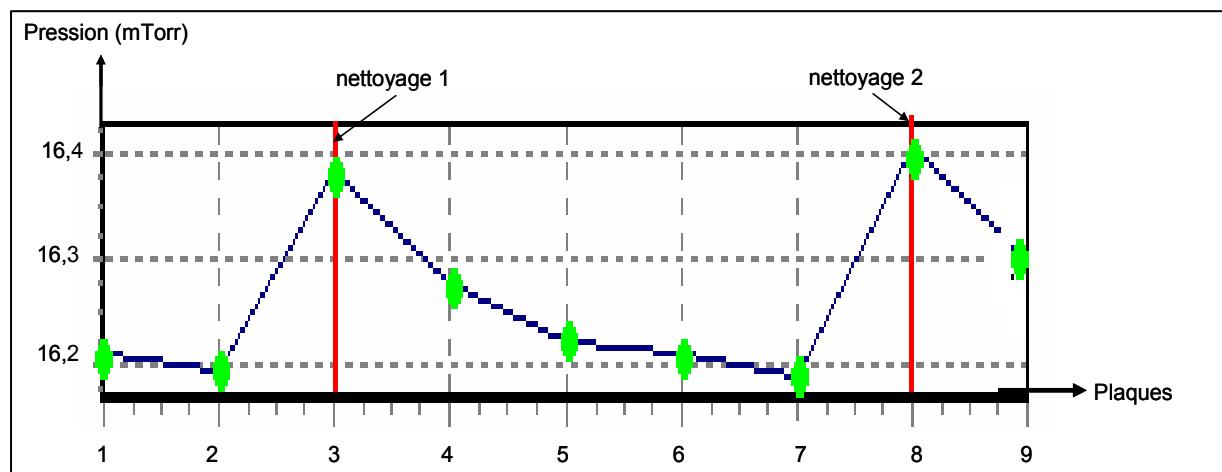


Figure 4.23- Evolution de la pression avant et après le nettoyage de la chambre

Il est évident que la variation de la pression de plaque à plaque va engendrer une variation de la vitesse de pulvérisation. Ce qui va se traduire par une variation de l'épaisseur déposée de plaque à plaque.

La puissance RF Bias

Pour que le rapport D/S soit dans les limites de contrôle, des tâches de qualifications sont effectuées périodiquement où quelques paramètres sont ajustés manuellement. Parmi ces paramètres, la puissance RF Bias. La figure 4.24 montre un exemple de tâche de qualification faite sur une chambre d'un équipement de dépôt où la puissance RF Bias opératoire avant la tâche de qualification était de 3500 W. Suite à cette tâche de qualification, la nouvelle valeur de la puissance RF Bias choisie est de 3650 W.

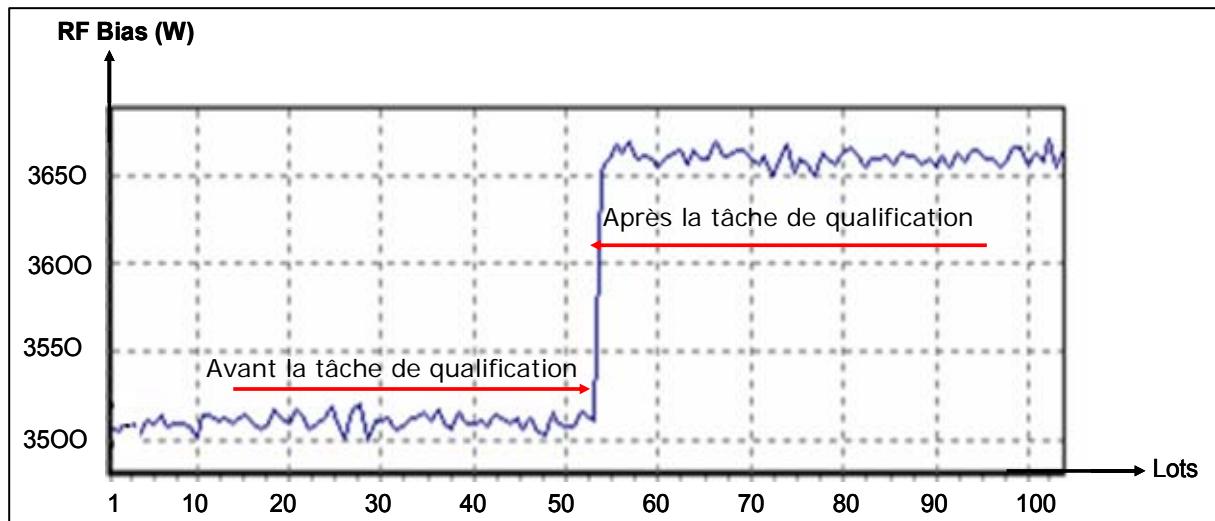


Figure 4.24- Evolution de la puissance RF Bias avant et après la tâche de qualification

Ce changement de la puissance RF Bias impacte énormément le dépôt du fait que ce paramètre intervient directement dans la détermination du flux ionique qui attaque la plaque. La figure 4.25 montre que l'épaisseur moyenne des 15 derniers lots traités avant la tâche de qualification était de 5550 Å. Après la tâche de qualification, la valeur de l'épaisseur moyenne des 15 premiers lots est de 5480Å.

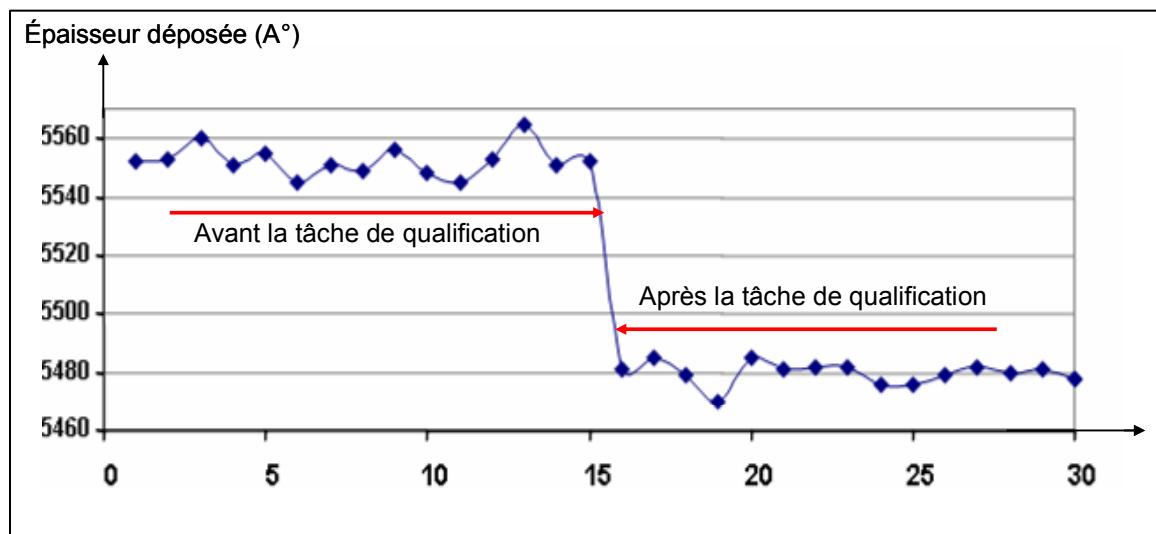


Figure 4.25- Evolution de l'épaisseur d'oxyde avant et après la tâche de qualification

Les changements opérés sur ce paramètre sur une large plage (3300-3700 W) ont une influence importante sur l'épaisseur de dépôt et ce paramètre est sélectionné comme paramètre à considérer dans la conception du plan d'expériences.

La température

Dans le cas du STI, une haute température est requise afin de fournir une bonne mobilité aux précurseurs dans la tranchée. La température de la plaque résulte d'un équilibre thermique entre le flux de chaleur généré par la pulvérisation des ions, la perte de chaleur au niveau de l'arrière de la plaque et aussi la perte de chaleur dans la chambre. Munekar et al [MUNGEKAR 2003] ont démontré que la température de la plaque est corrélée à la vitesse de pulvérisation pour les gaz (SF₆, CF₄, Ar, Cl₂...) utilisés en gravure sèche. Pour cette raison, la température ne sera pas prise en considération dans cette étude car la puissance RF Bias est bien corrélée à la vitesse de pulvérisation.

La puissance RF side

Le remplissage des tranchées STI nécessite une haute densité ionique, donc des valeurs de la puissance RF Side importantes. La vitesse de dépôt est proportionnelle à la puissance RF side [BABAYAN 2001]. L'augmentation de la densité ionique peut être atteinte aussi en augmentant le débit des gaz (silane). Afin de minimiser le nombre des paramètres pour cette étude, la RF side ne sera pas prise en considération car les débits de gaz Silane side et top sont déjà sélectionnés.

4.2.4.3.b Paramètres déduits directement

Certains paramètres sont sélectionnés directement [BELHARET 2006 a] comme les débits de gaz silane side et silane top qui sont responsables respectivement de la vitesse de dépôt et de l'uniformité. Lors de la régulation, il faut faire très attention à la stoechiométrie du film d'oxyde et donc respecter les limites de spécifications de la qualité d'oxyde. Pour cela, l'oxygène doit être en excès par rapport au silane afin de ne pas limiter la réaction. Le rapport débit O₂/débit Silane Side est de 2 environ.

4.2.5 Modélisation du procédé HDP-CVD

Une fois que les paramètres critiques sont identifiés, l'étape suivante est de construire le plan d'expériences. Les modèles recherchés sont celui de l'épaisseur et de l'étendue. On cherche à réduire la dispersion lot à lot des épaisseurs d'oxyde tout en faisant attention à ne pas dégrader l'étendue de l'épaisseur (range). Il faut donc concevoir une série d'expériences (Plan d'expériences), puis exécuter ce plan dans les conditions expérimentales définies. Une analyse statistique sera ensuite faite pour vérifier l'efficacité de ce modèle. Suite à ce résultat, le modèle sera testé pour pouvoir le comparer avec des données expérimentales.

4.2.5.1 Le plan d'expériences

Dans cette partie on cherchera à modéliser l'épaisseur moyenne et l'étendue de l'épaisseur avec les paramètres du procédé de dépôt. Nous chercherons à minimiser le nombre de réponses ou du moins la complexité de la réponse étudiée de manière à ne pas faire intervenir une exploitation trop lourde du modèle.

4.2.5.1.a La conception du modèle

La solution la plus simple consiste à étudier directement l'épaisseur moyenne et son étendue en fonction des facteurs influents déjà identifiés précédemment. Le plan d'expériences conçu est de type central composite à faces centrées. Le nombre de variables est de 3 (RF Bias, SiH₄ side et SiH₄ Top). Les valeurs de la puissance RF Bias varient de 3300 à 3700 W, les débits des gaz silane side et top varient de 165 à 195 Sccm et de 22 à 28 Sccm respectivement. Le choix des plages des paramètres se justifie en partie par les limites de tolérance de fonctionnement de l'équipement. Le nombre d'expériences est de l'ordre de 20 dont 6 expériences pour tester la répétitivité de la réponse. Le tableau 4.2 résume les différentes conditions expérimentales (paramètres fixes et paramètres variables).

Plaque	RF Bias(W)	SiH ₄ side (Sccm)	SiH ₄ Top(Sccm)	Pression (mTorr)	RF Side (W)	RF Top(W)	Ar Side (Sccm)	Ar Top(Sccm)	O2(Sccm)
1	3300	165	22	16	3100	1300	78	12	390
2	3300	165	28	16	3100	1300	78	12	390
3	3300	185	22	16	3100	1300	78	12	390
4	3300	195	28	16	3100	1300	78	12	390
5	3700	165	22	16	3100	1300	78	12	390
6	3700	165	28	16	3100	1300	78	12	390
7	3700	190	22	16	3100	1300	78	12	390
8	3700	190	28	16	3100	1300	78	12	390
9	3300	180	25	16	3100	1300	78	12	390
10	3700	180	25	16	3100	1300	78	12	390
11	3500	165	25	16	3100	1300	78	12	390
12	3500	190	25	16	3100	1300	78	12	390
13	3500	180	22	16	3100	1300	78	12	390
14	3500	180	28	16	3100	1300	78	12	390
15	3500	180	25	16	3100	1300	78	12	390
16	3500	180	25	16	3100	1300	78	12	390
17	3500	180	25	16	3100	1300	78	12	390
18	3500	180	25	16	3100	1300	78	12	390
19	3500	180	25	16	3100	1300	78	12	390
20	3500	180	25	16	3100	1300	78	12	390

Tableau 4.2- Récapitulatif des conditions expérimentales du plan d'expériences

En vue d'éviter les variabilités de la pression de plaque à plaque, nous avons modifié la fréquence de déclenchement du nettoyage. Ce dernier intervient après chaque plaque traitée. De ce fait, la pression opératoire de l'équipement (16mTorr) est la même pour chaque plaque traitée. La valeur du débit d'oxygène (390 Sccm) est choisie pour avoir un rapport (débit oxygène/ débit silane side) compris entre 2 et 2,36. Ce rapport permet d'avoir un film de bonne stoechiométrie.

4.2.5.1.b Mesures

La cartographie de mesure choisie est de type « couronne » qui apparaît comme la meilleure solution pour couvrir toutes les zones de dépôt (effet centre- bords des plaques) car les points de mesure sur la plaque sont répartis de façon à couvrir toutes les zones de dépôt. La représentation de la cartographie « couronne » 49 points est illustrée sur la figure 4.26.

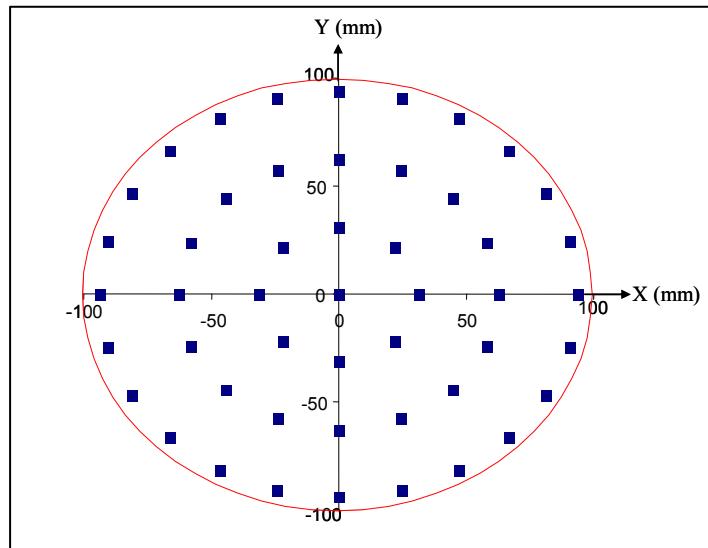


Figure 4.26- Cartographie de mesures en forme de couronne de 49 points

4.2.5.2 Le modèle de l'épaisseur

4.2.5.2.a Résultats du DOE

Le modèle de l'épaisseur obtenu est donné par la relation suivante:

$$Y_1 = a_1 * \text{RFBias} + a_2 * \text{SiH}_4\text{Side} + a_3 * \text{SiH}_4\text{Top} + a_{11} * \text{RFBias} * \text{RFBias} + a_{12} * \text{RFBias} * \text{SiH}_4\text{Side} + a_{13} * \text{RFBias} * \text{SiH}_4\text{Top} + a_{21} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Side} + a_{23} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Top} + a_{33} * \text{SiH}_4\text{Top} * \text{SiH}_4\text{Top}$$
 (4.8)

Où: Y_1 est l'épaisseur d'oxyde déposée

$a_1, a_{11}, a_2, a_{22}, a_3, a_{33}$ sont des constantes associées aux paramètres RF Bias, SiH₄ Side et SiH₄ Top.

$a_{12}, a_{13}, a_{21}, a_{23}, a_{31}$ et a_{32} sont les constantes associées aux interactions entre les 3 paramètres.

Nous avons analysé les résultats de DOE avec l'aide du logiciel SAS. Comme nous l'avons bien expliqué au chapitre 1 (partie 1.3.3.2.c), les résultats des indices statistiques sont déterminants pour juger la qualité du DOE effectué. Les premières analyses ont porté sur l'analyse des valeurs de la probabilité (P-value). Le tableau 4.3 donne les résultats des valeurs de (P-value) obtenus avec l'aide du logiciel SAS.

Effet	Valeur de P-value	Remarques
RF Bias	<0,0001	Significatif
SiH ₄ Side	<0,0001	Significatif
SiH ₄ Top	<0,0001	Significatif
RF Bias* RF Bias	0,2885	Rejeté car P-value>0,05
RF Bias* SiH ₄ Side	0,0189	Significatif
RF Bias* SiH ₄ Top	0,5535	Rejeté car P-value>0,05
SiH ₄ Side* SiH ₄ Side	0,0055	Significatif
SiH ₄ Side* SiH ₄ Top	0,0396	Significatif
SiH ₄ Top* SiH ₄ Top	0,7223	Rejeté car P-value>0,05

Tableau 4. 3- Valeurs de P-value obtenues sous SAS.

D'après le tableau 4.3, RF Bias* RF Bias, RF Bias* SiH₄ Top et SiH₄ Top* SiH₄ Top ne sont pas intégrés dans le modèle car leurs valeurs de P-value associées sont >0,05. L'épaisseur d'oxyde déposée peut être exprimée alors par la relation suivante :

$$Y_1 = a_1 * \text{RFBias} + a_2 * \text{SiH}_4\text{Side} + a_3 * \text{SiH}_4\text{Top} + a_{12} * \text{RFBias} * \text{SiH}_4\text{Side} + a_{22} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Side} + a_{23} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Top} \quad (4.9)$$

Le modèle de l'épaisseur établi est obtenu en fonction des paramètres critiques ainsi qu'avec les interactions entre les paramètres. Maintenant, il est possible de comprendre l'impact de chaque paramètre critique sur l'épaisseur. La réponse de surface de l'épaisseur d'oxyde est donnée dans la figure 4.27.

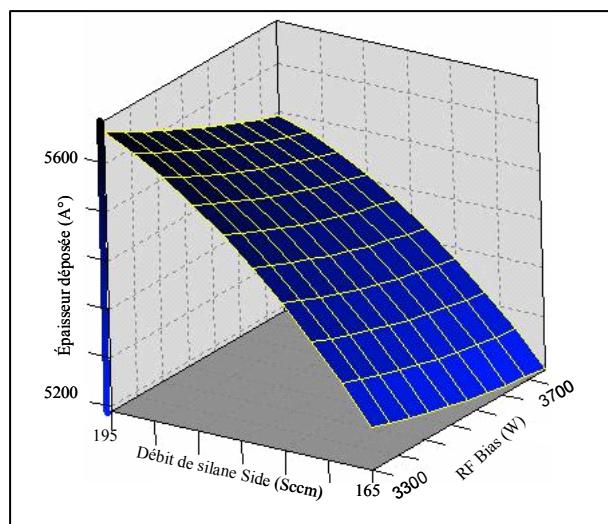


Figure 4.27- Evolution de l'épaisseur déposée en fonction de la RF Bias et du silane side pour une valeur de silane top de 25 Sccm.

La figure 4.27 montre l'évolution de l'épaisseur déposée en fonction des paramètres critiques. On observe que l'épaisseur décroît avec l'augmentation de la RF Bias. Mais, quand le débit de silane croît, l'épaisseur croît également. A partir de ce résultat, nous pouvons conclure que la dérive d'un paramètre (exemple la puissance RF Bias) peut être compensée

par l'ajustement des valeurs des autres variables (SiH_4 Side et SiH_4 Top) pour garantir la valeur cible de l'épaisseur d'oxyde déposée.

4.2.5.2.b Simulation du modèle

En vue de valider le modèle établi, une série de simulations est indispensable pour tester le modèle, en prenant tous les cas qui peuvent se produire en production. Les simulations ont été effectuées à l'aide du logiciel SAS. La figure 4.28 illustre un exemple de l'interface graphique SAS.

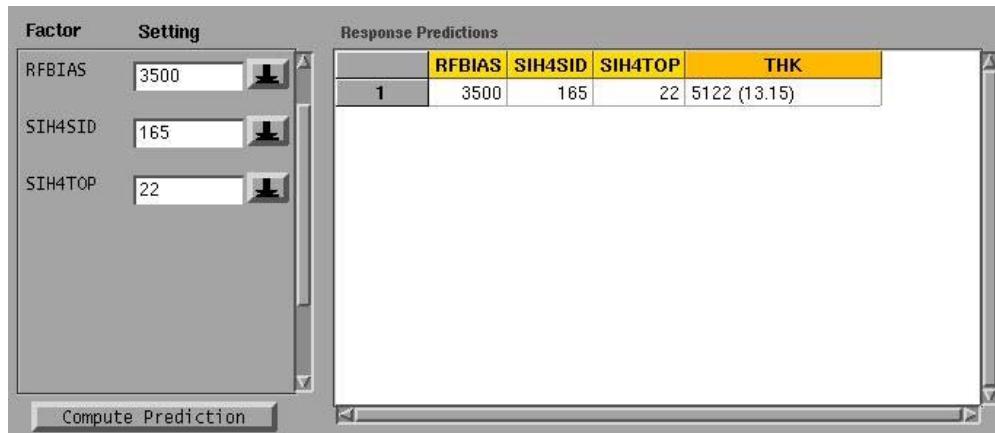


Figure 4.28- Exemple d'une simulation à l'aide du logiciel SAS.

Comme nous pouvons le constater sur la figure 4.29, nous pouvons simuler n'importe quelle épaisseur d'oxyde à partir des conditions du procédé. Dans notre cas, nous faisons varier les valeurs des paramètres critiques à savoir la puissance RF Bias (de 3300W à 3700W) et les débits de silane side (de 165 Sccm à 195 Sccm) et de silane top (de 22 Sccm à 28 Sccm). Une simulation est jugée bonne lorsque les simulations ont été effectuées avec des conditions similaires à celles des conditions expérimentales. De ce fait, les résultats des simulations seront comparés avec les données des mesures ou résultats expérimentaux. Nous avons comparé les conditions de mesures des épaisseurs d'oxyde avec des valeurs de simulations. Les résultats de cette comparaison sont illustrés sur la figure 4.29.

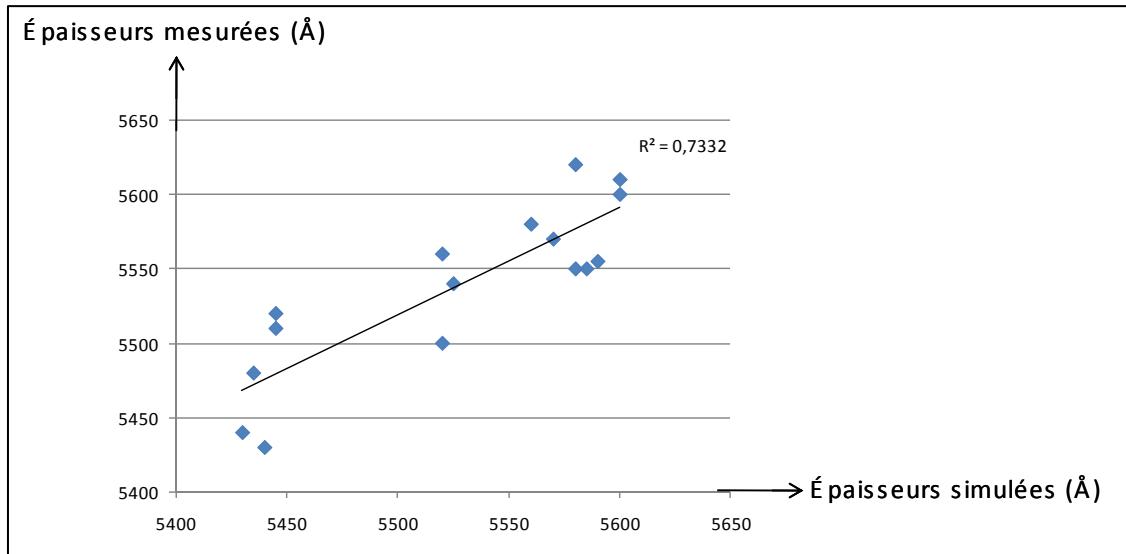


Figure 4.29- Comparaison entre les données de simulation et les données expérimentales

Les premiers résultats des simulations sont illustrés sur la figure 4.29. On observe que les résultats de simulations sont assez prédictifs car on arrive à simuler presque la même valeur trouvée en expérimental. Comme nous pouvons le constater sur la figure, l'erreur moyenne de prédiction (écart entre la valeur de mesure et de la valeur prédictive est de 30 Å).

4.2.5.3 Le modèle de l'étendue de l'épaisseur (range)

La deuxième réponse est l'étendue de l'épaisseur. L'uniformité du dépôt est souvent exprimée par l'indicateur de l'étendue d'épaisseur qui est la différence entre la valeur maximale de l'épaisseur déposée et sa valeur minimale, aussi appelée « range ».

4.2.5.3.a Les résultats du DOE

Le modèle de l'étendue d'épaisseur est donné par la relation suivante :

$$Y_2 = a_1 * \text{RFBias} + a_2 * \text{SiH}_4\text{Side} + a_3 * \text{SiH}_4\text{Top} + a_{11} * \text{RFBias} * \text{RFBias} + a_{12} * \text{RFBias} * \text{SiH}_4\text{Side} + a_{13} * \text{RFBias} * \text{SiH}_4\text{Top} + a_{22} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Side} + a_{23} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Top} + a_{33} * \text{SiH}_4\text{Top} * \text{SiH}_4\text{Top} \quad (4.10)$$

Où: Y_2 est l'étendue de l'épaisseur d'oxyde déposée

$a_1, a_{11}, a_2, a_{22}, a_3, a_{33}$ sont des constantes associées aux paramètres RF Bias, SiH₄ Side et SiH₄ Top.

$a_{12}, a_{13}, a_{21}, a_{23}, a_{31}$ et a_{32} sont les constantes associées aux interactions entre les 3 paramètres.

Nous avons procédé à l'analyse des valeurs de la probabilité (P-value). Le tableau 4.4 donne les résultats des valeurs de (P-value) obtenues avec l'aide du logiciel SAS.

Effet	Valeur de P-value	Remarques
RF Bias	0,2974	Rejeté car P-value>0,05
SiH ₄ Side	0,0004	Significatif
SiH ₄ Top	<0,0001	Significatif
RF Bias* RF Bias	0,4869	Rejeté car P-value>0,05
RF Bias* SiH ₄ Side	0,1468	Rejeté car P-value>0,05
RF Bias* SiH ₄ Top	0,3187	Rejeté car P-value>0,05
SiH ₄ Side* SiH ₄ Side	0,4550	Rejeté car P-value>0,05
SiH ₄ Side* SiH ₄ Top	0,0021	Significatif
SiH ₄ Top* SiH ₄ Top	0,0826	Rejeté car P-value>0,05

Tableau 4.4- Valeurs de P-value obtenues sous SAS

D'après le tableau 4.4, seuls les paramètres SiH₄ Side, SiH₄ Top et SiH₄ Side* SiH₄ Top sont sélectionnés dans le modèle. Les autres paramètres sont rejetés car les valeurs de P-value associées sont >0,05. L'étendue de l'épaisseur d'oxyde peut être exprimée par la relation suivante :

$$Y_2 = a_2 * \text{SiH}_4\text{Side} + a_3 * \text{SiH}_4\text{Top} + a_{23} * \text{SiH}_4\text{Side} * \text{SiH}_4\text{Top} \quad (4.11)$$

Un modèle est obtenu. Il met en relation l'étendue avec les paramètres d'entrée qui sont le débit de silane side et top. La figure 4.30 montre l'évolution de l'étendue de l'épaisseur déposée en fonction des paramètres critiques. On observe que la puissance RF Bias n'a aucun impact sur l'étendue. Notons que toutes les plaques ont été mesurées sur une même cartographie.

La réponse de surface de l'étendue de l'épaisseur d'oxyde est donnée dans la figure 4.30.

