



HAL
open science

Développement d'une méthodologie adaptée à l'industrie microélectronique pour la reconstruction topographique par imagerie SEM à faisceau inclinable

Charles Valade

► To cite this version:

Charles Valade. Développement d'une méthodologie adaptée à l'industrie microélectronique pour la reconstruction topographique par imagerie SEM à faisceau inclinable. Micro et nanotechnologies/Microélectronique. Université Grenoble Alpes [2020-..], 2020. Français. NNT : 2020GRALT015 . tel-03016206

HAL Id: tel-03016206

<https://theses.hal.science/tel-03016206>

Submitted on 20 Nov 2020

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **NANO ELECTRONIQUE ET NANO TECHNOLOGIES**

Arrêté ministériel : 25 mai 2016

Présentée par

Charles VALADE

Thèse dirigée par **Cécile GOURGON**, Directrice de recherche,
Communauté Université Grenoble Alpes

préparée au sein du **Laboratoire des technologies de la
Microélectronique (CNRS-LTM)** en collaboration avec le
**Laboratoire d'Electronique et des Technologies de
l'Information (CEA-LETI)** et **STMicroelectronics**,
dans l'**École Doctorale Electronique, Electrotechnique,
Automatique, Traitement du Signal (EEATS)**.

Développement d'une méthodologie adaptée à l'industrie microélectronique pour la reconstruction topographique par imagerie SEM à faisceau inclinable.

Thèse soutenue publiquement le **13/03/2020**,
devant le jury composé de :

Dr Alexandra DELVALLEE

Ingénieure de recherche au LNE, Examineur

Dr Patrick SCHIAVONE

Directeur de recherche au CNRS-LTM, Président du jury

Dr Tatiana NOVIKOVA

Ingénieure de recherche au CNRS-LPICM, Rapporteur

Pr Frédéric TRUCHETET

Professeur à l'université de Bourgogne, Rapporteur

Dr Cécile GOURGON

Directrice de recherche au CNRS-LTM, Membre (Directrice de thèse)

Dr Aurélien FAY

Ingénieur de recherche au CEA-LETI, Membre (Encadrant)

Dr Maxime BESACIER

Maitre de conférences au CNRS-LTM, Invité (Encadrant)

Dr Elodie SUNGAUER

Ingénieure à STMicroelectronics, Invitée (Encadrante)

Dr Jérôme HAZART

Ingénieur de recherche au CEA, Invité (Encadrant)



Table des matières

1	Introduction aux techniques de métrologie	9
1.1	La microélectronique	9
1.2	Microscopie électronique à balayage	10
1.3	Métrologie 3D	12
1.3.1	Microscope à force atomique (AFM)	12
1.3.2	Sonde ionique focalisée (FIB)	14
1.4	SEM à faisceau électronique inclinable	15
1.5	Imagerie photonique vs imagerie électronique	16
1.5.1	Source lumineuse, source électronique et formation de l'image	18
1.5.2	Matériau de l'objet observé	19
2	Etat de l'art : Reconstruction tridimensionnelle	23
2.1	Introduction	23
2.2	Reconstruction géométrique par appariement de points et triangulation	24
2.3	Reconstruction photométrique	29
2.4	Méthode de reconstruction hybride	31
2.5	Simulation de la formation d'une image SEM	33
2.5.1	Les modèles compacts	35
2.6	Résolution du problème inverse	36
2.7	Conclusion	38
3	Outils pour la reconstruction 3D	39
3.1	Paramétrisation de la géométrie : Approximation "8 paramètres"	39
3.1.1	Étude de la courbure	40
3.1.2	Définition du SWA	41
3.1.3	Définition des CR	42
3.1.4	Application du modèle paramétrique sur des données expérimentales	42
3.2	Analyse d'une image SEM	44
3.2.1	Transformation des images en profil et méthode du ruban	44
3.2.2	Limite du ruban	45
3.2.3	Les descripteurs	45
4	Détermination des tendances d'évolution des descripteurs, à partir de l'analyse de simulations physiques	47
4.1	Intérêt de la simulation SEM	47
4.2	Paramètres du simulateur JMONSEL	48
4.3	Influence de la hauteur et du SWA sur le signal SEM	49
4.3.1	Influence du SWA	49

4.3.2	Influence du SWA sur le CDt	50
4.3.3	Influence de la hauteur	52
4.3.4	Influence de la hauteur et du SWA sur l'EW	52
4.4	Influence des CR sur le signal SEM	53
4.4.1	Influence des CR sur le signal SEM global	53
4.4.2	Influence des CR sur l'EW	56
4.4.3	Influence des CR sur le CDt	58
4.5	Limite de la méthode pour les petits CD	59
4.6	Influence du tilt sur le signal SEM	61
4.6.1	Etude théorique	61
4.6.2	Cas pratique de simulation	62
4.7	Analyse de l'effet d'ombrage	64
4.8	Conclusion	66
5	Le modèle INVSEM3D : Création, calibration et reconstruction pour des lignes isolées en silicium	67
5.1	Création du modèle	67
5.1.1	Écriture du problème sous forme d'un problème linéaire	67
5.1.2	La régularisation	68
5.1.3	Optimisation des paramètres pour un motif de type ligne	70
5.1.4	Apport du modèle par rapport aux méthodes classiques	79
5.1.5	Conclusion	82
5.2	Calibration expérimentale du modèle INVSEM3D	83
5.2.1	Création de motifs en silicium gravé	83
5.2.2	Mesures de référence par AFM	84
5.2.3	Calibration	85
5.3	Reconstruction de motifs grâce au modèle INVSEM 3D	97
5.4	Conclusion	99
6	Extension du modèle aux tranchées isolées et aux denses	101
6.1	Extension du modèle aux tranchées	