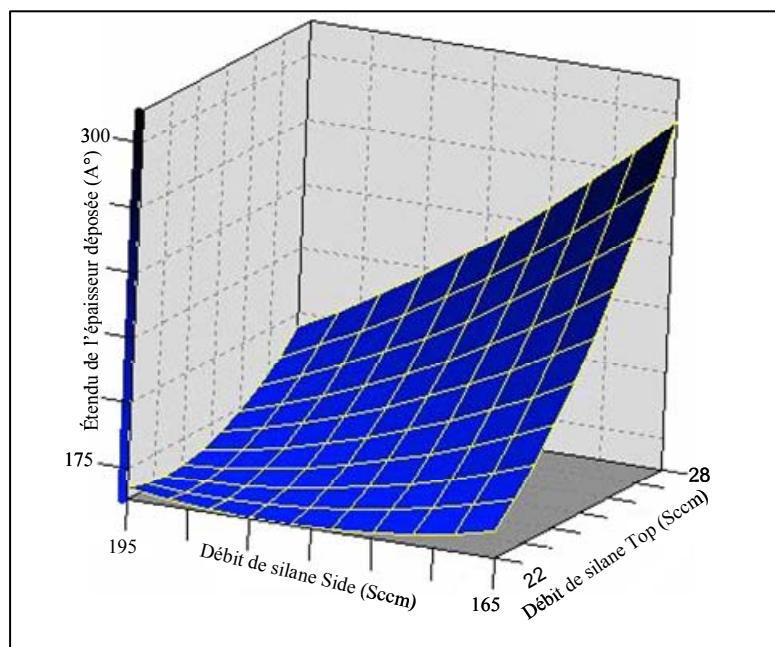


Figure 4.30- Evolution de l'étendue de l'épaisseur déposée en fonction du Silane Top et du Silane side pour une valeur de la RF Bias de 3500W.

La figure 4.30 montre l'évolution de l'étendue de l'épaisseur en fonction des paramètres critiques. On observe que l'étendue décroît avec l'augmentation de débit de silane Side. Mais, quand le débit de silane Top croît, l'étendue croît également. A partir de ce résultat, nous pouvons conclure que la dérive de SiH₄ Side peut être compensée par l'ajustement de la valeur de SiH₄ Top pour garantir la valeur cible de l'étendue de l'épaisseur d'oxyde.

4.2.5.3.b Simulation du modèle

En vue de valider le modèle, nous avons procédé à des simulations de la même façon que pour le modèle de l'épaisseur. Les résultats des simulations montrent que la prédiction en étendue d'épaisseur reste juste tant que la mesure se fait sur la même cartographie de référence (utilisée pour le plan d'expériences). Lorsqu'elle se fait sur une autre cartographie, la prédiction en étendue d'épaisseur est faussée car cette dernière n'est pas établie entre les mêmes zones de la cartographie. Dans ces conditions, l'optimisation ne peut pas fonctionner convenablement. Le modèle de l'étendue d'épaisseur n'est donc pas généralisable directement si ce problème de cartographie n'est pas résolu. Une approche pour prédire la valeur des sites de mesures non mesurables a été adoptée à l'aide du programme Wafer-Fit. Notre objectif est d'arriver à modéliser ce paramètre de sortie avec peu de points de mesures.

4.2.5.3.c Interpolation des points

L'uniformité est un paramètre difficile à modéliser à cause, généralement, du peu de sites de mesures après l'étape du procédé [GUO 1993]. La figure 4.31 montre un exemple de profil de dépôt symétrique en 3D. Cette plaque a été mesurée sur une cartographie de mesure contenant 49 points.

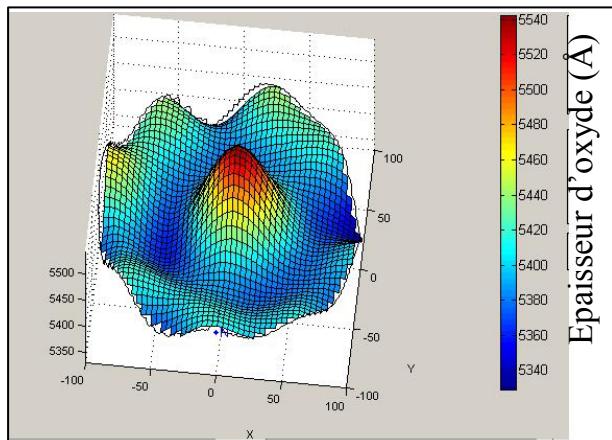


Figure 4.31- Profil de dépôt en 3-D mesuré en 49 points

Comme c'est illustré sur la figure 4.31, on distingue plusieurs zones avec différentes épaisseurs. Dans ce cas, l'épaisseur maximale se trouve dans la zone centrale de la plaque. Ce profil de dépôt est établi en fonction de la distribution des ions du plasma. Avoir un profil de dépôt avec une telle précision nécessite l'emploi d'une cartographie contenant plusieurs points de mesures. Ceci permettra de connaître également les effets de bords de plaque [ALEGRET 2005]. Dans le milieu de la production, ce type de cartographies n'est pas utilisé souvent sauf pour des tâches de qualifications ou pour des expériences spécifiques. La

cartographie de 9 points est souvent utilisée à ST-Rousset (200mm) et en particulier pour le procédé HDP-CVD.

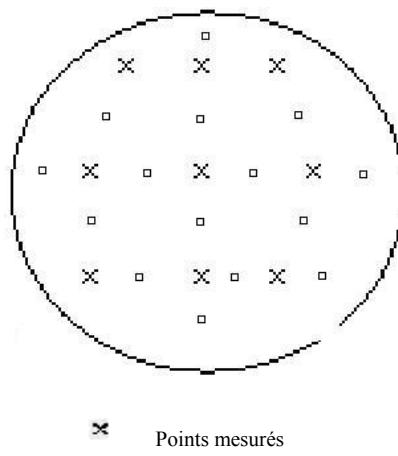


Figure 4.32- Représentation des points mesurés et points non mesurés

La figure 4.32 montre la cartographie des 9 points de mesures. Le choix de la localisation des 9 points est de tenir compte du centre de la plaque et des régions proches des bords de plaques. Pour une même technologie, la localisation des points sur la cartographie de mesures diffère d'un produit à un autre, suivant la localisation des structures de test où les mesures sont effectuées. Ces points de mesures sont légèrement translatés de quelques mm. Vu le manque d'information donné par une cartographie contenant peu de points, le défi est d'arriver à estimer le profil réel de la plaque sans utiliser beaucoup de points de mesures.

On a focalisé notre étude sur deux produits d'une même technologie ayant la même recette au niveau du procédé mais mesurés sur deux cartographies de mesures différentes. La cartographie de l'un des deux produits est la même que celle utilisée pour le DOE effectué pour établir le modèle.

4.2.5.3.d Simulations et analyses

Les résultats des simulations montrent que seulement pour les plaques mesurées (produits associés) avec la même cartographie que celle du DOE, la prédiction est bonne. Alors que pour les autres plaques ayant été mesurées sur d'autres cartographies différentes de celles utilisées pour le DOE, la prédiction est erronée [BELHARET 2006 b].

La prédiction de l'étendue d'épaisseur dépend également de la position des points de mesures sur la plaque. Les analyses de l'étendue d'épaisseur ont été effectuées avec l'aide de Wafer-Fit. Comme expliqué dans le chapitre 1, le principe de Wafer-Fit repose sur l'interpolation des points pour s'assurer que la surface est continue. Les différentes mesures faites avec une cartographie contenant 49 points de mesures servent pour alimenter la base de WaferFit et générer un algorithme qui sera capable par la suite d'apprendre la surface. Il est ainsi possible d'estimer la mesure à n'importe quel point sur la surface de la plaque.

4.2.5.3.e Validation de l'approche

Pour valider cette approche, des expériences ont été effectuées sur deux lots processés consécutivement. Ces deux lots sont ensuite mesurés par la cartographie standard et la

cartographie contenant 49 points de mesures. L'algorithme de reconstruction de surface du Wafer-Fit a été généré à partir de la cartographie à 49 points de mesures. Cet algorithme est par la suite utilisé pour générer les surfaces à partir des cartographies à 9 points et en déduire les indices statistiques plus précis.

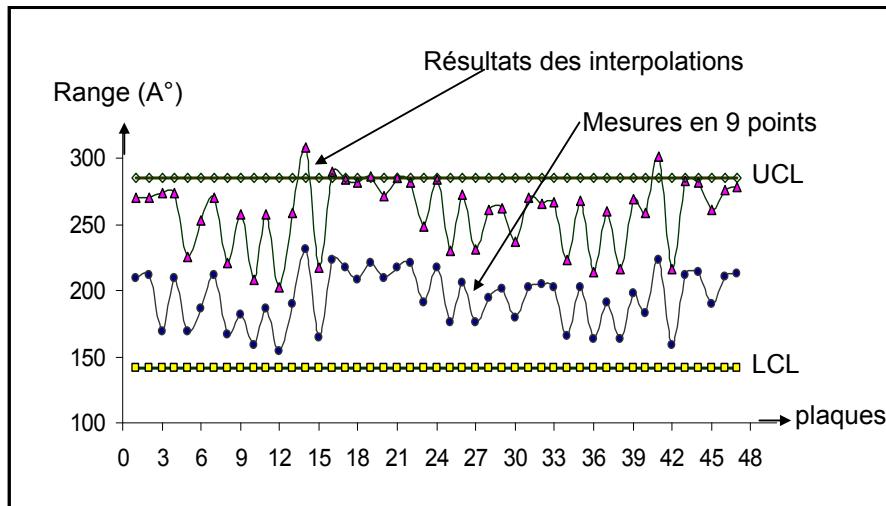


Figure 4.33- Evolution de l'étendue d'épaisseur avec le temps (mesures issues des 9points de mesures et les résultats de l'interpolation des 9 points de mesures)

Comme c'est illustré sur la figure 4.33, 6 plaques (14,16,17,19,21 et 41) sur un 2 lots de 50 plaques sont hors contrôle (en comparaison avec les limites de contrôle hautes UCL). Cet état n'est pas détecté par les mesures 9 points. Les mesures avec interpolation sont donc plus robustes que les mesures avec 9 points. Ce résultat est vérifié avec la cartographie contenant 49 points de mesures.

4.2.5.3.f L'effet des injecteurs de gaz

En plus de l'effet de la cartographie que nous avons cité précédemment, un autre effet est à prendre en considération pour la modélisation de l'étendue de l'épaisseur, il s'agit de l'encrassement des injecteurs de gaz. Les chambres de dépôt ont une construction à symétrie centrale mais peuvent souffrir d'une injection inhomogène du gaz sur la plaque suite à des défaillances d'injecteurs encrassés. La figure 4.34a montre la disposition des injecteurs de gaz. Comme le débit total est maintenu constant, la perte de débit dans un injecteur se répercute par une augmentation du débit dans les autres injecteurs. La perte ponctuelle qui s'observe sur la plaque se compense nécessairement par des épaisseurs plus hautes sur les autres points, ce qui se traduit par la suite par une non uniformité de l'épaisseur d'oxyde sur la plaque. La figure 4.34b illustre cette non uniformité d'oxyde. Cette non uniformité est le résultat de l'encrassement des injecteurs.

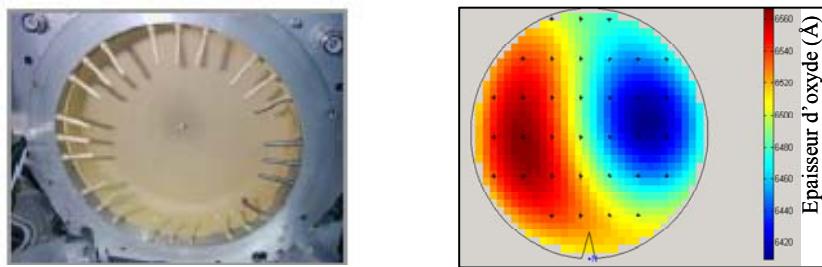


Figure 4.34- a) La répartition symétrique des injecteurs des gaz. b) profil après dépôt mesuré avec une cartographie de 49 points (présentant le problème d'injecteurs encrassés)

Nous observons une zone avec des épaisseurs d'oxyde de l'ordre de 6420Å et une autre zone avec des épaisseurs de l'ordre de 6550Å, l'étendue de l'épaisseur d'oxyde est de l'ordre de 130Å. Cette non uniformité d'oxyde va influencer les étapes de fabrication suivantes et contribue à la non uniformité de la hauteur de marche comme nous allons le constater dans les parties suivantes de ce chapitre.

4.2.6 Conclusion

Dans les paragraphes précédents, nous avons déterminé les types de variabilités dans le procédé HDP-CVD (lot à lot, plaque à plaque et intra-plaque). En comparant les différentes dispersions constatées (lot à lot, plaque à plaque et intra- plaque) avec les limites des cartes de contrôle (épaisseur d'oxyde, étendue), les dispersions sont très importantes et peuvent provoquer des dérives. Seules les variabilités lot à lot et intra- plaques ont été traitées dans ce manuscrit. Pour la variabilité intra- plaque, nous avons démontré que cette variabilité est liée à l'effet de la pression sur les deux premières plaques processées juste après le nettoyage. Nous avons démontré également qu'on peut réduire cette variabilité en rajoutant un temps de procédé de dépôt supplémentaire sur les deux premières plaques.

Avant de passer à la partie expérimentale, nous avons identifié les paramètres du procédé critiques. On a détaillé le choix des paramètres critiques (RF Bias, SiH₄ Side et SiH₄ Top) qui nous a permis d'obtenir le modèle du procédé de l'épaisseur et celui de l'étendue d'épaisseur. Ces paramètres seront impliqués dans une boucle de régulation R2R de type feedback. Nous avons modélisé l'épaisseur et l'étendue d'un dépôt HDP-CVD sur une plaque (200mm). La modélisation de l'épaisseur moyenne déposée donne entière satisfaction dans les tests de simulations effectuées. Ce modèle est prêt pour la pré- production. La modélisation de l'étendue d'épaisseur donne une bonne prédiction mais reste fortement dépendante de la cartographie de mesure utilisée, donc une dépendance avec le produit utilisé. Nous avons démontré l'intérêt de l'emploi d'une méthode d'interpolation basée sur l'apprentissage statistique propre à chaque produit. L'approche proposée répond aux exigences d'une bonne régulation à savoir alimenter les boucles de régulation avec des mesures précises.

Dans les prochains paragraphes, nous allons exposer nos travaux sur la modélisation de la non uniformité d'oxyde post-CMP sur les bords des plaques et nous allons proposer une solution pour réduire cette non uniformité d'oxyde.

Partie 2

Le procédé de polissage mécanico-chimique (CMP)

Dans cette partie nous allons exposer notre étude pour l'amélioration du rendement de fabrication au niveau des bords des plaques. Le mode STI visé par cette étude est le STI standard. Ce module est marqué par une non uniformité de la hauteur de marche. Nous allons proposer une méthode de réduction de cette non uniformité en agissant sur le procédé CMP. Il s'agit de réduire la non uniformité de l'oxyde après CMP. Il faut noter que l'autre mode STI à savoir le Direct STI n'est pas concerné par cette étude car la hauteur de marche présente une très bonne uniformité. Nous allons décrire rapidement les différents travaux scientifiques menés pour modéliser le procédé CMP et en particulier les travaux concernant la non uniformité de la plaque. C'est sur ce dernier point que nous allons nous focaliser pour y apporter notre contribution dans ce type de modélisation. La non uniformité sur la plaque provient généralement d'une différence de vitesse d'enlèvement entre le bord et le centre. Cette différence de vitesse est liée principalement à la différence de pression exercée dans les deux zones (bord-centre). Le travail consiste donc à comprendre la distribution de la pression sur la plaque et proposer un système de régulation de la pression sur les bords des plaques. L'objectif final de cette étude est de déduire un modèle pour la boucle de régulation R2R de type « feed-back ».

4.3.1 Introduction

Le polissage mécano-chimique (CMP) est un procédé de lissage et de planarisation des surfaces combinant des actions chimiques et mécaniques, un mélange de gravure chimique et de polissage mécanique à abrasif⁵⁸ libre. Le rodage mécanique seul, cause trop de dégâts sur les surfaces et la gravure humide seule ne permet pas d'obtenir une bonne planarisation. Les réactions chimiques étant généralement isotropes, elles attaquent les matériaux indifféremment dans toutes les directions. Le CMP combine les deux effets en même temps.

Les variations d'épaisseurs post-CMP impactent le rendement final de fabrication des puces au niveau électrique [TROUNG 1998]. La non uniformité [HUANG 1995] au sein de la plaque constitue depuis longtemps une limitation du niveau d'intégration des circuits intégrés en particulier, en bord des plaques. Au niveau du procédé CMP, ceci s'explique par la différence de la vitesse d'enlèvement à travers la surface de la plaque comme cela est illustré sur la figure 4.35. La figure 4.35 illustre un profil d'oxyde avant CMP. Nous avons pu constater dans la partie HDP-CVD que la non uniformité de dépôt donne lieu à l'apparition de différentes zones sur la plaque. Pour expliquer la notion de non uniformité d'oxyde, nous nous intéressons dans cet exemple à deux zones voisines sur la plaque. Si la vitesse d'enlèvement est la même sur tous les points des deux zones, le profil d'oxyde après CMP reste presque le même. Ceci est loin de la réalité physique comme nous allons le voir dans les prochains paragraphes, car, avec la différence des vitesses d'enlèvement dans les deux zones, les profils sont différents et cela va s'accompagner par la consommation d'une épaisseur d'oxyde additionnelle symbolisée sur la figure par Δ épaisseur d'oxyde.

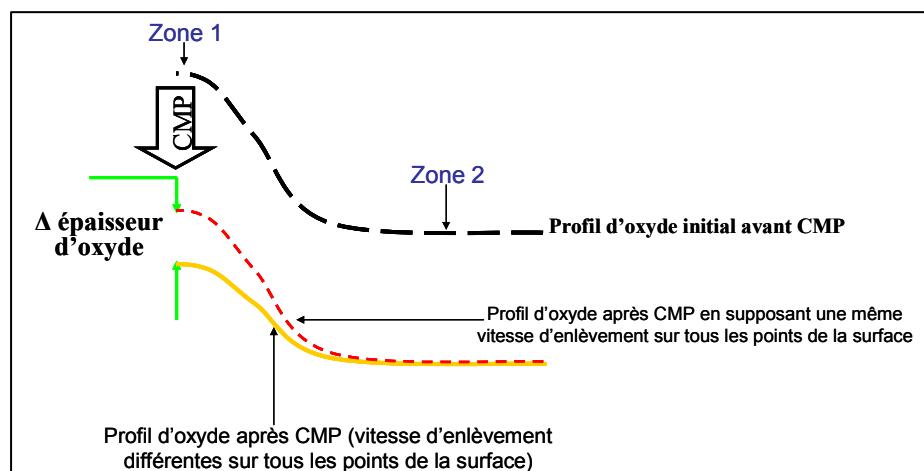


Figure 4.35- Profils d'oxyde après CMP

Les étapes concernées par la non uniformité sont généralement : le polissage d'oxyde au niveau du STI et le polissage du cuivre au niveau métal pour les technologies d'interconnexion de type « damacene ». Ces variations d'épaisseur centre-bord au niveau de la plaque vont se traduire par des variations des performances électriques. Ce phénomène impacte directement le rendement final. L'une des actions nécessaires pour améliorer le rendement est de comprendre les mécanismes de formation de cette non uniformité et de

⁵⁸ Une suspension aqueuse de particules pouvant être de différentes natures. On utilise par exemple des particules de silice, de cérium ou d'alumine, dont la taille varie de 10 nanomètres à 1 micromètre. Les abrasifs utilisés pour le polissage des diélectriques sont généralement des suspensions colloïdales de silice

proposer des solutions parmi lesquelles la réalisation des boucles de régulation de type R2R [BONINIG 1996, EL CHEMALI 2000]. Avant de détailler notre démarche pour la réduction de la non uniformité d'oxyde sur la plaque, nous allons rappeler d'abord les notions de base du procédé CMP.

4.3.2 Description du procédé CMP

4.3.2.1 Description de l'équipement CMP

La plaquette que l'on veut polir est fixée à l'intérieur de la tête de polissage. La force de maintien provient de gouttes d'eau présentes à l'interface entre la plaquette et la membrane interne de la tête. La plaquette est ensuite amenée sur le plateau de polissage. L'abrasif est une solution aqueuse de particules pouvant être de différentes natures. On utilise par exemple des particules de silice, de cérium ou d'alumine dont la taille varie de 10 nm à 1 µm. L'abrasif est déposé sur le plateau par un système de distribution simple composé la plupart du temps d'un injecteur dont on maîtrise le débit. Les différentes composantes de la machine sont illustrées sur la figure 4.36.

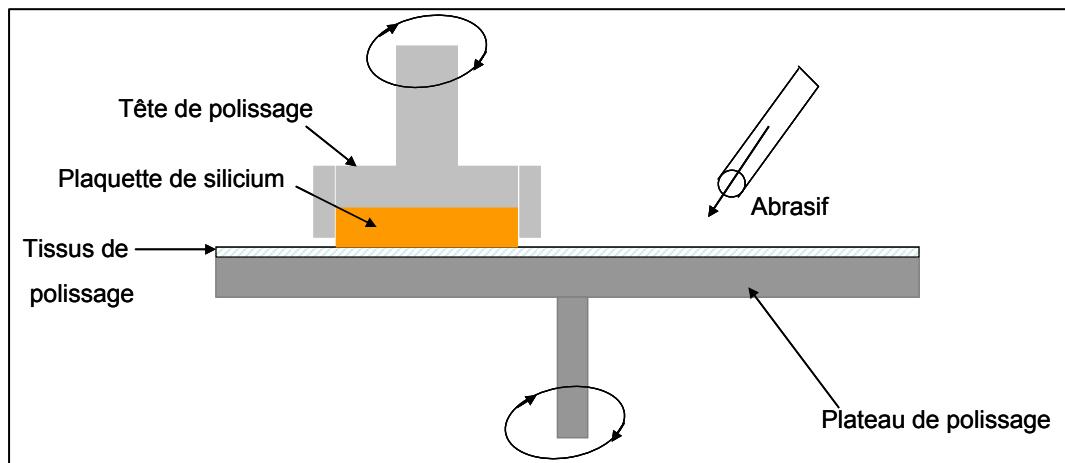


Figure 4.36- Représentation schématique de la machine de polissage

Les différentes expériences menées lors de cette étude ont été effectuées sur une machine de polissage composée de 3 plateaux de polissage et de 4 têtes de polissage. Le procédé de polissage est reproductible si l'état de surface de son tissu est constant. Ceci nécessite de l'aviver⁵⁹ entre chaque plaquette polie.

4.3.2.2 Mécanismes du procédé CMP

L'enlèvement de la matière s'effectue réellement de manière mécano-chimique. En effet, le labourage mécanique créé par les particules abrasives est essentiel à l'hydratation de la surface. La modification des propriétés de la surface résultant de cette hydratation permet

⁵⁹ Selon la machine, l'avivage est réalisé pendant ou après le polissage des plaquettes. Aviver le tissu consiste à enlever mécaniquement une certaine épaisseur du tissu, usé par le polissage de la plaquette. En pratique, on appuie sur le tissu un anneau sur lequel ont été fixés des éclats de diamant. On déplace l'anneau à la surface du tissu avec un double mouvement de rotation et de translation afin d'aviver le tissu le plus uniformément possible.

alors l'enlèvement ductile de matière par les particules abrasives. L'abrasion a lieu lorsqu'il y a contact entre le tissu, les particules abrasives et le diélectrique de la plaquette

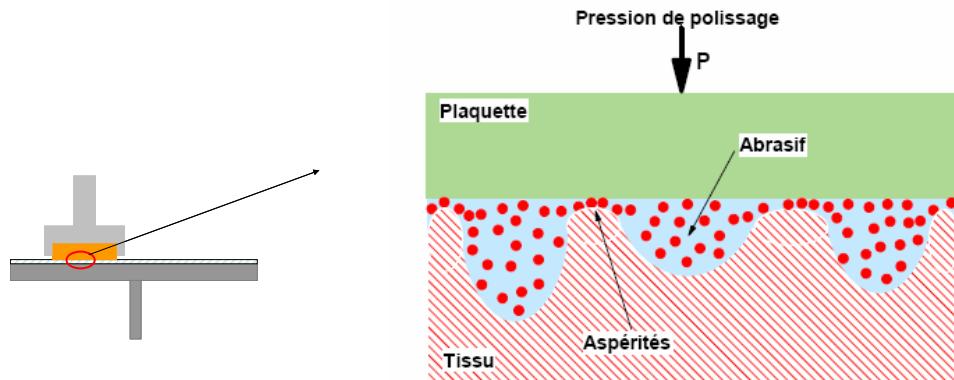


Figure 4.37- Représentation de l'interface plaquette- tissu de polissage.

4.3.2.3 La fin du polissage

En fonction de l'épaisseur moyenne d'oxyde à polir, le temps de polissage est calculé avec l'aide de la vitesse d'enlèvement d'oxyde. Pour s'assurer que l'oxyde est enlevé complètement lors du procédé de polissage, on rajoute un temps additionnel appelé « temps de sur- polissage ». Suivant la topographie du profil d'oxyde, certaines zones d'oxyde sont enlevées plus rapidement que d'autres. Si la vitesse d'enlèvement se dégrade avec le temps d'utilisation de l'équipement, ceci risque de ne pas enlever complètement l'oxyde et va faire apparaître dans certaines zones de l'oxyde résiduel comme cela est illustré sur la figure 4.38b. Dans la figure 4.38a est illustré un exemple d'une structure STI. Il s'agit de la région active proche de la tranchée STI (film silicium/oxyde piédestal/nitrate/oxyde).

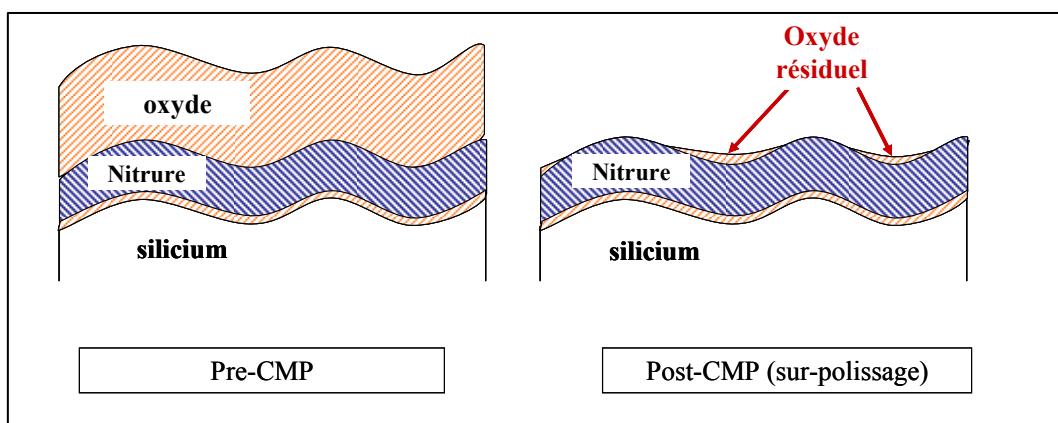


Figure 4.38- a) Structure d'un film avant CMP b) Après CMP : oxyde résiduel reste dans certains endroits.

Un système de fin de polissage (End point) est implémenté sur la machine de polissage. La fin du polissage est détectée après avoir consommé tout l'oxyde. Ensuite, un temps de sur-polissage est rajouté pour s'assurer de l'enlèvement total de l'oxyde.

Plusieurs travaux de recherches ont été effectués pour modéliser le procédé CMP. Avant de positionner notre contribution dans ce type de modélisations, nous allons récapituler les principaux axes de recherches sur la modélisation du procédé CMP.

4.3.3 La modélisation du polissage mécano-chimique

La modélisation du procédé CMP est très complexe car elle tient compte de plusieurs paramètres (température, vitesse...) et de plusieurs paramètres relatifs à la plaque comme la taille de la plaque, le profil de la plaque, les propriétés de la couche à polir... Les paramètres de sortie sont aussi multiples (la vitesse d'enlèvement, l'uniformité, les défauts). Une représentation d'une modélisation du procédé CMP est illustrée sur la figure 4.39.

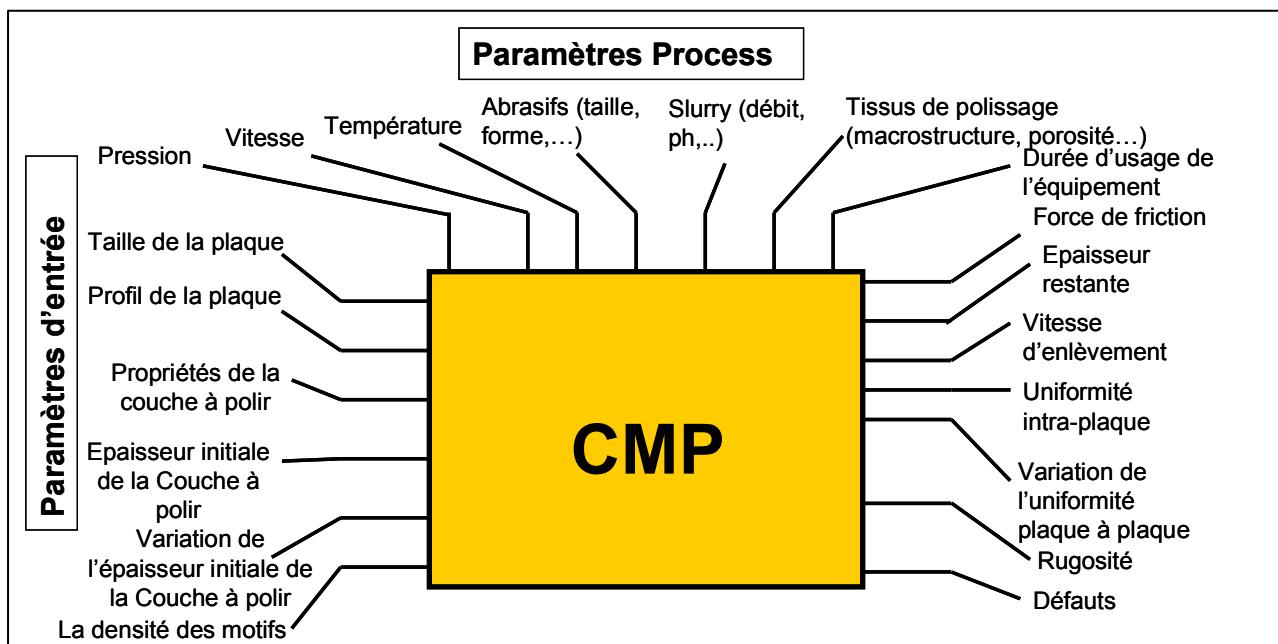


Figure 4.39- Modélisation du procédé CMP [LAI 2001]

Dans les paragraphes suivants, nous ne nous intéresserons qu'aux paramètres de sortie relatifs à l'uniformité intra-plaque. Le nombre de publications sur ce sujet augmente fortement depuis le début des années 1990⁶⁰. On peut classifier les différentes études en 4 axes principaux de recherches.

4.3.3.1 Les axes principaux de recherches consacrés à la modélisation du CMP

4.3.3.1.a Modélisation des phénomènes d'usure

Ces travaux consistent à analyser l'état de surface du tissu⁶¹ de polissage en contact avec la plaquette [YU 1994, PERRIN 2000]. Ces études sont une suite des premiers travaux

⁶⁰ Après avoir publié le premier papier sur le procédé CMP. La plupart des compagnies et équipementier se sont penchés sur l'étude des avantages qu'offre le CMP.

⁶¹ Le tissu est un élément clé du dispositif de polissage. Le choix des matériaux a été fait de manière très empirique par les polisseurs de verres optiques. Dès 1930, Preston montrait que les différentes étapes de polissage du verre nécessitaient des tissus aux propriétés mécaniques différentes selon le but recherché [Preston 1930]. Lorsque le polissage mécano-chimique est apparu en micro-électronique dans les années 80, les tissus disponibles étaient constitués de polyuréthane. Leur structure varie selon la rigidité souhaitée du tissu. Nous distinguons 2 sortes de tissus :

réalisés sur des verres optiques. Les travaux menés par L. COOK [COOK 1990] avaient pour objectif de rapprocher le savoir issu du travail expérimental mené sur les verres optiques et celui plus théorique fourni par la science des surfaces [Cook 90]. Par la suite, la science des contacts a été appliquée au polissage mécano-chimique [Liu 1996]. Une analyse détaillée a été proposée par T.-K. Yu en 1994 dans laquelle l'auteur a pris en compte l'état de surface du tissu de polissage au contact de la plaquette [Yu 1994]. Malgré un certain nombre de publications, peu de chercheurs ont abordé la modélisation des phénomènes d'usure pour le polissage mécano-chimique de l'oxyde de silicium en micro-électronique [PERRIN 2000].

4.3.3.1.b Modélisation de l'aplanissement à l'échelle des motifs

Ces études traitent de la relation entre la taille des motifs⁶² et leur aplanissement. La méthode générale consiste à appliquer en chaque point un coefficient d'accélération ou de décélération de la vitesse d'attaque en fonction de sa géométrie et de son environnement [BURKE 1991, WARNOCK 1991, NANZ 1995]. La validité de ces modèles sur un certain nombre réduit de motifs a été démontrée, mais leur manque de fondement physique ne permet pas de les généraliser sur les circuits [OTHA 1999]. D'autres études de modélisations ont été effectuées avec des représentations simplifiées du tissu de polissage [SIVARAM 1992, NANZ 1995]. Malgré de bonnes performances dans certains cas particuliers, ces modèles ne peuvent pas également être généralisés à cause de leur manque de prédition. On peut illustrer d'une manière générale le procédé CMP

4.3.3.1.c Modélisation de l'aplanissement à l'échelle de la puce

L'approche est presque la même que la précédente, sauf qu'ici au lieu de considérer un motif isolé ou un groupe de motifs, on considère les densités de motifs présents sur la puce. Des relations sont déduites pour décrire l'aplanissement des diverses zones de densité [TSENG 1997, PERRIN 2000].

4.3.3.1.d Modélisation de l'uniformité sur la plaque

La plupart des travaux présentés sur ce type de modélisations sont issus des résultats des plans d'expériences (DOE) où un grand nombre de paramètres liés au procédé est exploré pour voir leur impact sur la variation de l'épaisseur d'oxyde enlevé comme :

- les propriétés de l'abrasif utilisé (taille des abrasifs, qualité des abrasifs....) [LUO 2001, LUO 2003a, LUO 2003b, COPETTA 1998]
- les propriétés de tissu de polissage [BORBA 1995]
- la vitesse de rotation [BORBA 1995, GAN 2000]
- la répartition des contraintes mécaniques sur la plaque générées par les têtes de polissage [WANG 1997].....
- d'autres études ont été menées pour minimiser la non-uniformité d'oxyde sur la plaque en optimisant un ensemble de paramètres de procédés (la vitesse de rotation de

* Les tissus souples sont construits comme des feutres dont les fibres sont imprégnées de polyuréthane. Ceux-ci sont utilisés dans le polissage du verre pour réduire la rugosité finale de la surface, alors que la pièce possède déjà la forme voulue.

* Un tissu plus rigide est obtenu par moulage d'une matrice de polyuréthane. Des ajustements de propriétés mécaniques sont possibles en utilisant des additifs ou en faisant varier la densité de la matrice. Le matériau obtenu est poreux.

⁶² Ce sont les éléments de la plaque à aplanir.

plateau, la vitesse de rotation de la tête, la concentration des abrasifs, la pression de la tête principale) après avoir identifié la criticité de chaque paramètre sur la non-uniformité d'oxyde sur la plaque [PARK 2003].

L'étude présente se situe dans le prolongement de ces travaux de modélisation. Comme notre travail s'inscrit pour une application de boucle de régulation R2R et non pour un changement de procédé, on ne va donc pas s'intéresser aux aspects chimiques, mais seulement à l'aspect mécanique en agissant sur les paramètres ajustables. Tout d'abord on va rappeler les notions de base régissant la non uniformité de l'oxyde à travers la plaque.

4.3.3.2 Le degré d'aplanissement

Le degré d'aplanissement est défini comme étant le pourcentage de la hauteur de marche initiale (H_i) enlevée par le polissage. La valeur du degré d'aplanissement varie entre 0% pour un aplanissement nul et 100% pour un aplanissement total. La hauteur de marche ΔH désigne la topographie finale après polissage.

$$\text{Le degré d'aplanissement s'écrit : } DA = 1 - \frac{\Delta H}{H_i} \quad (4.12)$$

Où : ΔH est la hauteur de marche d'oxyde initiale et H_i est la hauteur de marche d'oxyde finale après polissage.

Le degré d'aplanissement peut varier suivant les conditions des procédés imposées lors de la réalisation de la brique STI. Nous allons donner différentes modélisations du degré d'aplanissement.

4.3.3.2.a Modélisation simple du degré d'aplanissement

Une représentation simple des étapes du procédé de la brique STI est donnée dans la figure 4.40. Après le dépôt d'oxyde piédestal et du nitre, la tranchée d'une profondeur H_i est définie après gravure. Ensuite, si on suppose que le dépôt HDP-CVD se fait d'une manière homogène, on obtient donc un premier profil de dépôt (avant polissage).

Soit :

- a est l'épaisseur d'oxyde (dans la tranchée) polie
- C est l'épaisseur STI avant polissage
- b est l'épaisseur d'oxyde polie (oxyde déposé sur le nitre)
- A est l'épaisseur STI après polissage
- B est l'épaisseur d'oxyde sur le nitre après polissage

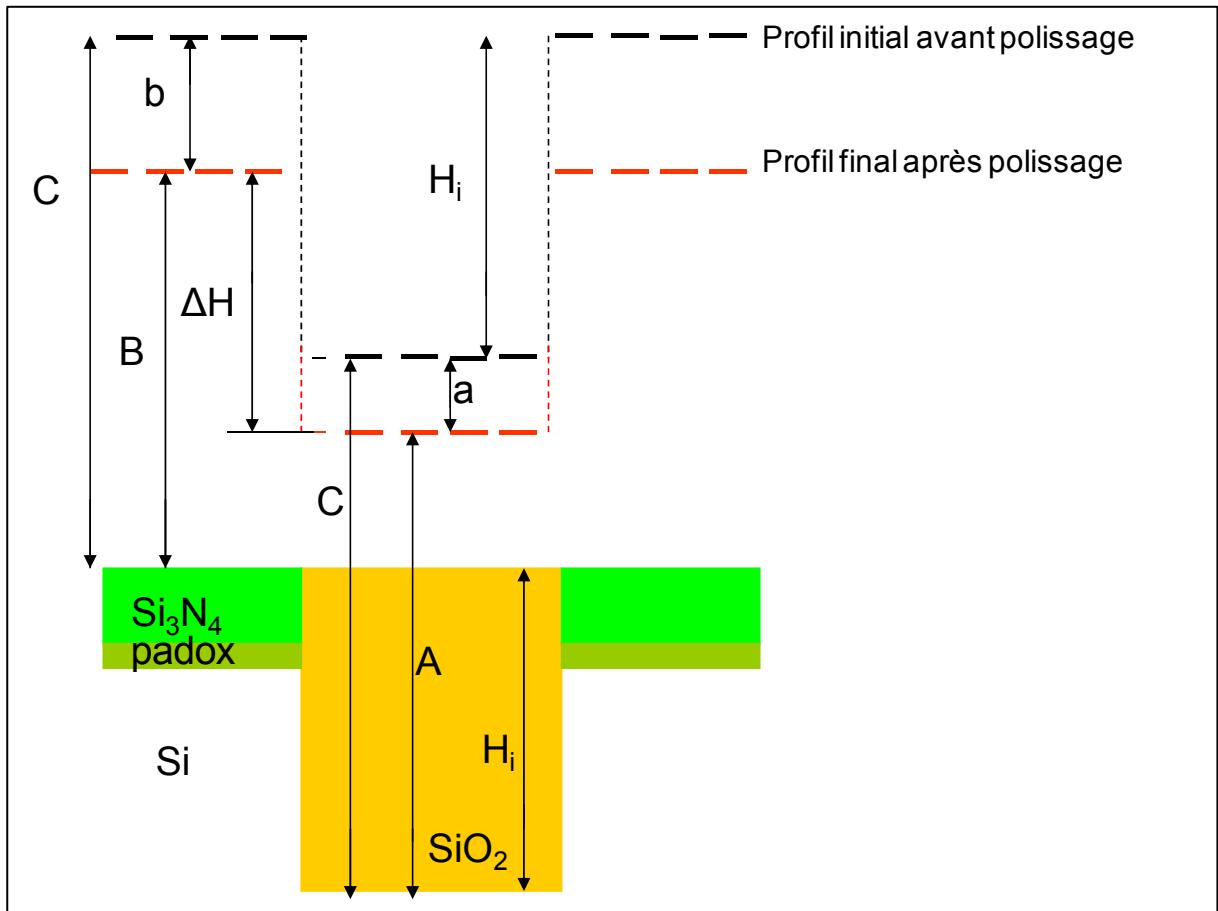


Figure 4.40- Représentation schématique des marches initiales et finales

Dans les structures à grandes dimensions (plusieurs microns) qui servent pour la mesure d'oxyde par ellipsométrie et pour la mesure de la hauteur de marche par la technique AFM, la valeur de l'épaisseur d'oxyde déposée dans la tranchée est la même que celle déposée au dessus de nitre. Comme cela est illustré sur la figure 4.40. La valeur de H_i est de l'ordre de 4700Å. La valeur de l'épaisseur d'oxyde enlevée dans la tranchée (a) est différente de celle enlevée dans la région voisine (b). Nous pouvons expliquer cette différence par le fait que le procédé CMP dépend fortement de la taille des motifs à polir. Les épaisseurs enlevées sont données par les relations suivantes :

$$a = C - A \quad (4.13)$$

Et

$$b = C - B \quad (4.14)$$

$$\text{La hauteur de marche finale } \Delta H \text{ est déduite : } \Delta H = H_i + a - b \quad (4.15)$$

$$\text{En remplaçant les valeurs de } a \text{ et } b, \text{ on a donc : } \Delta H = H_i + B - A \quad (4.16)$$

$$\text{Le degré d'aplanissement est donc } DA = 1 - \frac{\Delta H}{H_i} = \frac{A - B}{H_i} = \frac{b - a}{H_i} \quad (4.17)$$

Donc on voit clairement que le degré d'aplanissement est inversement proportionnel à la profondeur de la tranchée. Pour vérifier la non uniformité de la profondeur de la tranchée

sur la plaque, différentes plaques ont été mesurées après l'opération de la gravure de la tranchée. Un exemple de profil des profondeurs de la tranchée sur la plaque est illustré sur la figure 4.41. Nous observons différentes régions sur la plaque avec des plages d'épaisseurs différentes allant de 4750 à 5050 Å. Cette différence peut s'expliquer par le fait que des régions dans la plaque sont plus rapidement gravées que les autres. L'explication de ce phénomène nécessite une analyse très fine des paramètres équipements de la chambre de gravure (température, puissances RF, débits de gaz.....). La non-uniformité de la profondeur de tranchée induit directement par la suite, une non uniformité de profil de l'oxyde après CMP et donc une non uniformité de profil de la hauteur de marche du STI.

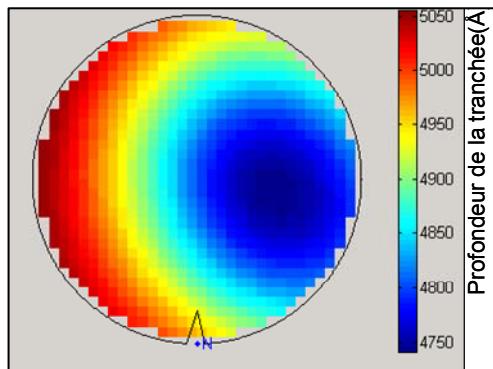


Figure 4.41- Profil de la profondeur de tranchée après gravure

Le déploiement de la boucle de régulation R2R gravure permet de réduire cette variabilité intra- plaque et donc d'améliorer l'uniformité du profil de la hauteur de marche.

4.3.3.2.b Modélisation du degré d'aplanissement en tenant compte de l'effet de la gravure inverse

En réalité, dans un procédé STI standard comme décrit au chapitre 2, après le dépôt HDP-CVD et le dépôt TEOS, l'uniformité est améliorée par gravure de l'oxyde avec un masque inversé de résine par rapport à la gravure active. Cette étape est très importante pour la qualité de l'étape de CMP et implique une importante épaisseur d'oxyde au dessus du silicium. Une fois l'oxyde gravé nous avons une surface beaucoup plus uniforme et par conséquent la représentation de la figure 4.40 n'est pas exacte. La marche initiale n'est plus H_i mais $H_i - H_0$, où H_0 est l'épaisseur d'oxyde gravée. Cela conduit à une nouvelle expression du degré d'aplanissement qui montre que le procédé de gravure inverse favorise l'aplanissement.

$$\boxed{DA = 1 - \frac{\Delta H}{H_i - H_0} = \frac{b - a}{H_i - H_0}} \quad (4.18)$$

Il a par ailleurs été démontré que l'étape de gravure HF n'introduisait pas une variabilité significative sur plaque [BELHARET 2006c] et que l'hypothèse $H_0 = \text{Constante}$ est tout à fait justifiable.

4.3.3.2.c Modélisation du degré d'aplanissement en tenant compte de la fin de polissage

En fait la modélisation précédente est simple car, dans le procédé CMP, on s'arrête après la révélation du nitre parce que les abrasifs employés pendant le procédé CMP présentent une faible sélectivité avec le nitre donc, une partie de la couche de nitre est attaquée par le CMP. De plus, le CMP détecte la fin du polissage après avoir consommé près de 500 Å de nitre au lieu de s'arrêter dessus. L'épaisseur de nitre consommée est symbolisée sur la figure 4.42 par « d ». Donc le CMP continue à attaquer l'oxyde et en plus les vitesses d'attaque de l'oxyde et du nitre sont différentes. Il faut noter qu'on peut avoir de l'oxyde résiduel au dessus de nitre après le procédé CMP. Ceci ne pose pas de problème car nous pouvons l'enlever au niveau de l'étape de fabrication qui suit (gravure humide).

La figure 4.42 illustre la forme de la hauteur de marche au moment d'atteindre le nitre, cette hauteur est symbolisée par (ΔH_1) . Pendant le surpolissage, la hauteur de marche change de (ΔH_1) à (ΔH) qui est de l'ordre de 50Å. Comme la vitesse d'enlèvement de nitre est plus rapide que celle d'oxyde (dans la tranchée), la valeur de (ΔH_1) est donc inférieure à 50Å. Nous concluons donc, que pendant le polissage, la hauteur de marche passe de 4700Å à moins de 50Å. Ce qui permet d'avoir un aplatissement quasi total et un relief quasi nul.

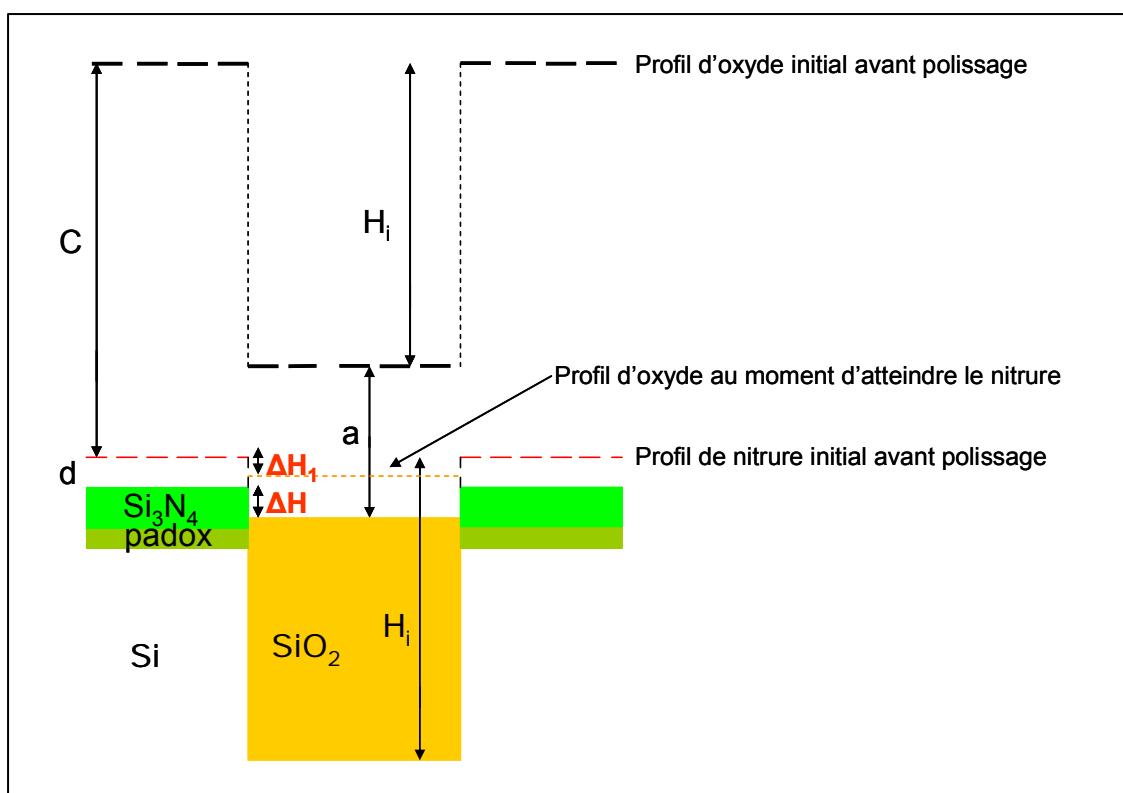


Figure 4.42- Hauteur de marche finale après polissage

L'épaisseur d'oxyde post CMP est donc donnée par la relation suivante:

$$\text{Epaisseur d'oxyde post-CMP} = H_i - \Delta H - d \quad (4.19)$$

Comme nous avons pu le démontrer dans la partie HDP-CVD de ce chapitre, l'épaisseur déposée dans la tranchée varie d'une région à une autre (exemple bord-centre), l'épaisseur d'oxyde post-CMP va varier de la même façon.

La non uniformité de l'épaisseur d'oxyde enlevée à travers la plaque peut être analysée par l'étude de La non uniformité de l'épaisseur d'oxyde post-CMP moyennant l'hypothèse que l'épaisseur déposée présente une bonne uniformité.

4.3.4 Définition de la non uniformité à travers la plaque (WIWNU : Within Wafer Non Uniformity)

Nous allons rappeler brièvement la définition de la non uniformité à travers la plaque pour bien comprendre les paramètres qui l'affectent. La non uniformité à travers la plaque (WIWNU) est généralement définie par les vitesses d'enlèvement maximale (V_{\max}), minimale (V_{\min}) et moyenne (V_{moy}). Le WIWNU peut être estimé à l'aide de la relation suivante :

$$\boxed{\text{WIWNU}(\%) = \frac{V_{\text{enlev. max}} - V_{\text{enlev. min}}}{V_{\text{enlev. moy}}} \times 100} \quad (4.20)$$

La non uniformité de la vitesse d'enlèvement à travers la plaque peut être attribuée à différents paramètres tels que la température et la distribution des abrasifs. Cependant, il est admis que la pression et la vitesse sont les paramètres les plus contributeurs à cette non uniformité [LUO 2003c]. En remplaçant dans la relation (4.21) la formule de la vitesse (relation de Preston), on obtient :

$$\boxed{\begin{aligned} \text{WIWNU}(\%) &= \frac{K_e(PV_{\text{enlev}})_{\max} + V_{\text{enlev.0}} - K_e(PV_{\text{enlev}})_{\min} - V_{\text{enlev.0}}}{K_e(PV)_{\text{moy}} + V_{\text{enlev.0}}} \times 100 \\ &= \frac{K_e[(PV_{\text{enlev}})_{\max} - (PV_{\text{enlev}})_{\min}]}{K_e(PV)_{\text{moy}} + V_{\text{enlev.0}}} \times 100 = \frac{K_e[(PV_{\text{enlev}})_{\max} - (PV_{\text{enlev}})_{\min}]}{V_{\text{enlev.moy}}} \times 100 \end{aligned}} \quad (4.21)$$

Où : K_e est la constante de Preston

P est la pression moyenne exercée au centre de la plaque.

$V_{\text{enlev.0}}$ est la vitesse d'enlèvement initiale.

Si on veut réduire le WIWNU, on a deux possibilités :

- soit de réduire le terme $[(PV_{\text{enlev}})_{\max} - (PV_{\text{enlev}})_{\min}]$. Cette solution n'est pas facile à réaliser car elle est relative à la conception de la machine (design). En d'autres termes, à la conception du plateau et des têtes de polissages.
- soit d'augmenter le terme $V_{\text{enlev.moy}}$, ce dernier dépend de la vitesse et de la pression.

Nous supposons dans la suite de cette étude que la vitesse est uniforme. Le paramètre restant donc à explorer est la pression.

Indicateurs statistiques pour caractériser la non-uniformité

Les indicateurs statistiques utilisés habituellement pour les études de l'uniformité des grandeurs physiques comme la moyenne, l'écart type (σ) et l'étendue sont insuffisants pour caractériser la non-uniformité de l'épaisseur d'oxyde [SMITH 2002]. D'autres indicateurs sont proposés pour bien accomplir l'étude de cette non uniformité parmi lesquels:

*) Ecart type normalisé (%std dev) de l'épaisseur d'oxyde post CMP:

$$\% \text{ std dev} = \frac{\sigma}{\text{épaisseur d'oxyde postCMP}} \quad (4.22)$$

*) Ecart type de l'épaisseur réellement enlevée :

$$\sigma_{\text{ep enlevée}} = \sigma(\text{épaisseur PreCMP} - \text{épaisseur postCMP}) \quad (4.23)$$

*) Ecart type normalisé de l'épaisseur réellement enlevée :

$$\% \text{ std.ep.enlev} = \frac{\sigma(\text{épaisseur Pre CMP} - \text{épaisseur postCMP})}{\text{épaisseur enlevée moyenne}} \quad (4.24)$$

*) D'autres indicateurs peuvent être utilisés comme l'écart type et l'écart type normalisé de la vitesse d'enlèvement.

4.3.5 La distribution de la pression

Pour étudier la distribution de la pression sur la plaque, nous nous appuierons sur les travaux scientifiques qui ont été développés sur les équipements CMP (300mm⁶³). Car peu de travaux sur ce sujet ont été réalisés sur les équipements 200mm. Ceci peut s'expliquer par le fait que la maturité du procédé CMP en fabrication des semi-conducteurs a coïncidé avec l'apparition des équipements 200mm et aussi par le fait que les travaux scientifiques visent des problématiques rencontrées dans les technologies très avancées produites souvent sur des équipements 300mm. Néanmoins, dans les deux cas, les principes physiques de base du procédé CMP restent les mêmes.

4.3.5.1 La distribution de la pression à travers la plaque pendant le procédé CMP

La distribution de la pression au niveau de l'interface tête de polissage / plaque - tissus de polissage est modélisée. Plusieurs travaux de modélisation ont été faits pour la détermination de la pression au niveau de cette interface. Nous pouvons distinguer 2 modèles principaux.

- Modèle contact solide- solide où la pression exercée est importante alors que la vitesse est basse.
- Modèle contact solide- fluide- solide où la vitesse est importante alors que la pression est basse.

Le modèle choisi dans notre étude est le modèle contact solide – solide. Le choix de ce modèle est justifié par le fait que la pression exercée par les abrasifs se trouvant entre la

⁶³ Sur le plan industriel, la réalisation de boucles de régulation en 300 mm ne posent pas de gros problèmes car, au niveau des équipements de CMP, on a la possibilité de réajuster les pressions de chaque zone de la plaque. Donc, il suffit d'analyser le profil de la plaque précédente pour pouvoir réajuster la pression sur une zone bien définie.

plaqué et le tissus de polissage (modèle solide- fluide - solide) est négligeable devant la pression exercée par les têtes de polissages [MURTHY 1997].

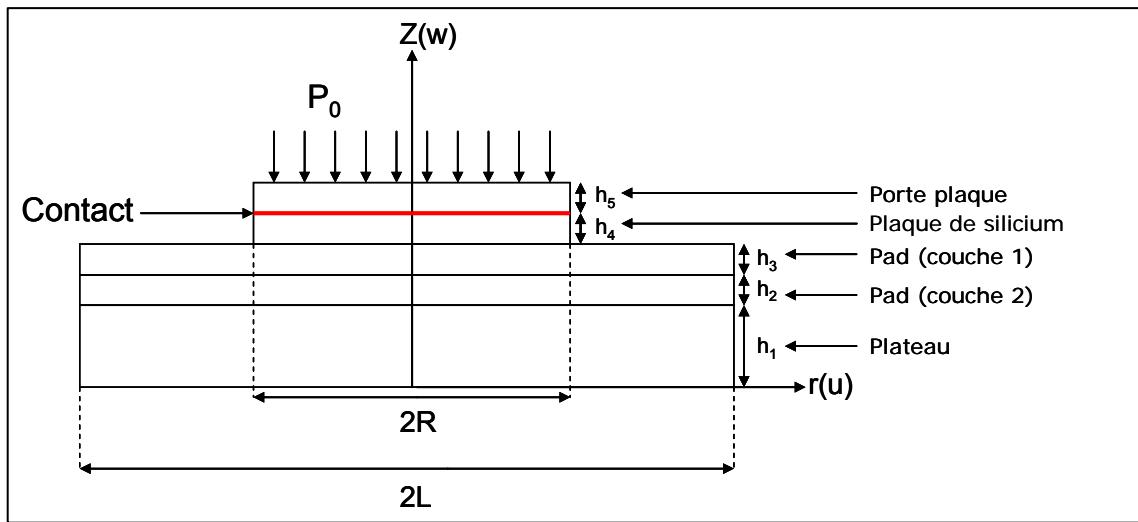


Figure 4.43- Schématisation de l'interface plaque/pad

Les matériaux (tête, anneau de maintien et plaque de silicium) sont considérés comme isotropes et déformables dans le régime élastique. Le tissu de polissage (pad) est modélisé comme un matériau composé de plusieurs couches avec des déplacements à l'interface (radial et axial) comme cela est illustré sur la figure 4.43. L'anneau de maintien est supposé rigide, ayant la même constante élastique que la plaquette de silicium. Les déformations de la plaque et du tissus de polissage sont asymétriques [XIN 1998]

Les composantes des contraintes mécaniques générées au niveau du contact (plaqué/tissus de polissage) peuvent être calculées [XIN 1998]. Les simulations de LUO [LUO 2003c] effectuées sur la répartition de la pression sur la plaque ont montré que la pression est plus importante sur les bords des plaques comme cela est illustré sur la figure 4.44.

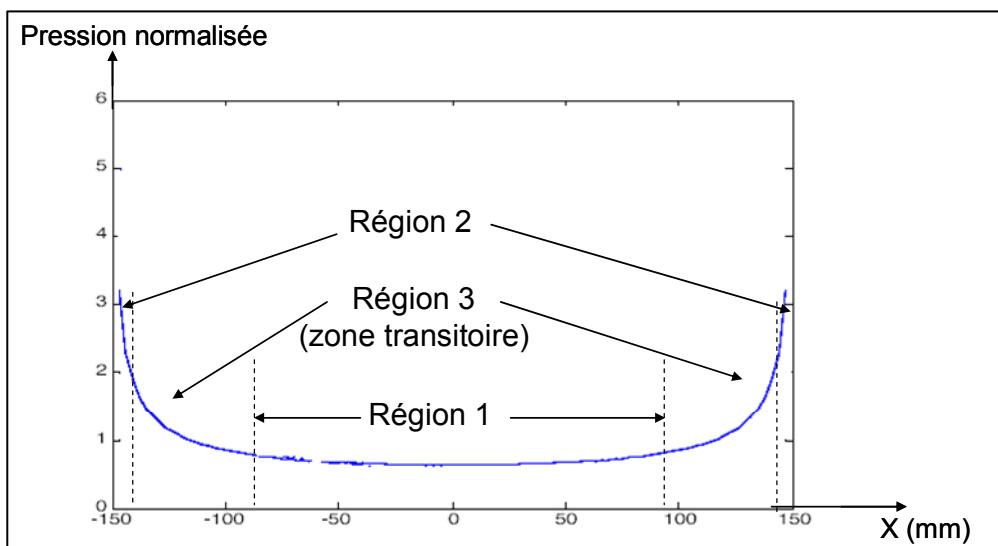


Figure 4.44- Distribution de la pression à travers la plaque [Luo 2003c]

Sur la figure 4.44, on distingue 3 régions ; la première région correspond à une pression distribuée uniformément sur une grande partie de la plaque, la seconde est une augmentation abrupte de la pression sur les bords de la plaque et la troisième région est une zone transitoire entre les deux premières régions. Notons que les plaques simulées par LUO sont de 300 mm, alors que les plaques étudiées dans ce manuscrit sont de 200 mm, mais les phénomènes physiques restent les mêmes. L'augmentation de la pression aux bords de la plaque peut s'expliquer par la différence de comportement des différentes régions face au procédé du CMP.

4.3.5.2 L'effet de la pression de l'anneau de maintien (pressure retaining ring)

La machine de polissage (en vue de face) est illustrée sur la figure 4.45. On distingue les 3 sources de pressions :

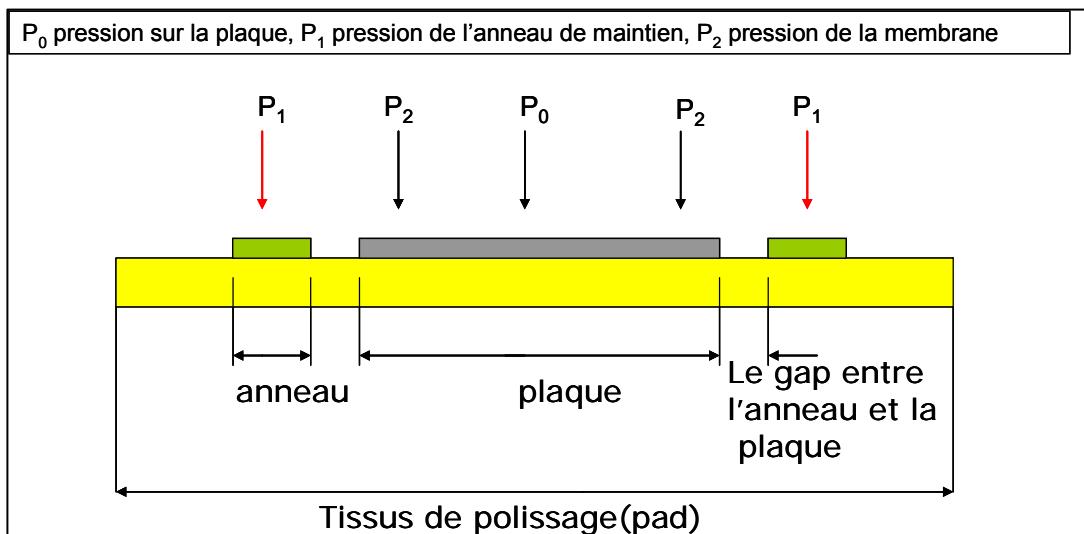


Figure 4.45- Coupe d'une machine de polissage CMP

- la pression principale : c'est la pression exercée au niveau du centre de la plaque.
- la pression de l'anneau de maintien agit sur les quelques mm des bords.
- la pression de la membrane : la valeur de cette pression est choisie de sorte à ce que la plaque ne soit pas éjectée du plateau. Généralement, il est utile de calculer la pression réellement appliquée sur l'anneau de maintien de façon à ce que le rapport :

$$\frac{P_{\text{réellement appliquée sur la surface de l'anneau}}}{P_{\text{membrane}}} \geq 0,5 \quad (4.25)$$

La plaque sera éjectée du plateau si la valeur du rapport est inférieure à 0,5. La valeur de ce rapport (0,5) correspond à valeur limite (limitations de la machine) donnée par le constructeur de la machine. De même pour que la plaque soit bien maintenue, la pression de l'anneau de maintien peut être calculée à l'aide de l'équation suivante :

$$P_{\text{réellement appliquée sur la surface de l'anneau}} = \alpha * P_{\text{anneau de maintien}} - \beta * P_{\text{membrane}} \quad (4.26)$$

Où : α et β sont des constantes. Les valeurs de ces constantes peuvent être déduites à partir d'une série d'expériences.

Les variations des valeurs de la pression de têtes d'anneaux de maintien modifient directement le profil de la pression à travers la plaque. Notre objectif est d'essayer d'obtenir un modèle permettant un réajustement de la pression de l'anneau de maintien pour avoir le même profil de la pression sur les bords. Nous avons pu démontrer à travers des analyses des données de mesures que le profil du CMP sur les bords dépend fortement de la pression et va impacter directement le profil de la hauteur de marche [BELHARET 2006].

4.3.6 Modélisation du procédé

4.3.6.1 les cartographies des mesures

4.3.6.1.a les cartographies des mesures pré et post CMP

Nous avons mentionné dans la partie précédente de ce chapitre consacré au procédé HDP-CVD que pour étudier avec finesse le profil d'oxyde après dépôt sur la plaque, il faut disposer d'un nombre suffisant de points de mesures couvrant toute la surface de la plaque. Nous avons utilisé les cartographies contenant 49 points de mesures. Ce type de cartographie est insuffisant pour étudier le profil d'oxyde après CMP car on s'intéressera en particulier aux extrêmes des bords de la plaque (jusqu'à 2mm). Il faut donc concevoir des cartographies de mesure capables de couvrir toute la surface de la plaque. La figure 4.46a montre la cartographie de 460 points proposée pour étudier le profil d'oxyde. Cette cartographie donne une très bonne satisfaction à nos besoins, néanmoins pour étudier un grand nombre de profils de plusieurs plaques, cela prendra un temps considérable.

Nous avons mentionné dans le chapitre 1 les applications possibles du logiciel Wafer-Fit. L'un des avantages qu'offre ce logiciel est l'optimisation des cartographies des mesures (réduction du nombre de points de mesures en quelques points tout en ayant la même qualité des indices statistiques). Nous avons réduit le nombre de points de mesure de 460 à 109 points de mesures (figure 4.46b) tout en gardant le même profil. L'optimisation de cette cartographie est rendue possible par l'élimination d'un certain nombre de points grâce en particulier à l'analyse de la symétrie des zones.

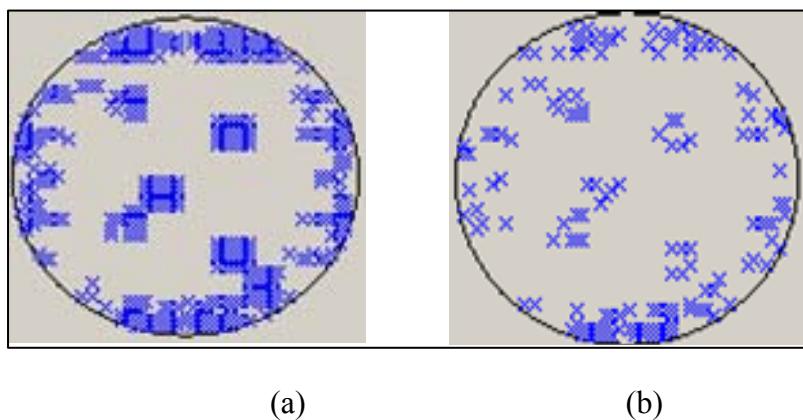


Figure 4.46- a) cartographie de mesure de 460 points b) cartographie de mesure réduite à 109 points

La nouvelle cartographie réalisée au cours de cette étude peut être utilisée en production pour des tâches de qualification. Il faut noter que les cartographies de mesures

existantes en production dédiées à cette tâche contiennent au maximum 70 points et elle ne couvre pas les 5 derniers mm des bords.

4.3.6.1.b Les cartographies des mesures de la hauteur de marche (Mesures AFM)

La méthode de mesure de la hauteur de marche repose sur le principe de la microscopie à force atomique (AFM : Atomic Force Microscopy). La mesure avec cette méthode est très lente, par exemple, pour une mesure d'une plaque avec une cartographie de 9 points de mesures, cela prend en moyenne une vingtaine de minutes. La mesure de la hauteur de marche d'un lot avec une cartographie complète (d'une centaine de points) paraît donc impossible. En vue de mesurer toutes les plaques qui serviront pour nos études expérimentales et aussi avoir une couverture de toute la surface de la plaque, nous avons opté pour une cartographie contenant 11 points de mesures. Cette cartographie est représentée sur la figure 4.47.

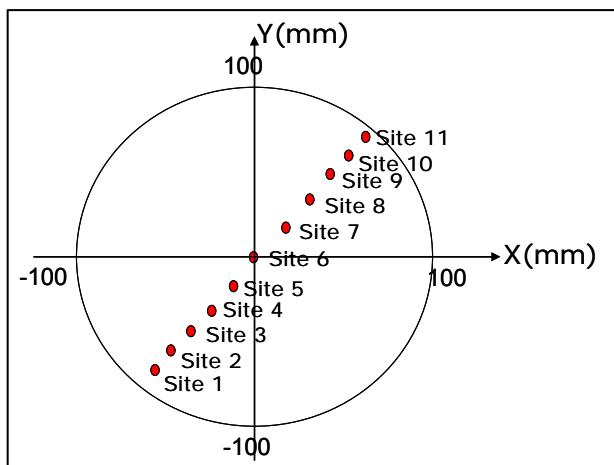


Figure 4.47- Exemple de cartographie de mesure de type radial des mesures de la hauteur de marche

4.3.6.2 Définition des plages de pression de l'anneau de maintien

Nous avons identifié la pression de l'anneau de maintien comme le paramètre critique du procédé sur la variation du profil d'oxyde après CMP en particulier sur les bords. Ce sera donc le seul paramètre du procédé ajustable pour les plans d'expériences conçus. Les plages d'un paramètre de procédé ajustable doivent répondre aux limitations de l'équipement mais aussi aux plages souhaitées des paramètres de sortie. Le paramètre de sortie est la mesure de la hauteur de marche. De ce fait, 6 plaques ont suivi un enchaînement de fabrication standard lors de la réalisation de la brique STI jusqu'à l'étape de procédé CMP, où chacune des plaques a vu sa pression de l'anneau de maintien modifiée comme l'illustre le tableau 4.5. Il faut rappeler qu'en microélectronique et en particulier dans le procédé CMP, il est d'usage d'exprimer la pression en PSI (Pound per Square Inch) sachant que 1 bar = 14,5 Psi.

Plaque	Pression Anneau de maintien (Psi)
1	4,8
2	5,8
3	6,9
4	8
5	9
6	10,5

Tableau 4.5- Conditions expérimentales

L'effet de la pression de l'anneau de maintien est très important sur la nature du profil final de la hauteur de marche. Suivant la valeur de la pression de l'anneau de maintien, on agit sur le profil de la hauteur de marche sur la plaque. En appliquant une pression minimale de 4,8 Psi pour la plaque 1, on a un profil de type convexe. Par contre en appliquant une pression assez importante (10,5 Psi) pour la plaque 6, le profil de la hauteur de marche devient concave comme cela est illustré sur la figure 4.48. A partir de ces deux résultats, on peut conclure que l'ajustement de la pression de l'anneau de maintien est un très bon moyen pour contrôler le profil de la hauteur de marche.

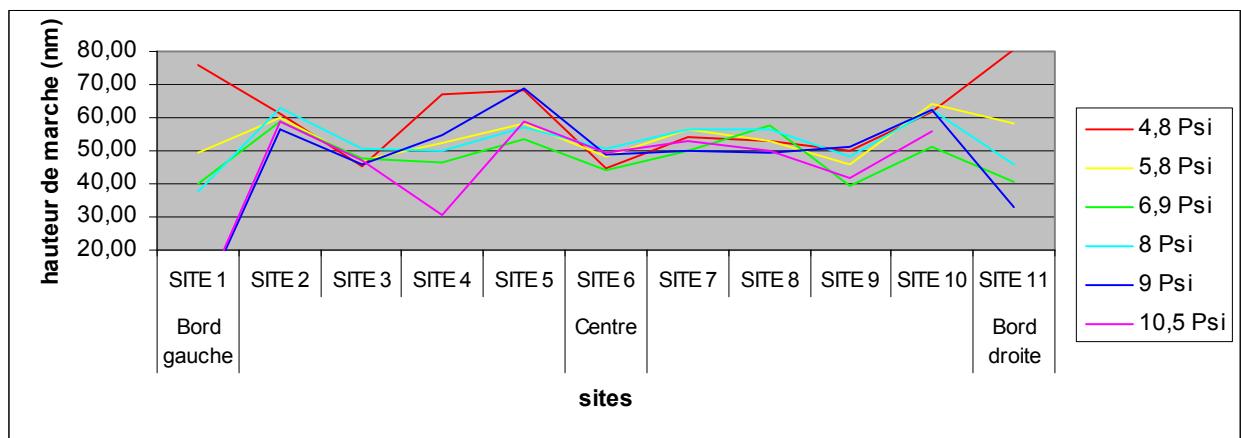


Figure 4.48- Variation de la hauteur de marche au niveau du produit suivant les valeurs de pression de l'anneau de maintien

Pour caractériser l'uniformité de la hauteur de marche, nous utilisons un des indicateurs dédiés à la caractérisation de l'uniformité : il s'agit de l'écart type normalisé. Les valeurs de l'écart type normalisé de la hauteur de marche sont données dans le tableau 4.6. Les fortes valeurs de l'indicateur sont observées pour les valeurs extrêmes de la pression de l'anneau de maintien.

PLAQUE	Ecart type normalisé de la hauteur de marche %stddev (%)
1	19,9
2	11
3	14,2
4	14,2
5	37,2
6	38,4

Tableau 4.6- Ecart type normalisé de la hauteur de marche des plaques

Ces pressions extrêmes dégradent donc l'uniformité de la hauteur de marche. Elles sont donc exclues pour le plan d'expériences. Les valeurs de l'anneau de maintien qui sont sélectionnés varient de 5,6 à 8,2 Psi.

4.3.6.3 Conditions expérimentales

On cherche à travers ce plan d'expériences à modéliser l'étendue (range) de l'épaisseur d'oxyde post-CMP avec la pression de l'anneau de maintien. Le plan d'expériences choisi est du premier ordre car nous avons une seule variable (paramètre ajustable). Le paramètre ajustable est la pression de l'anneau de maintien. Le nombre de plaques choisi est de 15. Ce choix est justifié par l'intervalle des valeurs sélectionnées de la pression. Les valeurs de la pression de l'anneau de maintien qui sont sélectionnées varient de 5,6 à 8,2 Psi avec un pas de 0,1 Psi. Les valeurs de la pression de la membrane est de 4 Psi. Cette valeur est suffisante dans tous les cas expérimentaux envisagés pour que la plaque ne soit pas éjectée du plateau de polissage. La pression centrale reste fixe, elle est de 6 Psi.

Les plaques sélectionnées ont suivi un enchainement de fabrication lors de la réalisation de la brique STI jusqu'à l'étape du CMP où chaque plaque a vu sa pression de l'anneau de maintien modifiée. Pour éviter un effet de plaque à plaque lié aux différentes chambres des équipements concernés par les procédés antérieurs, nous avons veillé à ce que les plaques soient traitées dans la même chambre de dépôt lors du procédé HDPCVD, ainsi que dans le même four lors du dépôt TEOS. En vue d'étudier l'impact du profil d'oxyde pré-CMP, une partie des plaques a été processée dans la chambre A de la machine de gravure et une autre partie dans la chambre B. Les conditions expérimentales sont résumées dans le tableau 4.7. Au niveau du procédé CMP, chaque plaque est processée sur un plateau individuellement.

Plaques	Plateau de polissage	Tête de polissage	Pression appliquée (Psi)	Pression membrane (Psi)	Pression Anneau de maintien (Psi)
Plaque 1	3	4	6	4	5,6
Plaque 2	2	1	6	4	5,8
Plaque 3	1	2	6	4	6
Plaque 4	3	3	6	4	6,2
Plaque 5	2	4	6	4	6,5
Plaque 6	1	1	6	4	6,7
Plaque 7	3	2	6	4	6,8
Plaque 8	2	3	6	4	6,9
Plaque 9	1	4	6	4	7
Plaque 10	3	1	6	4	7,1
Plaque 11	2	2	6	4	7,3
Plaque 12	1	3	6	4	7,5
Plaque 13	3	4	6	4	7,7
Plaque 14	2	1	6	4	8
Plaque 15	1	2	6	4	8,2

Tableau 4.7- Conditions expérimentales des plaques sélectionnées pour le plan d'expériences

4.3.6.4 Analyses des résultats

4.3.6.4.a Analyse du profil d'oxyde pré-CMP

Le profil d'oxyde dans la tranchée a été analysé à partir des données de mesures effectuées sur toutes les plaques à l'aide de cartographie de mesure optimisée de 109 points. La figure 4.49 illustre les profils d'oxyde avant CMP pour les 6 premières plaques.

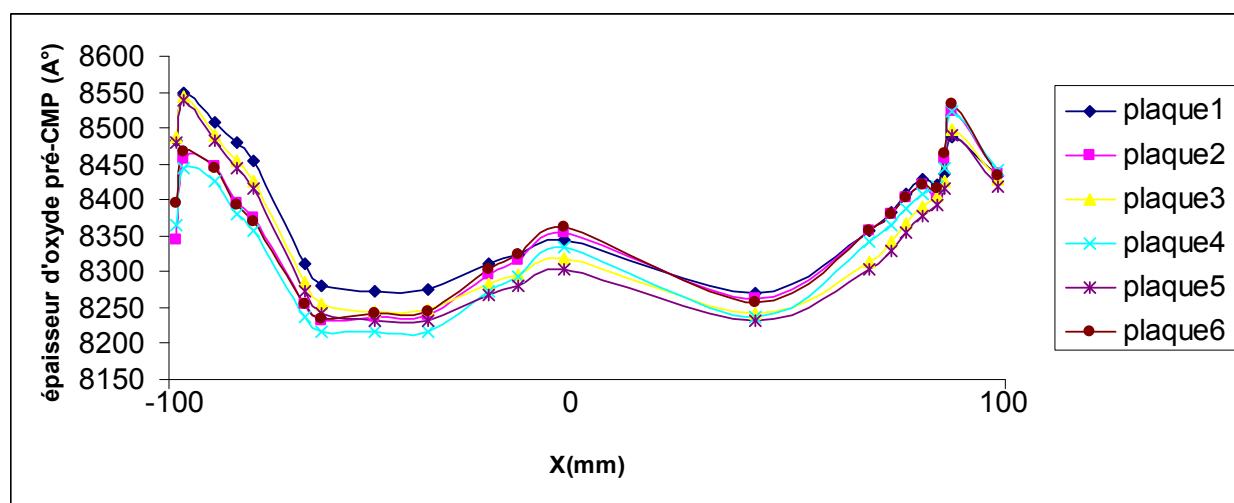


Figure 4.49- Evolution de l'épaisseur d'oxyde pré- CMP au sein de la plaque

Cette représentation illustre la répartition de l'oxyde à travers la plaque (centre – bords). Il en ressort que le profil d'oxyde est quasi-plat avec un écart type normalisé de 0,93 à 1,31. Tout de même, on distingue deux populations avec des ranges différents. Cet effet observé de plaque à plaque est attribué à ce stade à la gravure sèche (2 chambres de l'équipement de gravure A et B) comme nous pouvons le constater dans le tableau 4.8.

Plaques	Chambre de gravure	Ecart type normalisé de l'étendue d'épaisseur d'oxyde avant CMP
Plaque 1	A	1,22
Plaque 2	B	0,95
Plaque 3	A	1,29
Plaque 4	B	0,98
Plaque 5	A	1,31
Plaque 6	B	0,93
Plaque 7	A	1,25
Plaque 8	B	0,93
Plaque 9	A	1,28
Plaque 10	B	0,96
Plaque 11	A	1,30
Plaque 12	B	1
Plaque 13	A	1,26
Plaque 14	B	0,93
Plaque 15	A	1,23

Tableau 4.8- Ecart type normalisée de l'étendue d'épaisseur d'oxyde avant CMP.

Les fortes valeurs d'oxyde pré-CMP peuvent être expliquées par le fait que cet oxyde est composé d'une couche d'oxyde HDP-CVD (valeur cible 5500Å) et d'une couche TEOS d'environ 3000Å.

Nous pouvons constater dans le tableau 4.8 que l'étendue de l'épaisseur d'oxyde avant CMP d'une plaque dépend fortement des conditions du procédé de la gravure sèche (passage dans la chambre A ou B de l'équipement de gravure). C'est l'effet chambre (le même effet que nous avons déjà discuté dans la partie 4.2.4.2).

4.3.6.4.b Analyses du profil d'oxyde post-CMP

Comme l'effet de l'anneau de maintien n'a un effet que sur les bords des plaques, pour les analyses post- CMP, nous ne nous intéressons qu'aux bords des plaques (25 derniers mm) car, dans la région centrale, le profil d'oxyde post CMP est presque plat.

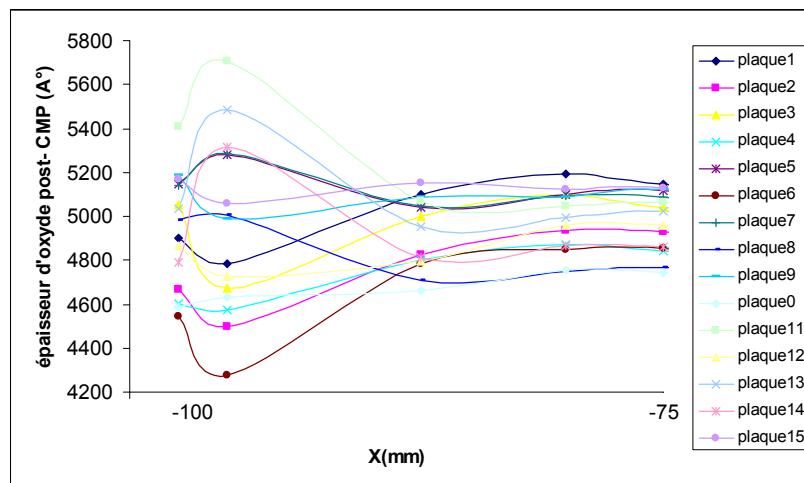


Figure 4.50- Evolution de l'épaisseur d'oxyde post- CMP au sein de la plaque (sur les 25 derniers mm)

L'épaisseur d'oxyde peut varier de 500Å en bord de plaque comme nous pouvons le constater sur la figure 4.50 (plaqué 1 et plaqué 6). Cette variation d'épaisseur est due à la variation de la pression exercée sur la plaque (bord-centre). Suivant la pression de l'anneau de maintien, le profil d'oxyde post-CMP sur les bords a deux formes (convexe et concave). Nous constatons également que le point (pic) correspondant à l'épaisseur maximale d'oxyde sur les bords décroît avec l'augmentation de la pression de l'anneau de maintien. Ce résultat vient confirmer les résultats des travaux effectués sur l'influence de la pression de l'anneau de maintien sur la pression globale sur la plaque ([SHENDON 1991, JACKSON 1995, GURTHRIE 1996, SHAMOUILIAN 1996, KIMURA 1998, XIN 1998]). L'ordre de décroissance n'est pas respecté car le temps effectif de polissage est différent d'une plaque à une autre et en plus comme chaque plaque est traitée sur un plateau individuellement, il y a un effet plateau qui se traduit par une différence de vitesse d'enlèvement d'un plateau à un autre (la vitesse d'enlèvement au centre $V_{enlev. centre}$ est de 32 Å/sec pour le plateau 1, 34Å/sec pour le plateau 2 et 33 Å/ sec pour le plateau 3). On traitera donc les plaques séparément par plateau.

Les résultats de DOE nous ont été utiles pour mieux comprendre la variation de l'épaisseur d'oxyde enlevé pendant le procédé du CMP sur toute la plaque et en particulier sur les bords pour les différentes pressions choisies.

4.3.6.5 Modélisation de l'étendue avec la pression de l'anneau de maintien

Connaissant les temps de polissage, nous avons calculé les vitesses de polissages pour tous les points de la plaque. Pour le calcul de l'étendue, nous allons nous contenter d'un point sur le bord (pic correspondant à l'épaisseur maximale) et du point centre de la plaque. La vitesse d'enlèvement au centre n'est pas impactée par la variation de la pression de l'anneau de maintien tandis que, la vitesse d'enlèvement sur les bords décroît linéairement avec la pression de l'anneau de maintien appliquée.

Nous voulons mettre en évidence l'étendue de la vitesse de polissage en fonction de la pression.

Soit :

- $V_{enlev. c}$ est la vitesse d'enlèvement au centre et $V_{enlev. b}$ est la vitesse d'enlèvement sur les bords
- t_0 est le temps de polissage
- l'étendue (R) est l'écart entre les épaisseurs enlevées au centre et sur les bords ($X_{enlev.c}$ et $X_{enlev.b}$).
- $V_{enlev. b0}$ est la vitesse de l'anneau de maintien initiale.

$$V_{enlev.b} = \alpha.P + V_{enlev.b0} \quad (4.27)$$

Où α est la valeur de la pente de la courbe donnant la vitesse d'enlèvement sur les bords de la plaque en fonction de la pression.

En remplaçant $X_{enlev.b}$

$$\frac{X_{enlev.b}}{t_0} = \alpha.P + V_{enlev.b0}$$

$$X_{enlev.bords} = (\alpha \cdot P + V_{enlev.b0}) * t_0 \quad (4.28)$$

Nous avons : $X_{enlev.c} = V_{enlev.c} t_0$ et donc

$$R = -\alpha t_0 P + (V_{enlev.c} - V_{enlev.b0}) t_0 \quad (4.29)$$

Comme nous l'avons indiqué dans le paragraphe précédent, la vitesse d'enlèvement au centre ($V_{enlev.c}$) est différente d'un plateau de polissage à un autre, l'étendue (R) varie d'un plateau à un autre. Nous allons donc appliquer 3 modèles différents. Nous allons donner le modèle l'étendue pour le plateau 1.

$V_{enlev.c} = 32 \text{ \AA/sec}$, $V_{enlev.b0} = 39 \text{ \AA/sec}$

La pente α a été calculée ($\alpha = -3,33$)

Le modèle de l'étendue (R) pour le plateau 1 est donnée par la relation suivante :

$$R = 3.33t_0P - 7t_0 \quad (4.30)$$

Nous concluons que l'étendue dépend du temps de polissage [NISHIGUCHI 2001] et de la pression. La vitesse d'enlèvement est dépendante aussi de la durée de vie et des propriétés des tissus de polissage (pads) et comme toutes les plaques ont été traitées sur des pads différents (effet plaque), les analyses ont été faites séparément. Nous avons pu constater un « effet tissus de polissage ».

Pour une meilleure synthèse d'un modèle très robuste, la meilleure solution est de traiter les plaques sur un seul plateau de polissage et de refaire le plan d'expériences au moins deux fois pour tenir en compte de la durée de vie du plateau de polissage (en début et en fin de durée de vie du tissus de polissage).

4.3.7 Corrélation du profil d'oxyde post-CMP avec les résultats des tests électriques

Les plaques qui ont servi pour la réalisation du DOE ont suivi par la suite les procédés standards post STI (module grille, interconnections,...). Les plaques ont été testées en test électrique. Comme nous avons déjà démontré dans le chapitre 3, la tension de seuil des transistors parasites (V_t) a servi d'indicateur de suivi des boucles de régulation R2R.

Pour les études de corrélations, nous avons analysé les données de mesure de ces plaques dont la valeur cible de l'épaisseur STI est de 4000Å et l'énergie d'implantation des caissons est de 600 keV.

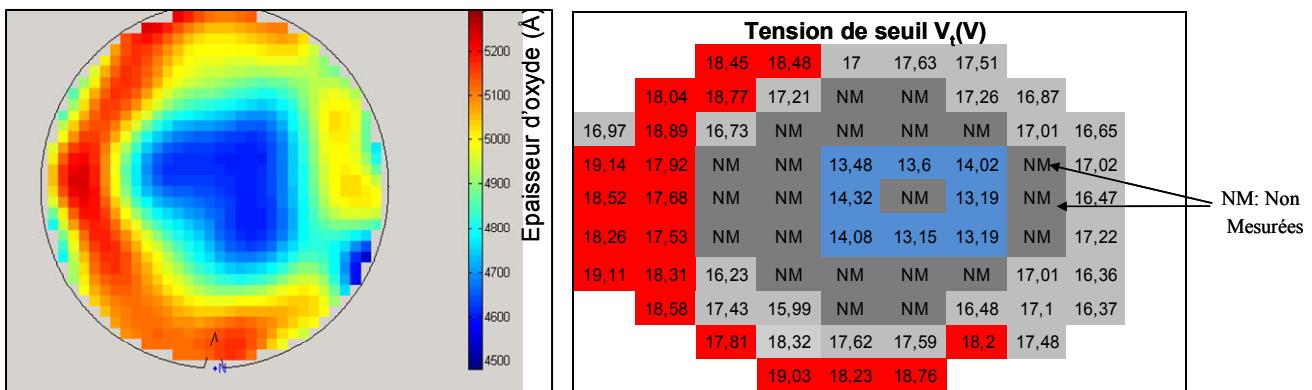


Figure 4.51- Corrélation entre le profil d’oxyde post CMP et les résultats de test électrique sur le paramètre de tension de seuil

On observe une forte corrélation [BELHARET 2007e] entre le profil d’oxyde post-CMP et le profil des résultats de V_t au niveau du test électrique où les fortes valeurs de V_t correspondent aux fortes valeurs de l’épaisseur d’oxyde post-CMP. Ceci démontre une fois de plus la robustesse de l’indicateur de STI (V_t). Grâce à ce résultat, il est donc possible de prédire les valeurs de V_t sur toutes les régions de la plaque. Les résultats des expériences montrent qu’en choisissant certaines valeurs de l’anneau de maintien, on améliore la dispersion de V_t . Il faut noter que la non symétrie du profil d’oxyde post-CMP peut s’expliquer par le profil de la gravure sèche déjà évoqué.

4.3.8 Corrélation entre le profil d’oxyde post- CMP et les défauts générés

L’apparition de particules et de défauts sur la surface de la plaque après une opération de fabrication est très fréquente en microélectronique. Certains de ces défauts peuvent être fatals pour le rendement final électrique des puces, tandis que d’autres ne sont pas très influents. Pour cette raison, des opérations d’inspection des défauts sont faites après un certain nombre d’opérations de fabrication. Le procédé CMP est un des procédés contributeurs à l’apparition de défauts sur la surface de la plaque. Juste après avoir réalisé la brique STI, les plaques qui ont servi à la réalisation du plan d’expériences ont été contrôlées en défectivité. Nous avons essayé de comparer les résultats de défectivité avec les résultats de mesure au niveau du STI après CMP. La figure 4.51 compare un profil d’oxyde post CMP avec celui de l’inspection de la défectivité.

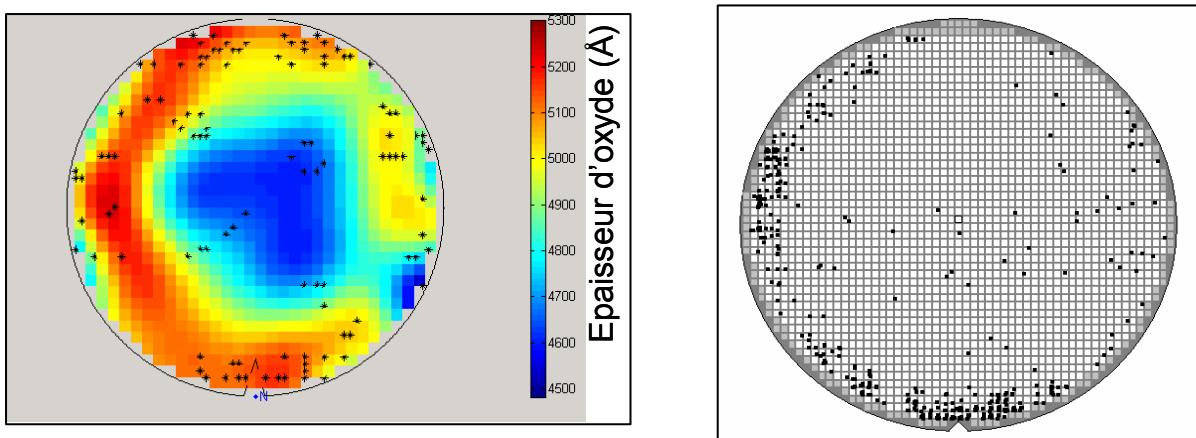


Figure 4.52- Comparaison du profil d'oxyde après CMP avec le profil de la densité des défauts

Nous observons une forte densité de défauts sur les bords des plaques et en particulier dans les zones les plus denses comme illustré sur la figure 4.52. Les épaisseurs importantes au niveau du CMP induisent un grand nombre de défauts [BELHARET 2007e]. Cet effet a été constaté sur toutes les plaques (15 plaques). Nous pouvons expliquer ce phénomène par l'effet de l'avivage car comme on l'a précisé au début de cette partie, l'avivage sert à maintenir la présence d'aspérités à la surface du tissu. Sans lui, tous les pores seraient remplis par les débris de l'usure après le polissage de quelques plaquettes et la circulation d'abrasif dans l'interface ne s'effectuerait plus correctement, dégradant fortement l'uniformité du procédé. La figure 4.53 illustre un cas de création de défauts, il s'agit de la condensation des grains de particules des abrasifs. Le contact plaque/tissus de polissage dans les endroits denses de la plaque (épaisseurs maximales de l'oxyde de dépôt) fait durcir le tissu de polissage en créant une condensation des particules de l'abrasif comme cela est illustré sur la figure 4.53.

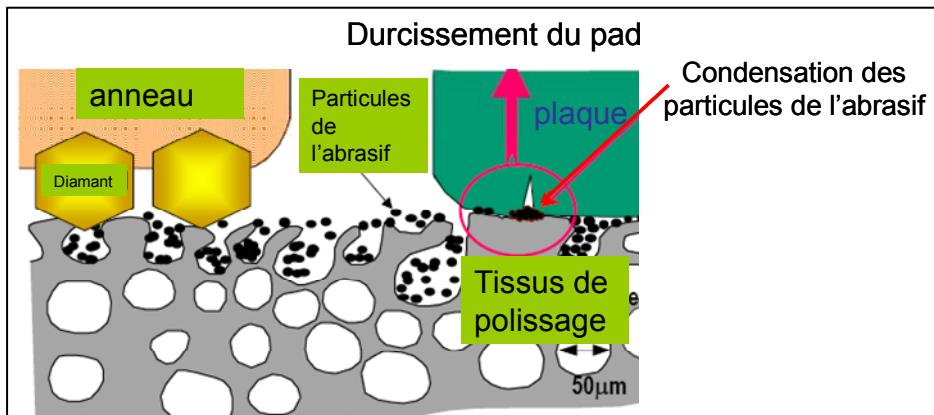


Figure 4.53- Condensation des particules de l'abrasif dans certaines zones du tissus de polissage en contact avec la plaque

De ce fait, il est désormais possible de prédire la densité de défauts connaissant le profil d'oxyde post-CMP. Comme le profil d'oxyde post CMP dépend du temps de polissage et de la pression de l'anneau de maintien avec un bon choix de ces deux paramètres, il est

possible de réduire la densité de défauts sur la surface de la plaque et ainsi améliorer le rendement.

4.3.9 Conclusion

La non uniformité de l'épaisseur d'oxyde post CMP et par conséquent la non uniformité de la hauteur de marche impactent directement les caractéristiques électriques et notamment la tension de seuil des transistors parasites (V_t). Cette non uniformité au niveau du module STI est attribuée au procédé de gravure sèche et CMP. La non uniformité produite après la gravure sèche est attribuée à l'effet des chambres de l'équipement. Alors que la non uniformité générée par le CMP est attribuée à la distribution de la pression sur la plaque. Nous avons démontré que la modification de la pression de l'anneau de maintien permet de modifier la pression globale sur la plaque et en particulier sur les bords. Un réajustement de la pression de l'anneau de maintien peut réduire la dispersion de l'uniformité de l'oxyde sur la plaque. Nous avons démontré qu'il y a une forte corrélation entre les mesures d'oxyde post-CMP et les paramètres électriques. Ce résultat nous permet de prédire les valeurs de V_t à partir des résultats d'oxyde post CMP. Ce résultat est extrêmement important pour la prédition de rendement. Nous avons également démontré que la corrélation entre les mesures d'oxyde post-CMP et les densités de défauts est très forte, l'effet observé montre que les défauts sont concentrés sur les bords de plaques pour les zones présentant des épaisseurs d'oxyde très importantes. Ce résultat constitue également un résultat très important car la répétitivité de cet effet sur toutes les plaques nous permettra de prédire la densité de défauts à partir des résultats des épaisseurs d'oxyde post-CMP.

Partie 3

Le polissage mécanico-chimique et la gravure humide

Dans cette partie, nous exposerons la démarche suivie pour la réalisation de la boucle de régulation entre le procédé CMP et la gravure humide. Nous allons exposer le contexte de l'étude que nous allons mener. Nous verrons que les variabilités des mesures post-CMP impactent fortement la hauteur de marche. Pour minimiser la variabilité de la hauteur de marche, on peut agir sur un procédé intermédiaire. Dans notre cas : la gravure humide. Il s'agit de définir dans un premier temps l'étape critique de l'opération de gravure humide et les paramètres critiques affectant la variabilité de la hauteur de marche. Nous nous intéresserons de près aux mécanismes de gravure humide et aux différentes composantes de l'équipement pour bien comprendre les paramètres et les étapes critiques. Nous détaillerons notre démarche expérimentale pour déduire un modèle qui lie la hauteur de marche avec les mesures post CMP et les paramètres de la gravure humide. Nous allons exposer également les premiers résultats de la pré- production et enfin, nous allons proposer une méthode assez originale pour sécuriser la boucle proposée quant à d'éventuelles dérives qui peuvent survenir sur les équipements.

4.4.1 Introduction

Les procédés de fabrication de la brique STI ont été décrits en détail au chapitre II. Nous nous intéressons dans cette partie qu'aux deux dernières étapes de la chaîne de fabrication de cette brique à savoir le polissage mécanico-chimique (CMP) et la gravure humide. Plusieurs modélisations de la brique STI ont été rapportées dans la littérature. Il s'agit généralement de modéliser la mesure de la hauteur de marche avec les différents procédés de fabrication. La figure 4.54 illustre une modélisation simple de la hauteur de marche avec les deux dernières étapes du procédé (CMP et gravure humide).

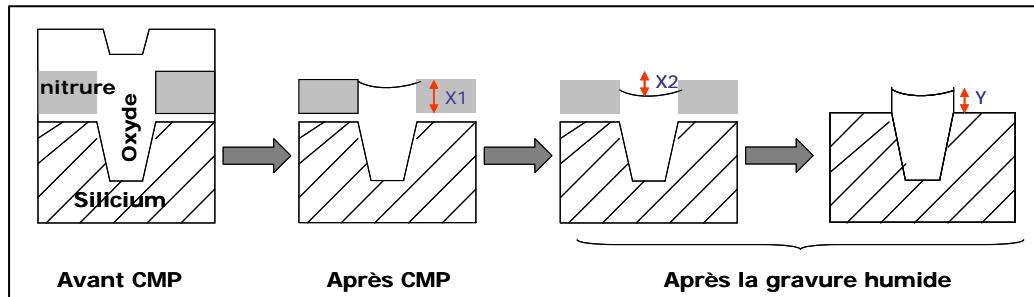


Figure 4.54- Représentation simplifiée des étapes des procédés STI

Où : X₁ représente l'épaisseur de nitrate après CMP (mesurée par ellipsométrie)
X₂ représente l'épaisseur d'oxyde enlevé après gravure humide (déduit par ellipsométrie)
Y représente la hauteur de marche (mesurée par AFM).

Nous constatons que la hauteur de marche dépend de l'épaisseur du film de nitrate et de l'épaisseur d'oxyde dans la tranchée. Une dispersion de l'un des deux paramètres influence directement la valeur de la hauteur de marche. Pour garantir des valeurs de hauteur de marche proches de la valeur cible, les dispersions des mesures post- CMP peuvent être compensées au niveau de la gravure humide en modifiant un paramètre de procédé (par exemple le temps de gravure d'oxyde). Cette méthode est employée par la plupart des compagnies d'une manière manuelle. La réalisation d'une boucle R2R de type feed-forward au niveau de ces deux procédés permettra un réajustement automatique de(s) paramètre(s) de gravure humide comme cela est illustré sur la figure 4.55.

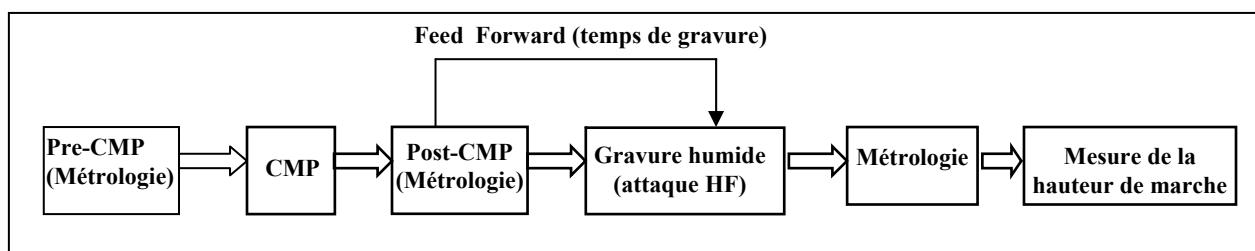


Figure 4.55- Schéma synoptique de la boucle CMP- gravure humide

La boucle feed-forward exploite les informations des mesures Post- CMP pour ensuite ajuster un paramètre de la gravure humide (temps de gravure par exemple). Il est à rappeler qu'au niveau de l'opération de la gravure humide, deux lots regroupés ensemble dans un bac⁶⁴ sont traitées simultanément, on parle alors de mode batch. Le grand obstacle ayant empêché la

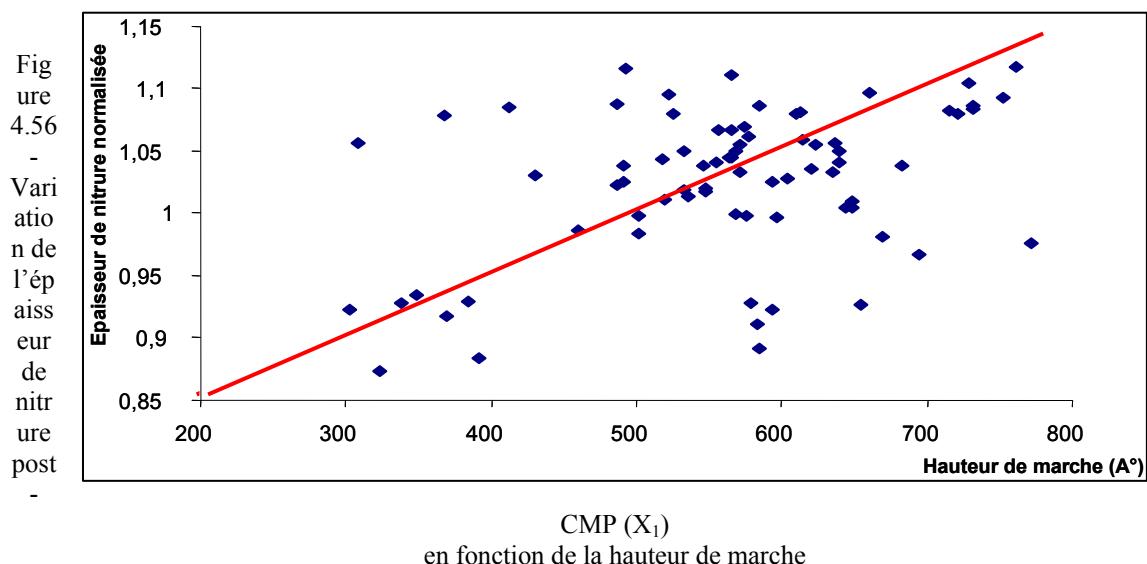
⁶⁴ Ensemble de deux lots, on parle aussi de batch.

plupart des compagnies de se lancer dans cette réalisation est le coût de l'opération de basculement du mode batch au mode lot, autrement dit, les contraintes de chargement des équipements de gravure humide. Dans notre cas, une étude préliminaire a été effectuée et a montré que sur les produits étudiés, le passage du mode batch au mode lot n'affecte pas la capacité de chargement des équipements de gravure humide. Nous considérons que la réalisation de cette boucle est une première à cette échelle de réalisation des boucles de régulation au niveau de la brique STI.

4.4.2 Contexte de l'étude

Des analyses de données ont été faites sur un historique de données de mesures (28 lots). Les paramètres analysés sont les mesures post- CMP (épaisseur d'oxyde et de nitre), les mesures de métrologie après la gravure humide et les mesures de hauteurs de marche. Il s'agit de corrélérer la hauteur de marche avec les données de mesures. Les données des deux modes STI sont exploitées.

4.4.2.1 Corrélation des mesures post-CMP avec la hauteur de marche dans le cas du STI standard



Nous observons sur la figure 4.56 que les dispersions de l'épaisseur de nitre impactent directement la hauteur de marche et donc une bonne corrélation entre les mesures post-CMP et la hauteur de marche.

4.4.2.2 Corrélation des mesures post-CMP avec la hauteur de marche dans le cas du Direct STI

La corrélation épaisseur de nitre - hauteur de marche est plus forte dans le cas du Direct STI comme c'est montré sur la figure 4.57.

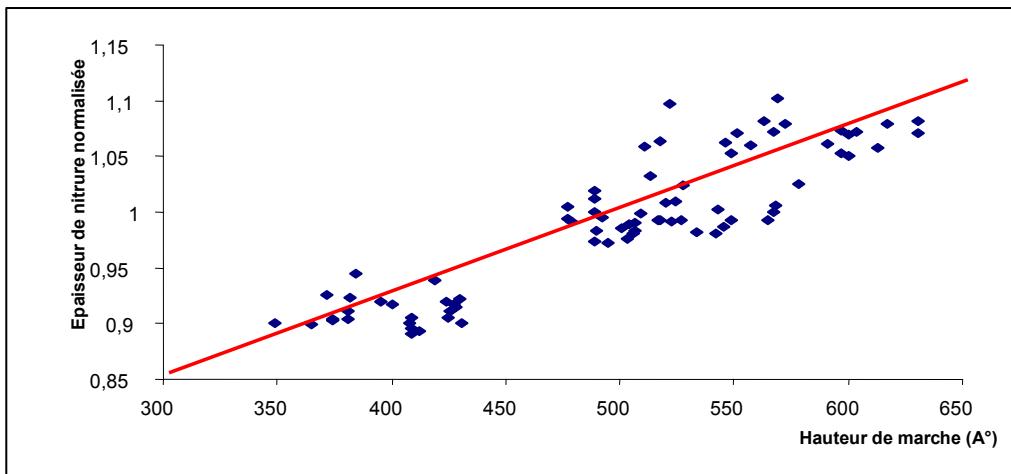


Figure 4.57- Variation de l'épaisseur de nitrule post-CMP (X_1) en fonction de la hauteur de marche dans le cas du direct STI

Dans le cas du Direct STI, la forte corrélation observée entre la hauteur de marche et les mesures du nitrule post-CMP peut être expliquée par le fait que dans le direct STI, la variabilité intra-plaque (uniformité de nitrule sur la plaque) est améliorée. Cette amélioration est obtenue grâce aux modifications apportées sur le procédé CMP en employant des abrasifs plus sélectifs comparés à ceux employés dans le procédé CMP standard. La figure 4.58 compare la distribution des mesures d'épaisseurs de nitrule post-CMP dans les deux cas du mode STI (direct et standard).

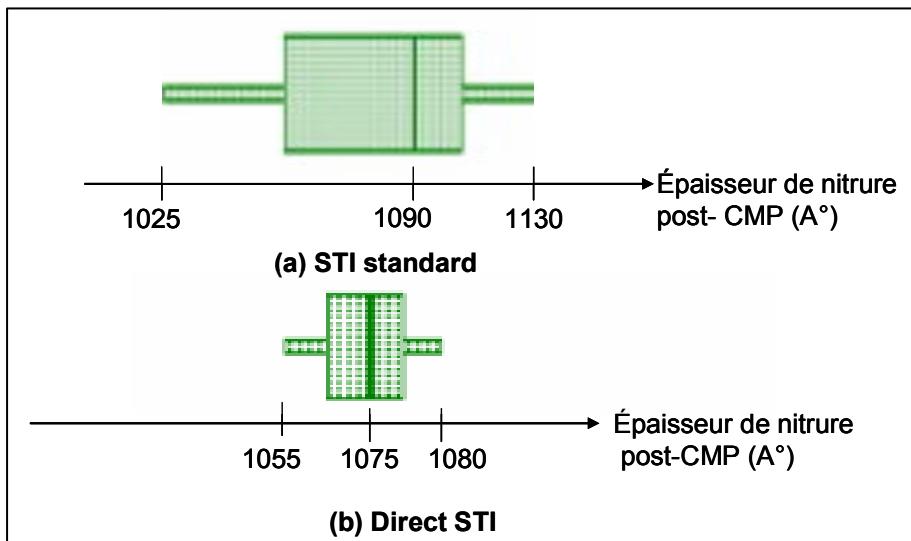


Figure 4.58- Comparaison de deux structures « box plot » représentant la distribution de l'épaisseur de nitrule post-CMP à travers la plaque dans le cas du STI standard et du Direct STI

La figure 4.58 montre bien que l'étendue est réduite de 105 Å dans le cas du STI standard à 35 Å dans le cas du Direct STI. Dans les prochains paragraphes, nous décrirons toute la démarche suivie pour réduire la dispersion de la hauteur de marche de plaque à plaque. Nous ne nous focaliserons que sur le Direct STI vu l'avantage qu'il présente.

Pour réduire la dispersion des mesures de la hauteur de marche de plaque à plaque, il faut donc réduire la dispersion des mesures post-CMP de plaque à plaque. La réalisation d'une boucle de type « feed-back » sur le procédé CMP en ajustant le temps de sur-polissage

du procédé CMP par exemple est difficile à réaliser car au niveau du CMP l'épaisseur de nitrate polie est fixe⁶⁵ (due à l'hyper-sélectivité des abrasifs employés lors du polissage). Une autre solution consiste à réaliser une boucle de régulation R2R type « feed forward ». Il s'agit d'agir sur le procédé intermédiaire à savoir la gravure humide en ajustant un paramètre du procédé de la gravure humide. Pour bien mener cette étude, il faut donc bien comprendre les mécanismes du procédé de la gravure humide pour choisir le bon paramètre ajustable.

4.4.3 Le procédé de gravure humide

4.4.3.1 Rappel sur les procédés chimiques utilisés dans la fabrication des circuits à semi-conducteurs

Les procédés chimiques interviennent tout au long de la fabrication des circuits. Les procédés chimiques sont utilisés pour nettoyer la plaque des différents contaminants⁶⁶ tels que les contaminations chimiques et les particulières présentes à la surface de la plaque. Ils sont également utilisés pour la désoxydation (enlèvement d'une couche d'oxyde), c'est le concept de la gravure par voie humide. Dans de très nombreuses filières technologiques, la gravure humide est utilisée majoritairement car elle est relativement simple à mettre en œuvre et permet de traiter en une seule opération. Des lots pouvant contenir jusqu'à 50 plaquettes peuvent être traités en une seule opération. La gravure par voie humide se fait par attaque chimique en solution aqueuse⁶⁷. Par exemple, l'oxyde de silicium est gravé par une solution partiellement diluée d'acide fluorhydrique (HF). Un exemple de bain de gravure est illustré sur la figure 4.59.

⁶⁵ Après le polissage de l'oxyde, le procédé CMP s'arrête après avoir poli environ 30 Å de nitrate.

⁶⁶ Nous pouvons distinguer 3 types de contaminants :

- Les particules : elles peuvent provenir du milieu externe de l'équipement comme la poussière, traces de l'eau, empreintes,...
- Les films sont généralement des couches différentes à la nature des matériaux constituant la plaque. Les films peuvent être organiques, inorganiques ou même métalliques. Par exemple, lors du procédé de la photolithographie, un film organique peut être déposé sur la plaque
- Les contaminants chimiques peuvent prendre la forme de particules ou de films. Les éléments chimiques peuvent être : sodium, potassium, etc... L'effet de ces contaminants est très critique sur le rendement électrique. Ils peuvent influencer le courant de fuite, modifier la tension de seuil.

⁶⁷ Bain contenant de l'eau.

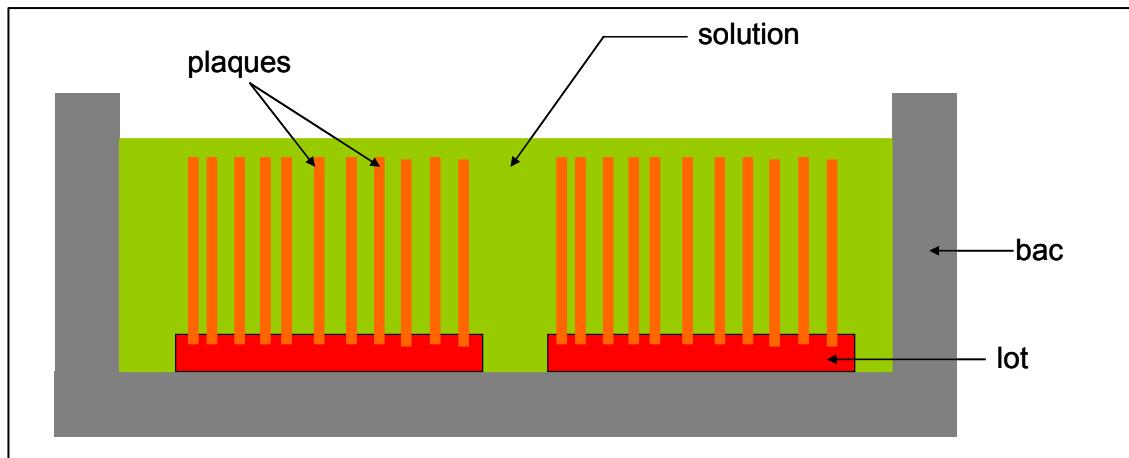


Figure 4.59- Gravure de deux lots par voie humide

4.4.3.2 Description de l'équipement de gravure humide

Pour l'opération de gravure par voie humide. Les bacs sont transférés d'un bain à un autre en mode automatique comme cela est illustré sur la figure 4.60. Le batch est trempé pendant un certain temps dans différents bains chimiques. Chaque bain est caractérisé par une certaine concentration et une certaine température.

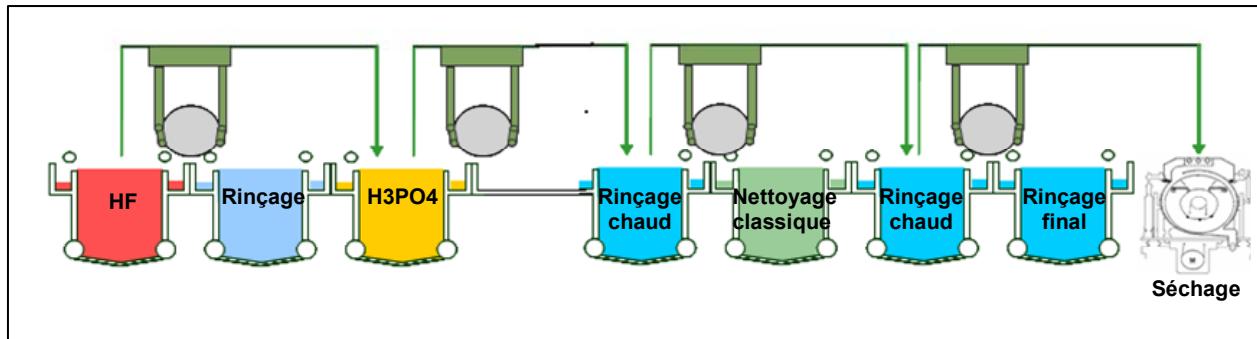


Figure 4.60- Les différents bains constituant la machine de gravure humide

Nous allons décrire les blocs (bains) principaux constituant l'équipement de gravure humide. Chaque bloc a une fonction particulière.

4.4.3.2.a Bain de gravure HF

Dans un premier temps, les plaquettes sont trempées dans un bain de HF partiellement dilué où une partie de l'oxyde est enlevé [FULLER 2006]. Cette opération de désoxydation est très importante pour définir la hauteur de marche car elle détermine l'épaisseur d'oxyde dans la tranchée.

4.4.3.2.b Bain de rinçage

Les résidus d’oxyde sont enlevés dans un bain d’eau déionisée⁶⁸. Si des impuretés ne sont pas nettoyées et arrivent à diffuser dans le silicium, dans les diélectriques ou les conducteurs, leur impact influence les paramètres électriques et donc le rendement électrique des circuits.

4.4.3.2.c Bain de gravure H₃PO₄

Les plaquettes sont trempées dans un bain d’acide ortophosphorique (H₃PO₄) afin d’enlever tout le film de nitrure de silicium (Si₃N₄). La gravure de nitrure par le H₃PO₄ est considérée comme sélective car le H₃PO₄ grave l’oxyde à basse vitesse de gravure. C’est l’un des avantages de ce procédé de gravure comparé à la gravure du nitrure par plasma.

4.4.3.2.d Bain de rinçage chaud

Après la gravure par H₃PO₄, il faut rincer abondamment [COWACHE 2000, BESSON 2001] et nettoyer les plaquettes. Le nettoyage avec des températures très élevées est plus efficace car les particules réagissent en grande quantité.

4.4.3.2.e Bain de nettoyage standard

Les bains d’acide phosphorique sont relativement visqueux et sont source de contamination particulière sur les plaques. Pour éliminer cette contamination, un bain de nettoyage succède au retrait nitrure. Le nettoyage se fait dans des bacs contenant des solutions diluées composées de NH₄OH/ H₂O. Deux solutions sont généralement utilisées, la première est une solution aqueuse, alcaline diluée avec l’eau (H₂O₂ : peroxyde d’hydrogène, NH₄OH : ammoniaque). Cette solution est très efficace pour l’enlèvement des films organiques. Ces derniers sont éliminés par les deux effets⁶⁹ de la solution (un effet de la forte oxydation de H₂O₂ et un effet des solvants NH₄OH⁷⁰). La deuxième solution est une solution aqueuse, acide diluée avec l’eau, H₂O₂ et du HCL (acide chlorhydrique). Cette solution est efficace pour l’enlèvement des métaux lourds.

4.4.3.2.f Bain de rinçage final

Le rinçage final se fait à l’eau dé ionisée. Cette étape permet de rincer la plaque et ainsi de s’assurer que toutes les traces d’acide sont complètement enlevées.

4.4.3.2.g Bain de séchage

Le séchage s’effectue dans de très grandes centrifugeuses ou par de l’air sec (1500 tours/minute). On utilise la centrifugation et l’azote pour sécher les plaquettes. Elles sont en rotation donc toute l’eau qui est sur les plaquettes est évacuée vers l’extérieur. Un flux d’azote chaud vient aider au séchage.

⁶⁸ Durant le rinçage à l’eau déionisée, la résistivité de l’eau est contrôlée afin de déterminer la quantité d’ions contaminants encore désorbés par les plaquettes.

⁶⁹ Les deux effets sont l’effet de l’oxydation de la surface par l’eau oxygénée où la croissance de l’oxyde se fait en consommant du silicium. Les particules se retrouvent ainsi piégées au milieu de l’oxyde. L’autre effet est l’attaque de l’oxyde par l’ammoniaque où les particules, prises dans l’oxyde, sont éliminées.

⁷⁰ Le NH₄OH est efficace pour le retrait des particules appartenant aux métaux I et II tels que le cuivre, l’argent, le cobalt et le nickel. Le NH₄OH permet de neutraliser les résiduels d’acide HF adsorbés sur la surface de la plaque.

4.4.4 Analyses multivariées

Nous venons de voir les étapes du procédé de gravure par voie humide. Chaque bain est caractérisé par une certaine concentration, une certaine température et un certain temps de procédé. Un exemple de fichier de données des paramètres du procédé est donné dans le tableau 4.9.

Etape	Temps de procédé effectif (secondes)	Temps de procédé demandé (secondes)	Température (°C)	Age du bain (minutes)	Ordre de passage (U.A)	Résistivité (MΩ.cm)	Pourcentage de la concentration HF (%)	Pourcentage de la concentration NH4OH (%)	Pourcentage de la concentration H2O2 (%)
(1) Gravure HF	163	156	21	416	10		3.02		
(2) Rinçage	3240	480			0	4.9			
(3) Gravure H ₃ PO ₄	3008	3000	160.3	667	7				15.88
(4) Rinçage chaud	493	480			0				
(5) Nettoyage classique	308	300	65.8	369	9			1.80	5.21
(6) Rinçage chaud	428	420			0	4.4			
(7) Rinçage final	488	480			0	17.8			
(8) Séchage	300	300			0				

Tableau 4.9- Exemple de conditions de procédé.

Les différentes colonnes du tableau résument les différents paramètres:

- l'étape de l'opération.
- le temps du procédé effectif (c'est le temps total durant lequel la plaque reste trempée dans le bain) [AKINWANDE 2003].
- le temps de procédé demandé. Pour des raisons d'automation, le temps de procédé demandé peut être différent du temps de procédé effectif.
- la température du bain chimique [WITVROUWA 2000].
- l'âge du bain. Généralement, la sélectivité par rapport à l'oxyde n'est pas la même au début du remplissage du bain et en fin de vie de celui-ci.
- l'ordre de passage des bacs dans le bain, certains produits (lots) sensibles à la sélectivité ne sont pas traités en début de vie d'un bain de H₃PO₄ par exemple mais après un certain temps. Ce dernier correspond au temps nécessaire pour saturer le bain [CALVIER 2005].
- la résistivité de la solution chimique du bain.
- la concentration de HF
- la concentration de NH₄OH
- la concentration de H₂O₂

4.4.4.1 Les analyses sous Simca-P

Des analyses multivariées ont été effectuées sur un ensemble de données des paramètres du procédé en vue de déterminer les paramètres qui influencent la variation de la hauteur de marche. Les résultats des analyses de la méthode des moindres carrés (PLS) sont donnés dans la figure 4.61.

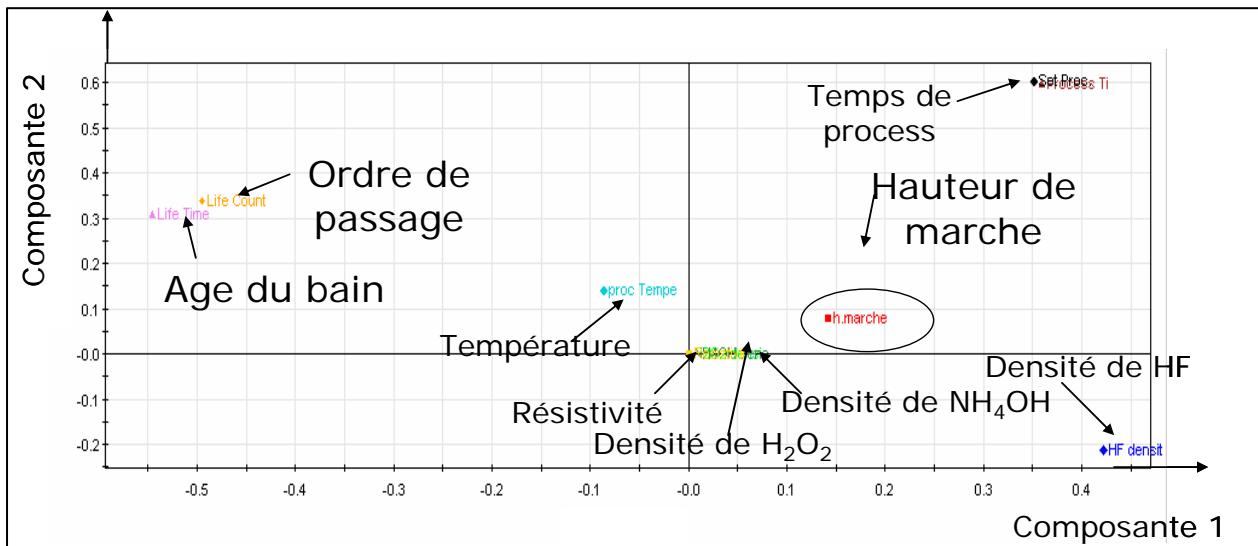


Figure 4.61- Résultats de l'analyse PLS

Comme nous l'avons déjà expliqué dans le chapitre I (partie III.2.2.2.), la méthode des moindres carrées (PLS) est une méthode qui permet d'estimer les coefficients d'un polynôme d'ajustement entre les variables à expliquer Y_k (dans ce cas la hauteur de marche) et un ensemble de variables explicatives X_i (paramètres du procédé de la gravure humide). Les variables présentes dans le graphe sont pondérées par des coefficients. La valeur des coefficients est comprise entre 0 et 1, si la valeur des coefficients est proche de 0 (c'est le cas de la température, la résistivité, la densité de NH_4OH et de H_2O_2), ces variables sont peu influentes sur la hauteur de marche. Par contre, la concentration de HF et les temps de procédé ainsi que l'âge du bain sont les variables influentes sur la hauteur de marche. La température, la résistivité, la densité de NH_4OH et de H_2O_2 ne seront pas retenus dans la partie expérimentale.

4.4.4.2 Choix des paramètres

Les résultats de la méthode PLS ont montré que certains paramètres du procédé sont très influents sur la hauteur de marche. Malgré cela, certains paramètres ne seront pas choisis comme paramètres ajustables pour plusieurs raisons :

- la concentration de HF est rigoureusement contrôlée (les limites de contrôle de la concentration de HF sont très serrées) donc ce paramètre ne peut être choisi comme paramètre ajustable.

- l'âge du bain ainsi que l'ordre de passage des lots ne seront pas pris comme paramètres ajustables, car sur les produits étudiés, les lots sont procédés dans le bain de H_3PO_4 une fois que celui ci est saturé.

Le temps de procédé est le seul paramètre restant ajustable, donc c'est ce paramètre qui sera considéré.

4.4.5 Modélisation du procédé

4.4.5.1 Le dépôt de nitrure

Nous avons cité dans le chapitre 2 les opérations fondamentales pour la réalisation du STI. Les deux premières opérations consistent à faire croître un oxyde piédestal et ensuite un film de nitrure. L'épaisseur de nitrure dans le cas du STI standard est de 1600Å. On peut disposer facilement de différentes épaisseurs de nitrure (1100Å à 1300Å) pendant le procédé CMP en modifiant le temps de polissage. Dans le cas du direct STI, l'épaisseur de nitrure de départ est de 1200 Å. Comme on utilise des abrasifs sélectifs au niveau du procédé CMP, l'épaisseur du film de nitrure reste presque inchangée après le procédé CMP. Dans ce cas, il n'est pas possible d'avoir différents films de nitrure (1100Å à 1300Å) au niveau du procédé CMP. Deux alternatives se présentent :

- soit le dépôt de différents films (des plaques avec 1100Å, 1200Å et 1300 Å), cela nécessite l'utilisation d'un four LPCVD en mode batch (procédant généralement 6 lots ou bien 150 plaques) pendant un certain temps (quelques heures). Cette opération est très coûteuse et pénalise la production, donc cette solution est rejetée.
- la seconde solution consiste à graver dans un bain de H_3PO_4 les 13 plaques ayant des films de nitrure avec des épaisseurs de 1600 Å. Cette solution est moins coûteuse par contre elle requiert une grande vigilance sur le contrôle de la vitesse d'attaque du H_3PO_4 . Nous avons vérifié l'uniformité de nitrure sur la plaque après la gravure de H_3PO_4 sur 6 plaques par une analyse cartographique. Le choix des 6 plaques est arbitraire et permet juste de voir la reproductibilité. Cette analyse consiste à comparer les profils de mesure de nitrure avant et après gravure. Les résultats de cette analyse sont illustrés dans le tableau 4.10.

Plaques	Moyenne de l'épaisseur de nitrure avant gravure (Å)	Etendue de l'épaisseur de nitrure avant gravure(Å)	Moyenne de l'épaisseur de nitrure après gravure (Å)	Etendue de l'épaisseur de nitrure après gravure (Å)
1	1622	63	1198	63
2	1616	40	1183	35
3	1574	74	1166	72
4	1601	48	1260	47
5	1610	58	1054	51
6	1610	51	1204	46

Tableau 4.10- Mesures avant et après gravure H_3PO_4

Les mesures de l'étendue avant et après gravure H_3PO_4 montrent que l'étendue de l'épaisseur reste quasi identique. Ceci, peut s'expliquer par le caractère de la gravure humide (même vitesse d'attaque en tout point de la plaque). Ce résultat est

confirmé par les travaux d'analyses des profils d'épaisseurs comme cela est illustré sur la figure 4.62.

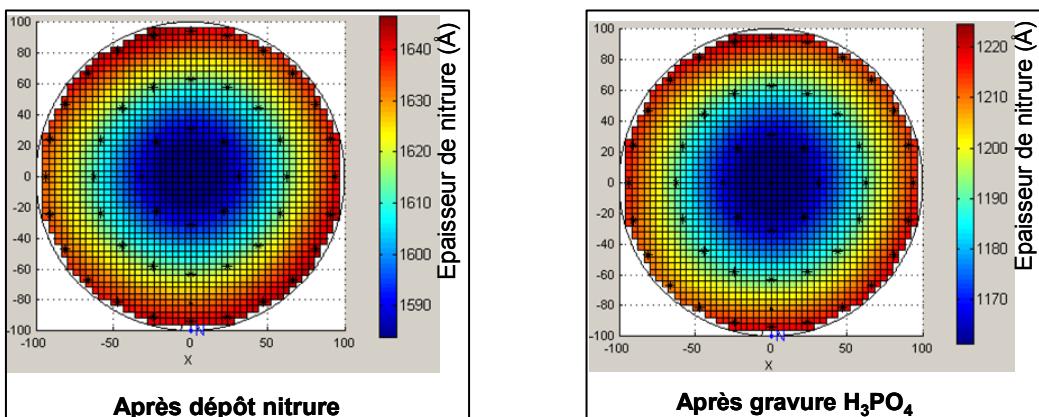


Figure 4.62- Comparaison des profils : a)après CMP b) après gravure humide

La figure 4.62 illustre les profils de niture avant et après gravure H₃PO₄. Les deux profils obtenus sont le résultat de la reconstruction de la surface (réalisée par le logiciel Wafer-Fit) à partir des mesures effectués en 49 points. Le profil de la plaque après dépôt niture correspond parfaitement aux profils de dépôts obtenus par LPCVD avec une épaisseur maximale sur les bords et une épaisseur minimale au niveau du centre. Ce profil est expliqué généralement par la répartition de la température sur la plaque. Après gravure, ce profil reste presque le même comme on peut le constater sur la figure 4.62. La gravure par H₃PO₄ est donc un très bon moyen pour réaliser des épaisseurs différentes de niture pour le plan d'expériences souhaité.

4.4.5.2 Le plan d'expériences

Le plan d'expériences conçu est de type central composite à faces centrées. Les épaisseurs de niture varient de 1100 à 1300 Å, sachant que la valeur cible est 1200 Å. Les temps de gravure varient également de 175 à 195 secondes de sorte à avoir des hauteurs de marche allant de 400 à 600 Å. Le type de plan d'expériences choisi est un « central composite à faces centrées ». Vu ce type de plan d'expériences et le nombre des variables étudiées (2), le nombre total d'expériences est donc 13 (dont 5 expériences pour tester la répétitivité de la réponse). Nous estimons que 3 expériences sont largement suffisantes pour tester la répétitivité. Les deux autres plaques ont été exploitées pour étudier deux cas extrêmes (hauteur de marche très petite et très grande).

Plaques	Epaisseur de nitrure (Å)	Temps du procédé gravure humide [secondes]
1	1100	175
2	1100	195
3	1300	175
4	1300	195
5	1100	185
6	1300	185
7	1200	175
8	1200	195
9	1200	185
10	1200	185
11	1200	185
12	1200	80
13	1200	280

Tableau 4.11- Conditions expérimentales visées pour le plan d'expériences.

4.4.5.3 Résultats

Une fois que les variables d'entrées du procédé sont définies, un plan d'expériences s'impose pour trouver un modèle qui met en évidence la relation entre la hauteur de marche et les différents paramètres d'entrée CMP et gravure humide. Comme cela est illustré dans le tableau 4.12, les paramètres mesurés après chaque étape du procédé sont :

- au niveau du CMP : mesure du film de nitrure et d'oxyde dans la tranchée
- au niveau de la gravure humide : mesure de l'épaisseur d'oxyde dans la tranchée
- mesure de la hauteur de marche

Le tableau 4.12 résume les différentes mesures à chaque étape pour toutes les plaques.

Plaques	Mesures post-CMP		Mesures après gravure humide			Mesures hauteur de marche et oxyde résiduel (Å)	
	Epaisseur d'oxyde dans la tranchée (Å)	Epaisseur de nitrule (Å)	Epaisseur d'oxyde dans la tranchée (Å)	Epaisseur d'oxyde enlevée dans la tranchée (Å)	Vitesse de Gravure humide [attaque HF] (Å/secondes)	hauteur de marche (Å)	Epaisseur D'oxyde résiduelle Dans la tranchée (Å)
1	4503	977	4018	485	2,77	415	3503
2	4605	978	4049	556	2,85	378	3571
3	4745	1144	4261	484	2,76	612	3549
4	4716	1146	4161	555	2,84	542	3519
5	4597	971	4085	512	2,76	413	3572
6	4746	1145	4222	524	2,83	579	3543
7	4752	1064	4267	485	2,62	555	3612
8	4728	1075	4172	556	2,85	488	3584
9	4744	1064	4217	527	2,84	527	3590
10	4699	1066	4173	526	2,84	512	3561
11	4653	1054	4129	524	2,83	474	3555
12	4744	1064	4490	254	2,82	791	3599
13	4734	1057	3923	811	2,83	163	3660

Tableau 4.12- Résultats expérimentaux

Nous pouvons constater dans le tableau 4.12 que la vitesse de gravure HF est quasi constante.

Le modèle de la hauteur de marche obtenu est donné par la relation suivante :

$$Y = a_0 + a_1 X_1 + a_2 t + a_{12} X_1 t + a_{11} X_1^2 + a_{22} t^2 \quad (4.31)$$

Où : Y est la hauteur de marche

X₁ est l'épaisseur de nitrule post- CMP et t est le temps de procédé de la gravure HF.

Les résultats des analyses statistiques effectuées à l'aide du logiciel SAS sont illustrés dans le tableau 4.13

Effet	Valeur de P-value	Remarques
X ₁	0,0234	Significatif
t	0,5324	Rejeté car P-value >0,05
X ₁ *X ₁	0,0543	Rejeté car P-value >0,05
X ₁ *t	0,2551	Rejeté car P-value >0,05
t*t	0,6730	Rejeté car P-value >0,05

Tableau 4. 13- Valeurs de (P-value) obtenus sous SAS

Les résultats du tableau 4. 13 montrent que X₁ est le seul paramètre significatif car la valeur de P-value <0.05. Les coefficients du modèle ont été calculés à l'aide du logiciel SAS. Le modèle donnant la hauteur de marche en fonction de l'épaisseur de nitrule post-CMP est :

$Y = 1.033$	$X_1 = 598$
-------------	-------------

(4.32)

4.4.5.4 Simulation avec le modèle du procédé

Des simulations ont été faites sur un historique de données (170 lots) en vue de simuler des hauteurs de marche. Les résultats des simulations ont été comparés avec les résultats expérimentaux comme cela est illustré sur la figure 4.63.

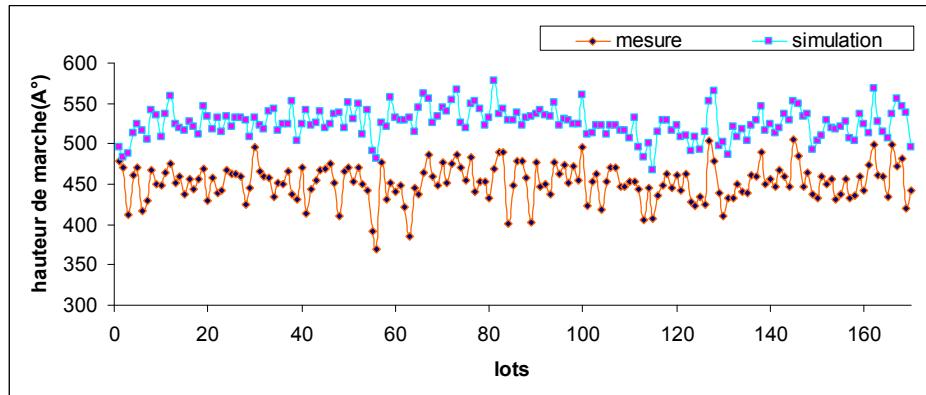


Figure 4.63- Comparaison entre les valeurs de mesure et de simulation

Nous observons sur la figure 4.63 que les résultats de simulations ne sont pas assez prédictifs. De ce fait, ce modèle ne peut pas être employé pour la boucle « feed-forward ». Le fait que ce modèle manque de prédiction peut s’expliquer par le fait que ce modèle ne tient pas en compte des variations des paramètres équipement CMP et donc au moins une autre variable manque dans le modèle. L’hypothèse émise est la variation de la vitesse d’enlèvement dans le procédé CMP causée par des dérives⁷¹ généralement observées au cours de la durée de vie des plateaux et des têtes. Pour vérifier la validité de cette hypothèse, nous avons effectué des analyses sur un ensemble de données regroupant des paramètres équipements CMP, des mesures pré-CMP et des mesures post-CMP.

Nous nous sommes intéressés à l’épaisseur de nitrure qu’on appellera par la suite Δ nitrure (Δ nitrure = épaisseur de nitrure préCMP – épaisseur de nitrure postCMP). Cette grandeur est illustrée sur la figure 4.64. Nous distinguons également sur la même figure l’épaisseur d’oxyde poli dans la tranchée symbolisée par Δ oxyde :

$$\Delta\text{oxyde} = \text{épaisseur d'oxyde préCMP} - \text{épaisseur d'oxyde postCMP} \quad (4.33)$$

⁷¹ Les dérives sont dues essentiellement à l’usure de certains éléments constituant l’équipement CMP par exemple les plateaux de polissage, les têtes de polissage et les tissus de polissage.

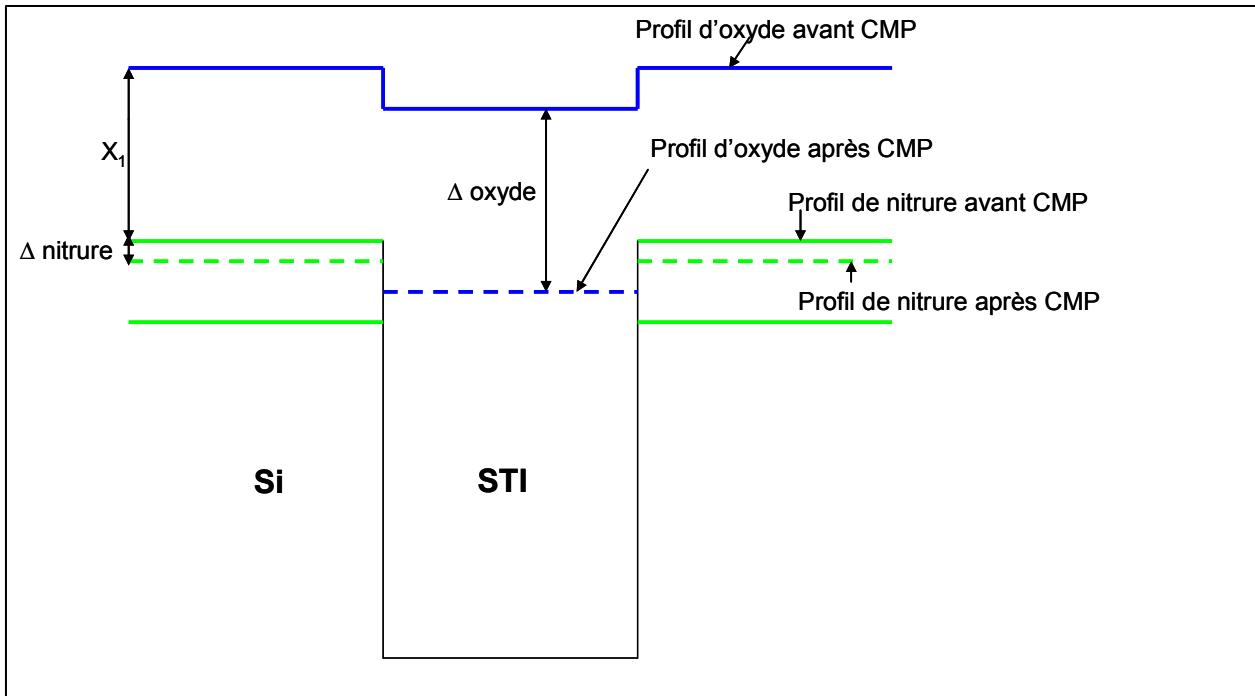


Figure 4.64- Illustration de l'épaisseur polie après le procédé CMP

4.4.5.5 Modélisation des procédés CMP - gravure humide

Nous allons rappeler brièvement les opérations élémentaires des deux procédés (CMP et gravure humide). Le temps total de polissage est la somme du temps de polissage d'oxyde (déposé sur le nitrate) et du temps de gravure d'une couche de nitrate comme cela est illustré sur la figure 4.65. Ce temps correspond au temps de « sur-polissage ». Pendant le temps de « sur-polissage », une petite couche de nitrate est uniquement enlevée. Ceci est dû à la haute sélectivité des abrasifs (vitesse d'enlèvement d'oxyde/ vitesse d'enlèvement de nitrate) et donc le point de fin de polissage d'oxyde est plus profond comparé à celui de nitrate. Le procédé suivant est la gravure humide où on grave une quantité d'oxyde à l'aide de l'acide fluorhydrique (HF) et ensuite le film de nitrate est éliminé par l'acide orthophosphorique. La hauteur de marche est mesurée après le retrait de nitrate et le retrait du film piédestal.

On définit :

- X_0 correspond à l'écart initial entre l'épaisseur d'oxyde dans la tranchée et l'épaisseur d'oxyde. Pour les structures de test dédiées pour la mesure d'oxyde par ellipsométrie et la hauteur de marche par AFM dont les dimensions sont grandes. La marche $X_0 = X_8 + X_4 + X_2$. Pour des structures dont les dimensions sont petites (largeur $< 1\mu\text{m}$), X_0 devient plus petite et par conséquent la tranchée est quasi pleine.
- X_1 : épaisseur d'oxyde déposée sur le nitrate, la valeur cible de X_1 est de 6300\AA
- X_2 : épaisseur initiale de film de nitrate (avant CMP), la valeur cible de X_2 est de 1200\AA
- X_3 : épaisseur finale de film de nitrate (après CMP)

- Δ_{nit} : épaisseur de film de nitre réellement polie (après CMP) [$\Delta_{\text{nit}}=X_2-X_3$], généralement $\Delta_{\text{nit}}<50 \text{ \AA}$
 - X_4 : épaisseur d'oxyde (SiO_2) piédestal, la valeur cible de X_4 est de 100 \AA
 - X_5 : épaisseur de film d'oxyde poli pendant le temps correspondant au temps de polissage de nitre
 - X_6 : épaisseur de film d'oxyde gravé après la gravure humide
 - X_7 correspond à une partie d'oxyde gravée dans la tranchée (X_7) lors de la gravure HF de l'oxyde piédestal (dans ce cas $X_4=X_7$).
 - X_8 : correspond à la profondeur de la tranchée (épaisseur de silicium enlevée après gravure sèche), la valeur cible de X_8 est de 3500 \AA
 - $t_{\text{grav.HF}}$ est le temps de gravure par HF dans la tranchée.
 - h_m est la mesure de la valeur de la hauteur de marche (mesurée après le retrait nitre et du film d'oxyde piédestal).

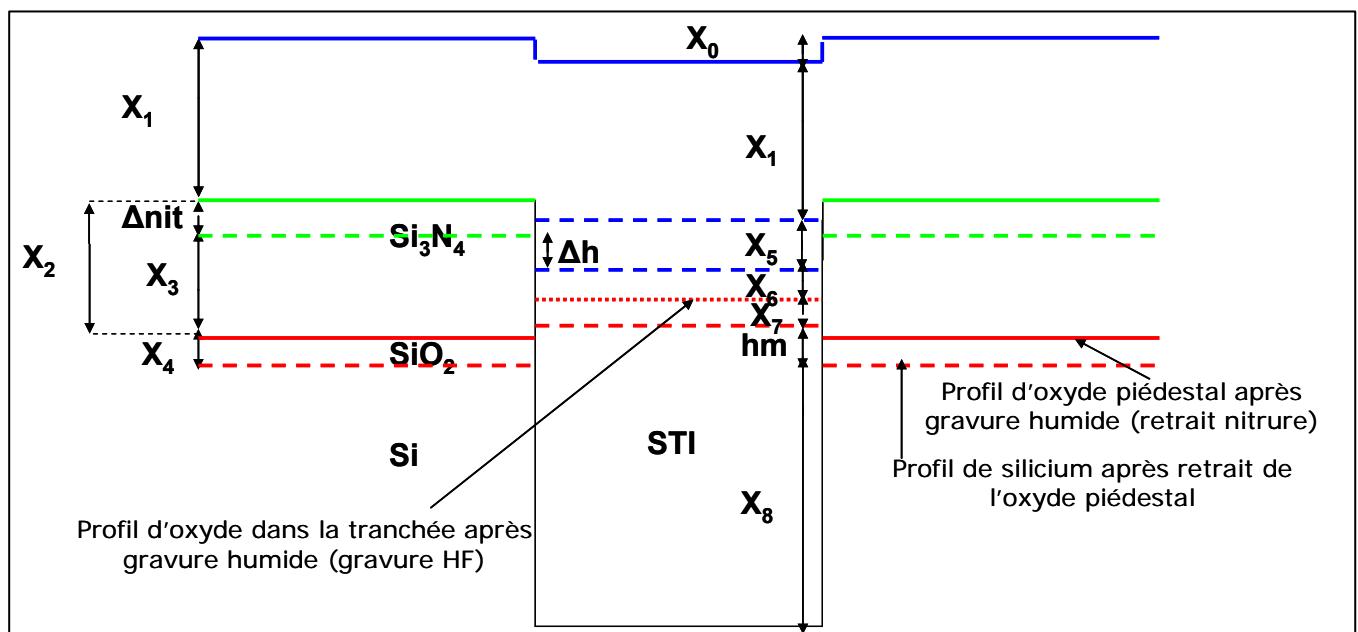


Figure 4.65- Représentation schématique du modèle.

$$X_3 + X_4 = \Delta h + X_6 + X_7 + hm \quad (4.34)$$

Comme : $X_4 = X_7$, alors $X_3 = \Delta h + X_6 + hm$

$$X_6 = t_{HF} * V_{gravureHF} = X_3 - \Delta h - hm \quad (4.35)$$

La vitesse de gravure HF est de 2,82 Å/sec. Le temps de gravure t_{HF} est donné par la relation suivante :

$$t_{HF} = \frac{X_3 - \Delta h - hm}{2,82} \text{ (secondes)} \quad (4.36)$$

Si la dispersion de X_2 est faible, la dispersion de la hauteur de marche n'est pas impactée par X_2 mais par la dispersion de l'épaisseur d'oxyde post-CMP ($X_6 + X_7 + hm + X_8$). Nous pouvons donc réduire cette dispersion qu'au niveau de la tranchée en ajustant X_6 et donc le temps de gravure.

D'après la figure 4.65 :

$$hm = (X_6 + X_7 + hm + X_8)_{mesuré} - (X_6 + X_7 + X_8) \quad (4.37)$$

Si la hauteur de marche visée est $hm_{visée}$, la formule (4.37) devient:

$$hm_{visée} = (X_6 + X_7 + hm + X_8)_{mesuré} - (X_6 + X_7 + X_8) \quad (4.38)$$

$$hm_{visée} = (X_6 + X_7 + hm + X_8)_{mesuré} - (V_{gravureHF} * t_{HF} + X_7 + X_8) \quad (4.39)$$

$$t_{HF} = \frac{(X_6 + X_7 + hm + X_8)_{mesuré} - (hm_{visée} + X_7 + X_8)}{V_{grav.HF}} \text{ (secondes)} \quad (4.40)$$

4.4.5.6 Simulations avec le nouveau modèle

Nous pouvons simuler n'importe quelle hauteur de marche à partir des conditions du procédé. Dans notre cas, nous faisons varier les valeurs de l'épaisseur de nitre post CMP (de 1020 Å à 1100 Å). Une simulation est jugée bonne lorsque les simulations ont été effectuées avec des conditions similaires à celles des conditions expérimentales. De ce fait, les résultats des simulations sont comparés avec les données des mesures ou résultats expérimentaux. Nous avons comparé les conditions de mesures des hauteurs de marche avec des valeurs de simulations. Les résultats de cette comparaison sont illustrés sur la figure 4.66.

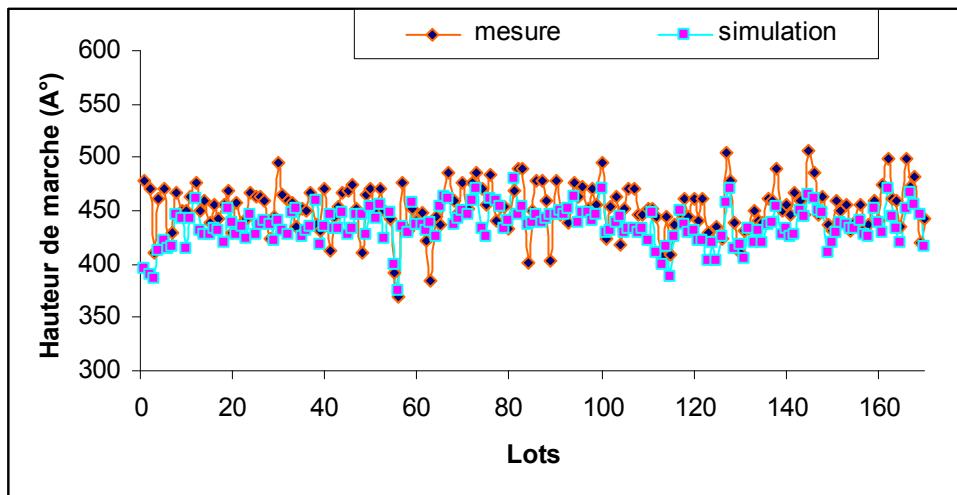


Figure 4.66- Comparaison entre les données mesurées et la simulation.

Nous observons que les résultats de simulations sont assez prédictifs car on arrive à simuler presque la même valeur trouvée en expérimental comme nous pouvons le constater sur la figure 4.66. En modifiant le temps de gravure humide, nous pouvons contrôler la valeur de la hauteur de marche et s'approcher de sa valeur cible. Comme nous l'avons déjà expliqué dans le chapitre 1 (partie 1.3.6), la notion C_{pk} « capabilité des procédés » est utilisée pour analyser la variabilité des procédés en regard des spécifications associées. Nous pouvons donc conclure qu'avec ce modèle nous pouvons améliorer le C_{pk} de la hauteur de marche.

Ce dernier modèle est assez prédictif contrairement au premier modèle proposé car il prend en compte les variations des données équipements CMP qui impactent directement la vitesse d'enlèvement. C'est ce modèle qui est proposé pour être déployé en production.

4.4.6 Premiers résultats de la pré- production

Avant de lancer ce modèle en production et malgré les bons résultats de simulation, une étape de tests en production ou bien la pré- production est nécessaire pour valider le modèle et vérifier le gain apporté par les boucles de régulation en test électrique. Au moment de la pré- production, seuls 6 lots ont été à notre disposition pour le test électrique, les autres lots ont été bloqués au niveau d'autres opérations de fabrication.

4.4.6.1 Mesures de hauteur de marche

Les lots sont mesurés au niveau des opérations pré- CMP, post- CMP et au niveau de la hauteur de marche.

LOTS	Temps de procédé de gravure HF (secondes)	Hauteur de marche visée(Å)	Hauteur de marche mesurée (Å)
1	188	450	468
2	202	450	463
3	202	450	468
4	202	450	451
5	202	450	450
6	204	450	462

Tableau 4.14- Les résultats des mesures sur les lots de la pré- production

Comme c'est indiqué sur le tableau 4.14, il n'y a que deux lots sur six qui ont été ajusté automatiquement. La correction sur les deux lots a été faite avec succès. En fait, sur le premier lot, au lieu d'avoir une hauteur de marche de 415 Å avec un temps standard (soit un écart de 35 Å par rapport à la valeur cible), avec la correction $t_{HF}=188$ secondes, la valeur de la hauteur de marche est de 468 Å donc plus proche de la cible. De la même façon, pour le sixième lot, au lieu d'avoir une hauteur de marche de 475 Å, avec le rajustement, la valeur de la hauteur de marche est de 462 Å.

Grâce au modèle proposé, nous pouvons prédire la hauteur de marche. Les temps de gravure sont ajustés automatiquement. Les résultats de la pré-production sont donc très encourageants.

4.4.6.2 Paramètres électriques

Dans le chapitre 3, on a démontré que le paramètre électrique à surveiller dans les boucles de régulation est la tension de seuil des transistors parasites (Vt). Les lots qui ont servi pour le test en pré- production ont continué tous les procédés de fabrication, en vue de les tester en tests électriques. Malheureusement, au moment de la rédaction, ces lots ne sont pas encore arrivés à cette étape. Nous nous contenterons des résultats de simulations. En se basant sur les résultats des corrélations déduites au chapitre 3. Nous pouvons prédire la valeur du Vt parasite.

Nous avons constaté que les valeurs de Vt pour chaque lot sont bien centrées sur la valeur cible. Néanmoins, il faut noter que ce nombre de lots est insuffisant pour expliquer le gain apporté par les boucles de régulation. Il faut donc au moins, une dizaine de lots et de mesures effectives pour bien expliquer cet effet.

4.4.7 Surveillance de la boucle de régulation

En vue de bien surveiller la boucle de régulation R2R, il est nécessaire de calibrer la boucle périodiquement, ceci pour tenir compte d'éventuelles perturbations. Si par exemple, un changement de procédé a lieu au niveau du module de la grille, ceci impactera d'une manière directe ou indirecte les paramètres électriques comme l'indicateur du STI :Vt . Des dispersions sur ce paramètre ont été observées. Ces dispersions peuvent être expliquées par l'effet produit.

4.4.7.1 Effet produit

Chaque technologie produite en FAB est caractérisée par ses propres limites SPC imposées (valeur cible, limites hautes et basses de contrôle des paramètres du procédé). Il s'avère que les produits appartenant à la même technologie ne sont pas tous sensibles aux paramètres d'entrée et de sortie des variables du procédé [KADOSH 2005].

La figure 4.67 montre un exemple de l'évolution de la tension de seuil des transistors parasites en fonction de l'épaisseur de STI. Il est très clair sur cet exemple que le paramètre est dépendant de la nature du produit.

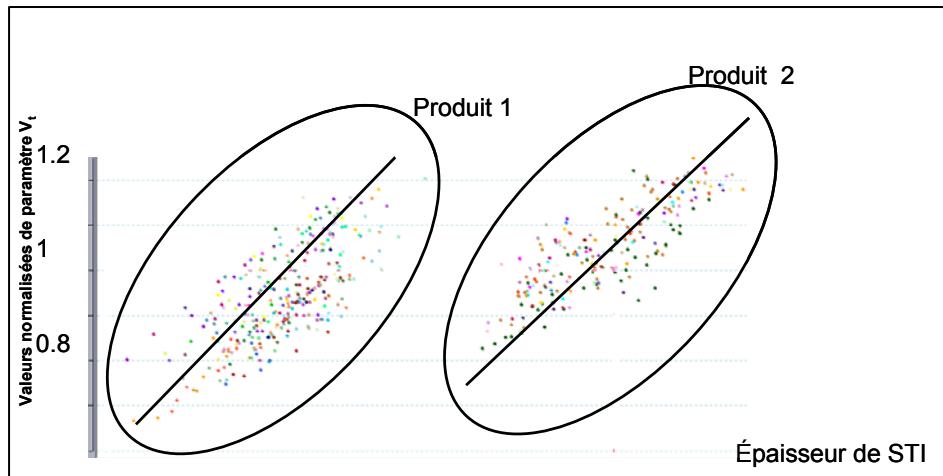


Figure 4.67- Evolution de la tension de seuil en fonction de l'épaisseur STI

4.4.7.2 Superviseur pour le contrôle

En plus de l'effet produit qui induit certaines formes de dérives, des perturbations peuvent avoir lieu et sont observées par la suite au niveau des paramètres électriques. Ces dérives peuvent avoir comme origine la variation des caractéristiques physiques des composants (concentration de dopants, propriétés de matériau, l'interface d'une couche...). Un superviseur pour les boucles de R2R a été conçu pour mieux contrôler les variations de certains paramètres électriques déjà identifiés comme indicateurs pour monitorer le STI. Les dérives induites par le procédé ou équipement peuvent être compensées à l'aide du nombre important des données des expériences et données de production qui sont fournies régulièrement pour le superviseur [CHIOU 2005, HARRISON 2004, GOOD 2003].

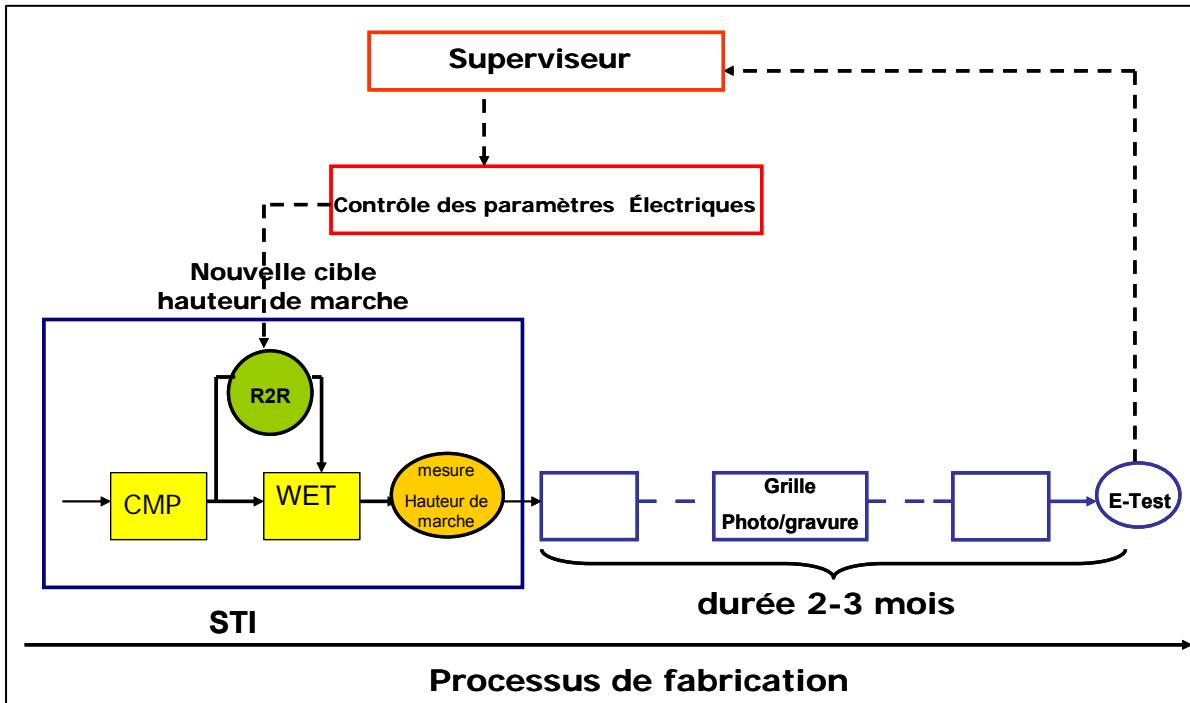


Figure 4.68- Schéma synoptique du superviseur

La figure 4.68 montre le schéma synoptique du superviseur. On distingue sur le graphe les mesures de métrologie après les opérations du procédé au niveau du module STI ainsi que les données de tests électriques (faits 2 à 3 mois après la réalisation du module STI). Ces données sont envoyées au superviseur. Le modèle prédictif compare les données de mesures ainsi que les valeurs prédites. Le superviseur recommande par la suite une autre valeur cible de hauteur de marche. Dans ce cas, par exemple, la boucle feed forward CMP- gravure humide [BELHARET 2007d], verra son modèle de procédé modifié pour être adapté à la nouvelle valeur cible de la hauteur de marche. En plus, le superviseur peut être employé pour plusieurs boucles de régulations R2R suivant la dépendance des paramètres électriques avec les paramètres géométriques du STI. Par exemple, si on observe une variation des valeurs d'un autre paramètre électrique qui dépend fortement de la profondeur de tranchée, dans ce cas, le superviseur recommande une nouvelle valeur cible de profondeur de tranchée.

4.4.8 Conclusion

La dispersion de la hauteur de marche est due en partie à la dispersion de l'épaisseur du nitrate post-CMP. La solution que nous avons proposée est la réalisation d'une boucle R2R de type « feed forward ». Cette dernière permet de réduire la dispersion de la hauteur de marche de plaque à plaque en ajustant le temps de gravure HF. Le modèle de la boucle tient en compte les mesures de nitrate post-CMP ainsi que de l'épaisseur de nitrate polie. Ce dernier paramètre est un très bon indicateur de la variation de la vitesse d'enlèvement. Nos travaux de recherches ont confirmé l'importance de ce paramètre à travers les corrélations obtenues. Les résultats de simulations avec ce modèle de procédé ont été satisfaisants. Nous avons testé ce modèle en pré-production sur six lots. Les résultats de la pré-production sont très bons, car toutes les hauteurs de marche mesurées ont été toutes proches de la valeur cible. Pour calibrer cette boucle et les boucles proposées dans ce chapitre, nous avons proposé un superviseur qui tient compte des modifications faites sur les procédés ou bien des dérives constatées sur les équipements. En comparant les résultats des tests électriques et les résultats issus des modèles prédictifs établis, une valeur cible de la mesure au niveau d'un procédé peut être modifiée.

Références bibliographiques

- [ABDELGADIR 2006] Mahjoub Ali Abdelgadir, “HDP-CVD STI Oxide Process With *in situ* Postdeposition Laterally Enhanced Sputter Etchback for the Reduction of Pattern-Dependent Film Topography in Deep Submicron Technologies”, *IEEE transactions on semiconductor manufacturing*, Volume: 19, issue 1, pp. 130- 137, February 2006.
- [ADIVIKOLANU 1998] Adivikolanu, S. and E.Zafirou (1998). “Robust run to run control manufacturing: An internal model control approach”. *Proceedings of the American Control Conference*, pp. 3687-3691.
- [AKINWANDE 2003] Tayo Akinwande, Cours de AKINWANDE “Wet Etching-Etching and Pattern Transfer, *Fall Term 2003 Lecture 13*”, MIT, 2003.
- [ALEGRET 2005] Alegret, C., F. Pasqualini, R. Fernandez (2005). “Machine learning for metrology applications in front end manufacturing: Adaptive surface reconstruction and abnormal profile detection”. *2nd ISMI Symposium on Manufacturing Effectiveness* Austin.
- [BABAYAN 2001] S.E. Babayan, J.Y. Jeong, A. Schutze, V.J. Tu, Maryam Moravej, G.S. Selwyn and R.F. Hicks, “Deposition of silicon dioxide films with a non-equilibrium atmospheric-pressure plasma jet”, *Plasma Sources Sci. Technol.* 10 (2001) 573–578
- [BELHARET 2005] D.BELHARET, P.COLLOT, J.PINATON, L.BUCELLE, J.CHOLVY, E.DEHAN, “Correlation beetwen Fault Detection results and Process Mechanisme: Application to HDP-CVD Process”, *AEC/APC 2005 Asian conference*, Hsin Chu (Taiwan), decembre 1&2, 2005
- [BELHARET 2006 a] D.Belharet, P.Collot, J.Pinaton, “Run TO RUN Process Model for HDP-CVD”, *AEC/APC European Conference*, Aix en Provence, France, March 29-31 2006
- [BELHARET 2006 b] D. Belharet, P.Collot, P. Fornara,J.Pinaton, J.Cholvy, A.Roussy, “OXIDE HDP-CVD MODELING FOR SHALLOW TRENCH ISOLATION”, *Advanced process control for semiconductor manufacturing*, Singapoure 4-5 Décembre 2006.
- [BELHARET 2006c] Djaffar BELHARET, Pascal FORNARA, Jérôme Cires, Philippe COLLOT, Jacques PINATON, Agnès ROUSSY, “An Improved Process for Direct Shallow Trench Isolation (STI) Wet Etch Using Post-CMP Metrology Information”, *AEC/APC Symposium XVIII*, Colorado 2006.
- [BELHARET 2007d] Djaffar Belharet, Pascal Fornara, Philippe Collot, Agnès Roussy, Jacques Pinaton, Loic Charrier, “Model Predictive Control of Critical Electrical Parameters Using CMOS Isolation Module Information in the Semiconductor Manufacturing”, *IEEE APC 2007-Applications for industry*, Vancouver, 14-16 May 2007.
- [BELHARET 2007e] Djaffar BELHARET, Agnès ROUSSY, Pascal FORNARA, Philippe COLLOT, Jacques PINATON, “Yield enhancement on the wafer edge by adjusting the retainer pressure during the CMP process”, *The 10th Technical and Scientific Meeting of ARCSIS*. 6-7 Décembre, Fuveau (France) 2007.
- [BESSON 2001]P.Besson, C.Cowache , J.M.Fabbri, F.Tardif and A.Beverina, “Low Consumption Front End Of the Line Cleaning : LC-FEOL”, *Solid state phenomena* Vol 76-77 (2001), pp.199-202
- [BHUSHAN 1995] M. Bhushan, R. Rouse, and J. E. Lukens, “Chemical-Mechanical Polishing in Semidirect Contact Mode” , *J. Electrochem*, Vol. 142, No.11, November 1995
- [BONINIG 1996] Duane Boning, William Moyne, Taber Smith, James Moyne, Roland Telfeyan, Arnon Hurwitz, Scott Shellman, “Run by Run Control of Chemical-Mechanical Polishing”, *IEEE*, Vol 19, issu 4, pages 307-314, 1996.
- [BORBA 1995] M. Borba, T. Myers, M. Stell, D. Scherber and M. Fury, “Effect of Oxide CMP Slurry Properties on Polishing Performance”, *DUMIC*, 1995.
- [BURKE 1991] P. A. Burke, “Semi-Empirical Modeling of SiO₂ Chemical-Mechanical Polishing Planarisation”, *VMIC*, pp. 379-384, 1991.

- [CALVIER 2005] Jean-Jerome Calvier, Virginie Piard, Colin Henderson, Philippe Garnier, Gautier Horellou, Marcel Geomini, “Improve hot phosphoric process robustness in 300mm nitride strip wet bench, Semiconductor Manufacturing”, *ISSM 2005, IEEE International Symposium on Volume, Issue , 13-15 Sept. 2005* Page(s): 438 - 441
- [CHEN 2001] K.W. Chen, Y.L.Wang, L.Chang, CW Liu, Y.K. Lin, T.C. Wang, and S.T. Chang, K.Y. Lo, “Novel Strategies of FSG - CMP for Within-Wafer Uniformity Improvement and Wafer Edge Yield Enhancement Beyond 0.18 Micro Technologies”, *IEEE*, 2001
- [CHIOU 2005] H.W. Chiou, and T. Wang, Supervisal “Run to Run Control of Device Electrical Characteristics”, *AEC/APC Asian symposium*, Hsin Shu (Taiwan), 1-2 Dec. 2005.
- [CHOI 2006] Jihong Choi, Shantanu Tripathi, David A. Hansen, David Dornfeld “Chip Scale Prediction of Nitride Erosion in High Selectivity STI CMP”, rapport technique, Université de Berkeley, 2006
- [COTE 1995] D. R. Cote, S. V. Nguyen,W. J. Cote, S. L. Pennington, A. K. Stamper, D. V. Podlesnik, “low temperature chemical vapour deposition processes and dielectrics for microelectronic circuit manufacturing” , *IBM J. RES. DEVELOP.* VOL. 39 NO. 4 JULY 1995.
- [COOK 1990] L. M. Cook, “Chemical Processes in Glass Polishing”, *Journal of Non-Crystalline Solids*, 120, pp. 152-171, 1990.
- [COPETTA 1998] J. Copetta, C. Rogers, A. Philipossian and F. B. Kaufman, L. C. Racz, “Pad Effects on Slurry Transport Beneath a Wafer during Polishing”, *CMP-MIC*, Feb. 1998.
- [COWACHE 2000] C. Cowache, P. Boelen, I. Kashkoush, P. Besson, F. Tardif, “Evaluation of advanced pre-gate cleanings”, *Proceedings of the Sixth International Symposium, Electrochemical Society Proceedings*, vol. 99-36, pp. 59-68 (2000).
- [EL CHEMALI 2000] C El Chemali et al., "Multizone Uniformity Control of a Chemical Mechanical Polishing Process Utilizing a Pre- and Postmeasurement Strategy," *Journal of Vacuum Science and Technology A* 18, no. 4 (July/August 2000): 1287–1296.
- [HOYER 2004] Gunther Hoyer, “Matching of HDP-Chambers using Golden Tool Methodology & VDS Measurement Tool”, *IEEE advanced semiconductor manufacturing conference and workshop*, pp.199-204, 4-6 mai 2004. Boston, USA
- [HUANG 1995] Vincent S.K. Huang, Cuong T. Nguyen, Alice B.Y. Chan, Curtis C. Ling, and S . Simon Wong, “NON-RADIAL NON-UNIFORMITY IN CHEMO-MECHANICAL POLISHING”, *IEEE*, 1995
- [FULLER 2006] Lynn Fuller, ”Wet Etch for Microelectronics”, http://www.rit.edu/~lf3eee/wet_etch.pdf
- [GAN 2000] Terence Gan, “Modeling of Chemical Mechanical Polishing for Shallow Trench Isolation”, *Thèse soutenue au Laboratoire MIT*, 2000 .
- [GOUGOUSI 2000] T. Gougousi, R. Sreenivasan, Y. Xu, L. Henn-Lecordier, G. W. Rubloff, J. N. Kidder, Jr., and E. Zafiriou, “In-situ Sensing Using Mass Spectrometry and its Use for Run-To-Run Control on a W-CVD Cluster Tool”, *Characterization and Metrology for ULSI Technology: 2000 International Conference*, 26-29 Juin2000, Maryland, USA.
- [GOOD 2003] R. Good, “The Control of Electrical Properties of Semiconductors”, *Texas-Wisconsin Modeling and Control Consortium Spring Conference*, 10-11 Feb. 2003.
- [GOWER-HALL 2002] Aaron E. Gower-Hall, Duane S. Boning, Peter Rosenthal, and Ann Waldhauer, “Model-Based Uniformity Control for Epitaxial Silicon Deposition”, *IEEE transactions on semiconductor manufacturing*, Volume: 15, issue 3, pp. 295- 309, august 2002.
- [GOWER 2002] Aaron E. Gower-Hall, Duane S. Boning, “Model-Based Uniformity Control for Epitaxial Silicon Deposition”, *IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 15, NO. 3, AUGUST 2002*.

- [GOUGOUSI 2000] T. Gougousi, R. Sreenivasan, Y. Xu, L. Henn-Lecordier, G. W. Rubloff, J. N. Kidder, Jr., and E. Zafiriou, “In-situ Sensing Using Mass Spectrometry and its Use for Run-To-Run Control on a W-CVD Cluster Tool, CP550”, *Characterization and Metrology for ULSI Technology: 2000 International Conference*, 2000
- [GUO 1993] Ruey-Shan Guo and Emanuel Sachs, “Modeling, Optimization and Control of Spatial Uniformity in Manufacturing Processes”, *IEEE transactions on semiconductor manufacturing*, Volume: 6, issue 1, pp. 41- 57, fevrier 1993.
- [GUTHRIE 1996] Guthrie, W and Cheng, T. “Method and apparatus for using a retaining ring to control the edge effect”. *Applied Materials, Inc., US Patent No.5,795,215*: 1996.
- [HARRISON 2004] C.A. Harrison, R. Good, D. Kadosh, and S.J. Qin, “A multi-step supervisory control strategy for semiconductor device manufacturing”, *Decision and Control, 2004. CDC. 43rd IEEE Conference*, vol 4, 14-17, pp. 4237 – 4242, Dec. 2004
- [KADOSH 2005] D. Kadosh, K. Chamness, and R. Good, “Implementation of Optimized MPC-Based Supervisory Control In a Semiconductor Manufacturing Environment”, *6th European AEC/APC conference*, Dublin (Ireland), 6-8 Apr. 2005.
- [JACKSON 1995] Jackson, P and Schultz, S. “Pneumatic polishing head for CMP apparatus. Integrated Process Equipment Corp”, *US Patent No. 5,643,061*: 1995.
- [JOUBERT 2004] O. Joubert, G.Cunge, B.Pelissier, L.Vallier, M.kogelschatz and E. Pargon, “Monitoring of chamber walls coating deposited during plasma processes: Application to silicon gate etch processes”, *Journal of vaccum & sciences technology*, Vol.22(3), may 2004
- [KIM 2006] Nam-Hoon Kim , Min-Ho Choi , Sang-Yong Kim , Eui-Goo Chang, “Design of experiment (DOE) method considering interaction effect of process parameters for optimization of copper chemical mechanical polishing (CMP) process”, *Microelectronic Engineering* 83 (2006) 506–512
- [KIMURA 1998] Kimura, N and Yasuda, H. “Polishing apparatus”. *Ebara Corp., European Patent No. 0870576-A2*: 1998.
- [LENSING 2001] Kevin R. Lensing, Richard J. Markle, Broc Stirton and Michael A. Laughery, “Shallow Trench Isolation Scatterometry Metrology in a High Volume Fab”, *IEEE international Semiconductor manufacturing symposium* , pp.195-198, 2001
- [LAI 2001] Jiun-Yu Lai, “MECHANICS, MECHANISMS, AND MODELING OF THE CHEMICAL MECHANICAL POLISHING PROCESS”, Thèse soutenue au Laboratoire MIT, 2001
- [LIU 1996] C-W. Liu, B-T. Dai, W-T. Tseng and C-F. Yeh, “Modeling of the Wear Mechanism during CMP”, *J. of Electrochem. Soc.*, Vol. 143, pp. 716-720, 1996.
- [LUO 2001] J. Luo and D. A. Dornfeld, “Material removal mechanism in chemical mechanical polishing: theory and modeling,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 14, No. 2, pp. 112-123, 2001.
- [LUO 2003 a] J. Luo and D. A. Dornfeld, “Material removal regions in chemical mechanical planarization for sub-micron integrated circuit fabrication: coupling effects of slurry chemicals, abrasive size distribution and wafer-pad contact area,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 16, No. 1, pp. 45-56, 2003.
- [LUO 2003 b] J. Luo and D. A. Dornfeld, “Effects of abrasive size distribution in chemical mechanical planarization: modeling and verification,” *IEEE Transactions on Semiconductor Manufacturing*, in press, 2003.
- [LUO 2003 c] J. Luo, “Wafer-Scale CMP Modeling of With-in Wafer Non-Uniformity”, *rapport technique à l'université de Berkeley*, 2003.
- [MOYNE 2001] Moyne, J., J. Kim, M. Beachy and T. Parikh (2001). “Determining RTR control deployment strategies from process run data”. *2nd European Advanced Equipment Control / Advanced Process Control (AEC/APC) Conference*, Dresden.

- [MUNGEKAR 2003] Mungekar. H., B. Geoffrion, B. Kapoor, N. Dubey, M. Salimian, M. Cox, and P. Krishnaraj (2003). “heat and mass transport in HDP-CVD chamber”, *Proceedings of the 2003 ASME Summer Heat Transfer Conference*, Las Vegas, NV.
- [MUNGEKAR 2005] Mungekar, H., Y.S Lee, and S. Venkataraman (2005). “Feature evolution during sub 100nm gap-fill and etch”, *Proceedings of HT05:2005 Summer Heat Transfer Conference*, July 17-22, , Westin St. Francis, San Francisco, CA,USA.
- [MUNGEKAR 2006] Mungekar, H. and Y.S. Lee (2006), “High density plasma chemical vapor deposition gap-fill mechanisms”, *American Vacuum Society*. 24(2), 11-15.
- [NAKAMURA 2001] Takashi Nakamura, “Novel Pulse Pressure CVD for Void Free STI Trench TEOS Fill”, *IEEE semiconductor manufacturing symposium*, pp.117-120, San Jose, CA, USA, 2001.
- [MURTHY 1997] C. Srinivasa-Murthy , D. Wanga, S.P. Beaudoin, T. Bibby, K. Holland, T.S. Cale, “Stress distribution in chemical mechanical polishing”, *Thin Solid Films* 308–309 (1997) 533–537.
- [NANZ 1995] Gerd Nanz and Lawrence E. Camilletti, “Modeling of Chemical-Mechanical Polishing:A Review”, *IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING*, VOL. 8, NO. 4, NOVEMBER 1995.
- [NISHIGUCHI 2001] Takashi Nishiguchi, “CMP Consumable Technologies in PERL HITACHI”, *ERC RETREAT & IAB meeting*, 2001
- [OTHA 1999] T. Otha, T. Toda and H. Ueno, “A Practical CMP Profile Model for LSI Design Application”, *SISPAD*, pp. 195-198, 1999
- [PARK 2003] Sung-Woo Park , Chul-Bok Kim , Sang-Yong Kim , Yong-Jin Seo, “Design of experimental optimization for ULSI CMP process applications”, *Microelectronic Engineering* 66 (2003) 488–495
- [PASQUALINI 2001] François Pasqualini, Martial Baudrier, “Application of Multivariate Statistical Process Control in real time for equipment parameters using PLS models”, *European AEC/APC conference*, 18-20 Avril 2001.Dresden.
- [PERRIN 2000] Emmanuel PERRIN, thèse de doctorat, “Etude et modélisation de l’aplanissement par polissage mécano-chimique des diélectriques dans les interconnexions des circuits intégrés”, Ecole centrale de Lyon, 2000.
- [PIERSON 1999] Hugh O. Pierson, “HANDBOOK OF CHEMICAL VAPOR DEPOSITION (CVD), Principles, Technology, and Applications”, *Second Edition, Noyes publications/william Andrew publishing,1999*, New York,(USA).
- [POTTHOFF 2006] Ulrich Potthoff, “MIMO Run-to-Run Control for LPCVD”, *7th AEC/APC European conference*, Aix en Provence, 29-31 Mars 2006
- [RAJAGOPAL 2006] Ramkumar Rajagopal, Kavitha Pathy, Kenneth Blancher, Clint Bowles, Chris Miles, “APC System for Process Capability Improvement of HDP Deposition” , *European AEC/APC conference*, Aix en provence, 29-31 Mars 2006.
- [SMITH 2002] Taber Smith, Duane Boning, Simon Fang,Greg Shinn, and Jerry Stefani, “A STUDY OF WITHIN-WAFER NON-UNIFORMITY METRICS”, 2002
- [SREENIVASAN 2001] Sreenivasan, Ramaswamy; Gougousi, Theodosia; Xu, Yiheng; Kidder, John, Jr.; Zafiriou, Evangelos; Rubloff, Gary W. “Run to run control in tungsten chemical vapor deposition using H₂ ’WF6 at low pressures”, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, Volume 19, Issue 5, September 2001, pp.1931-1941.
- [STEIGERWALD 1997] Joseph M. Steigerwald, Shyam P.Muraka, Ronald J. Gutmann, “Chemical mechanical planarization of microelectronic materials”, *John wiley & sons, Inc.*, New York, 1997

- [STEFANI 1996] Stefani, J.A, S. Poarch, S. Saxena and P.K. Mozumder (1996), “Advanced process control of a CVD tungsten reactor”. *IEEE Transactions on semiconductor Manufacturing* 9(3), 366-383.
- [STEFANI 2006] Jerry Stefani, “Run by Run APC Deployment in 65nm Thin Film Module”, *AEC/APC Symposium XVIII*, 30 Septembre- 4 Octobre 2006, Colorado(USA).
- [SUH 2003 a] S.-M. Suh, S. L. Girshick and M. R. Zachariah, “The Role Of Total Pressure In Gas Phase Nucleation: A Diffusion Effect”, *Journal of Chemical Physics* 118, 736-745 (2003).
- [SUH 2003 b] S.-M. Suh, S. L. Girshick,a) and U. R. Kortshagen, M. R. Zachariah, “Modeling gas-phase nucleation in inductively coupled silane-oxygen plasmas”, *Journal of Vacuum Science Technology*, Volume 21, Issue 1, pp. 251-264 (2003).
- [SREENIVASAN 2001] Ramaswamy Sreenivasan, Theodosia Gougousi, Yiheng Xu, and John Kidder, Jr. Evangelos Zafiriou, Gary W. Rubloff, “Run to run control in tungsten chemical vapor deposition using H₂ /WF₆ at low pressures”, *J. Vac. Sci. Technol. B* 19.5, Sep / Oct 2001,
- [SHAMOUILIAN 1996] Shamoulian, S and Shendon, N. Carrier “head with a layer of conformable material for a chemical mechanical polishing system”. *Applied Materials, Inc., US patent No. 58204448*: 1996.
- [SHENDON 1991] Shendon, N, Struven, K and Kolenkow, R. “Wafer polisher head having floating retainer ring”. *Cybeq System, Inc., US Patent No. 5,205,082*: 1991.
- [TROUNG 1998] P Troung and LR Blanchard, "Utilizing Pad Shaping as a Method to Stabilize Removal Rate, Improve Non-uniformity, and Increase Pad Life for Oxide CMP," in *Proceedings of the Third International Chemical Mechanical Planarization for ULSI Multilevel Interconnection Conference (CMP-MIC)* (Tampa, FL: IMIC, 1998), 351–356.
- [TSENG 1997] E. Tseng , C. Yi and H. C. Chen, “A Mechanical Model for DRAM Dielectric CMP Process”, *CMP-MIC*, pp. 258-264, Feb. 1997.
- [WANG 1997] D. Wang, J. Lee, K. Holland, T. Bibby, S. Beaudoin, and T. Calle, “Von Mises Stress in Chemical-Mechanical Polishing Processes”, *J. Electrochem*, Vol. 144, No.3, March 1997
- [WARNOCK 1991] J. Warnock, “A Two Dimensional Process Model for Chemimechanical Polish Planarisation”, *J. of Electrochem. Soc.*, pp. 2398-2402, August 1991.
- [WITVROUWA 2000] A. Witvrouw, B. Du Bois, P. De Moor, A. Verbist, C. Van Hoof, H. Bender , Kris Baert, “A comparison between wet HF etching and vapor HF etching for sacrificial oxide removal”, *Proceedings SPIE*, Vol.4174, 2000
- [XIN 1998] Yun-Biao Xin, “Modeling of Pad-Wafer Contact Pressure Distribution in Chemical-Mechanical Polishing”
- [XU 2000]Yiheng Xu, “Real-time in-situ chemical sensing, sensor-based film thickness metrology, and process control in W CVD process”, *Thèse de doctorat, Université Maryland*, 2000.
- [YU 1994] T-K. Yu, C. Yu and M. Orlowski, “Modeling of Chemical-Mechanical Polishing”, *Adv. Met. For ULSI App., MRS* 1994.
- [XU 2000] Yiheng Xu,“Real-time in-situ chemical sensing, sensor-based film thickness metrology, and process control in W CVD process”, *thèse de doctorat, Université de Maryland*, 2000.

Conclusion et perspectives

Avec le développement des nouvelles technologies (réduction des dimensions des circuits), des problèmes liés aux procédés de fabrication apparaissent et impactent négativement le rendement électrique final des circuits fabriqués. Ceci requiert donc l'introduction de nouvelles stratégies de contrôle des procédés. Dans le premier chapitre de cette thèse, nos recherches bibliographiques nous ont amenés à découvrir le domaine du contrôle des procédés dans le domaine de la fabrication des dispositifs à semi-conducteurs. Les techniques de contrôle statistique des procédés (SPC) ont été rappelées. Nous avons également discuté de l'application de la technique de détection et de classification des défauts FDC. L'incorporation de nouvelles méthodes telles que les boucles de régulation permet de contrôler le procédé en temps réel en réajustant les paramètres de procédé d'une façon automatique pour garantir les valeurs spécifiques physiques et électriques des produits. Le contrôle avancé des procédés est constitué donc de ces trois techniques de contrôle. Nous avons également décrit dans ce chapitre la méthodologie établie ayant servi pour l'implémentation des boucles de régulation. L'élaboration de cette méthode est faite à base de recherches bibliographiques mais aussi des contraintes spécifiques de l'usine de fabrication de ST- Microelectronics de Rousset (200mm) telles que l'échantillonnage des mesures et le type de métrologie disponible.

Dans cette thèse, nous avons appliqué les techniques de mise en place de boucles de régulation sur la brique technologique d'isolation des transistors. Un chapitre a été consacré à cette brique. Nous avons rappelé toutes les solutions technologiques proposées pour assurer une bonne isolation entre les transistors adjacents dans la technologie CMOS. Nous avons vu que la solution classique consiste à faire une oxydation locale du silicium (structure LOCOS). Avec la réduction des dimensions des transistors, cette structure présentait différents inconvénients. Le principal inconvénient du LOCOS est celui de la perte dimensionnelle créée par la croissance du bec d'oiseau (la largeur de la zone isolante est supérieur à celle définie par l'ouverture dans la couche de nitre de silicium. Il y a donc une baisse de la densité d'intégration des circuits). En plus du problème dimensionnel, cette isolation pose des problèmes électriques. Le plus grand problème électrique qui se manifeste dans le circuit est le phénomène de déverrouillage (ou latchup). Nous avons constaté que, pour éviter les

passages de courants parasites, il y a deux possibilités : soit d'éloigner les transistors les uns des autres, ce qui limite le niveau d'intégration du circuit, soit de disposer une zone isolante entre les transistors. C'est la deuxième solution qui est la plus efficace pour maintenir et accroître la densité d'intégration. Cette solution consiste à réaliser des tranchées d'isolation peu profondes (STI). Nous avons montré que le STI ne répond pas seulement aux besoins technologiques (l'élimination du phénomène du bec d'oiseau et la réduction de phénomène de déverrouillage) mais aussi, il offre une bonne planarité de la surface. Nous avons décrit le protocole de fabrication de la brique STI où après chaque opération de fabrication, une étape de mesures s'enchaîne pour vérifier si les grandeurs physiques sont comprises dans les limites de contrôle. Nous avons déduit que la hauteur de marche (la distance entre la région active et l'oxyde) est le paramètre de mesure qui caractérise la robustesse de l'isolation. Les variations de la valeur de ce paramètre influent directement le rendement électrique final.

Dans ce travail de thèse, nous avons axé nos recherches pour mettre en évidence l'influence du STI sur les paramètres électriques. Dans le chapitre 3, nous avons montré de forts taux de corrélations entre le STI et les paramètres électriques. Les corrélations entre la résistance carrée des caissons (Nwell STI) et l'épaisseur du STI pour les transistors PMOS sont très fortes. Ces fortes corrélations n'ont pas été observées pour les transistors NMOS. Les simulations de l'implantation ionique nous ont permis de démontrer l'origine de la différence entre les deux caissons. En analysant finement les profils de dopage pour les deux types de transistors, nous avons observé que la position du pic d'implantation dans le cas de Nwell (à $0,09\mu\text{m}$ de l'interface Si/STI) est différente de celle du Pwell (à $0,21\mu\text{m}$ de l'interface Si/STI). En modifiant les conditions des opérations du procédé d'implantation ionique (l'énergie d'implantation), nous avons constaté une forte linéarité entre la résistance Pwell STI et l'épaisseur STI. L'autre paramètre électrique qui est corrélé fortement avec la morphologie du STI est la tension de seuil des transistors parasites. Ce dernier présente une forte variabilité (environ 65% quand on fait varier la hauteur de marche de 150 \AA à 850 \AA). Cette forte variabilité a été un argument suffisant pour choisir ce paramètre comme un indicateur du suivi des boucles de régulation.

Dans la partie consacrée aux contraintes mécaniques, nous avons confirmé l'effet de la distance entre le STI et la grille (distance caractérisant l'éloignement de la tranchée par rapport à la grille (distance « a »)). Pour une diminution de cette distance, la mobilité des porteurs se dégrade dans le cas du NMOS alors qu'elle augmente dans le cas du PMOS. Nous avons interprété ces résultats par le fait que, quand le STI se rapproche de la zone active, il y a augmentation des contraintes mécaniques et en particulier en compression. Nous avons pu expliquer cette différence de comportement des transistors en procédant à des mesures sous flexion (nous avons imposé à un transistor une contrainte mécanique connue). L'objectif de ces expériences est d'analyser la variation de la mobilité en fonction du stress mécanique. Les résultats de ces analyses nous ont permis d'extraire les valeurs des coefficients piezorésistifs pour les deux types de transistors étudiés (NMOS et PMOS). Les valeurs et les signes des coefficients piezorésistifs extraits expliquent bien les différences de comportement des transistors NMOS et PMOS. Les mesures sous flexion ainsi que les résultats électriques nous ont permis de quantifier le stress mécanique en fonction de la distance « a ». Les résultats expérimentaux sont en accord avec les résultats de la modélisation numérique effectuée. Nous avons établi un modèle prédictif donnant le niveau des contraintes en fonction de la distance « a ».

Après avoir quantifié l'impact du STI sur les paramètres électriques, nous avons abordé dans le chapitre 4, les solutions pour réduire la variabilité de la hauteur de marche.

Les travaux de cette thèse ont été l'occasion d'acquérir une meilleure compréhension des variabilités générées au niveau de la brique STI. Grâce aux analyses de variabilités, nous avons pu quantifier la contribution de chaque étape du procédé sur la variabilité de la morphologie du STI. Les résultats de ces analyses ont montré que les étapes de procédés les plus critiques sont : la gravure de la tranchée, le remplissage des tranchées par l'oxyde, le polissage mécano-chimique (CMP). De ce fait, trois boucles ont été conçues :

- une boucle « feed forward » CMP- gravure humide (WET).
- une boucle feed back sur le procédé HDP-CVD
- une boucle feed back sur le procédé CMP

Pour compléter l'ensemble des actions entreprises sur la brique STI, une boucle de régulation R2R de type feed back sur le procédé de gravure (sèche) de la tranchée est en cours d'élaboration.

Pour la boucle sur le procédé HDP-CVD, deux modèles ont été recherchés ; le modèle d'épaisseur et le modèle de l'étendue d'épaisseur. En analysant les résultats de la méthode des moindres carrés (PLS), nous avons quantifié l'impact de chaque paramètre sur l'épaisseur déposée et nous avons sélectionné les variables critiques (la puissance RF Bias, les débits de gaz side et top). Les travaux de simulation ont démontré la robustesse du modèle de l'épaisseur. Il y a une forte corrélation entre les mesures prédictives et les mesures expérimentales. Les tests de robustesse du modèle de l'étendue d'épaisseur ont montré que ce modèle donne satisfaction seulement si les cartographies de mesures sont identiques à celles employées dans les conditions expérimentales et donc une dépendance avec le type de produit. L'outil statistique (Wafer Fit) nous a permis de reconstruire la surface de l'oxyde à travers la plaque. Cette reconstruction de surface a été faite grâce à des algorithmes d'apprentissage statistique basés sur l'interpolation des points de mesures. Nous avons démontré qu'avec l'emploi de cette méthode on peut détecter des situations « hors contrôle ».

Dans la seconde partie du chapitre 4, nous nous sommes préoccupés de la non-uniformité d'oxyde après CMP à travers la plaque dans le cas du STI standard. Nous avons démontré que la non uniformité générée par le CMP a été attribuée à la distribution de pression sur la plaque. L'optimisation du polissage d'oxyde, en particulier sur les bords des plaques en ajustant la pression nous permet d'avoir les mêmes caractéristiques électriques sur la plaque et ainsi de réduire la perte de rendement enregistrée sur les bords des plaques. Les analyses des données faites entre les mesures post-CMP et V_t montrent une forte corrélation. Nous nous sommes enfin préoccupés de la densité de défauts sur les bords des plaques. Les résultats des analyses des corrélations entre les mesures d'épaisseurs d'oxyde post-CMP et la densité de défauts sur les bords nous ont permis de déduire l'origine des défauts. Cette étude nous a permis de prédire les défauts au niveau des premières opérations du procédé.

Dans la dernière partie de ce chapitre, nous avons étudié le modèle de la boucle « feed-forward » entre le CMP et la gravure humide. Nous avons été les premiers à travailler sur la conception de la boucle R2R entre ces deux procédés. Le modèle de cette boucle est basé sur la recommandation d'une nouvelle valeur de temps de gravure HF pour le lot suivant les résultats post CMP du lot précédent. Le modèle établi a donné d'excellents résultats sur le plan industriel. Ce modèle a été testé en pré-production où des lots ont été sélectionnés pour le test. Certains lots ont vu leur temps de HF modifié et après la mesure de la hauteur de marche, on a pu constater que tous les résultats de la hauteur de marche sont proches de la valeur cible.

Nous pouvons conclure que ces travaux peuvent servir de référence pour d'autres travaux sur le STI pour d'autres technologies, en particulier pour les technologies les plus modernes (90, 65 nm et même en deça) où les problèmes de contraintes mécaniques sont importants avec la réduction des dimensions des transistors. Les perspectives de nos travaux peuvent s'orienter vers trois directions, la première concerne le domaine de la surveillance des boucles de régulation. Une première approche de superviseur a été proposée dans ce manuscrit, d'autres actions sont à l'étude. La deuxième perspective s'inscrit dans le domaine de la métrologie virtuelle. Car, comme dans tous les sites de production, l'un des défis majeurs est le gain en temps et en qualité. L'échantillonnage des mesures pose toujours un problème et notamment pour les boucles de régulation qui doivent être alimentées par des données de mesure. C'est pour cette raison que la métrologie virtuelle (prédition des valeurs de mesures) peut être un élément majeur pour la robustesse des boucles de régulation et une contribution à l'amélioration du temps de cycle. Le temps d'attente de mesures des lots et le temps de transport dans le cas de la non disponibilité de la métrologie intégrée peuvent impacter négativement les boucles de régulation et donc le temps de cycle. La troisième perspective de ces travaux concerne un futur axe de recherche où les travaux d'ordonnancement (Scheduling) peuvent être associés pour améliorer les modèles proposés. Cette dernière approche peut s'inscrire dans un nouvel axe de recherche très prometteur APC-Scheduling pour une application dans les usines de fabrication.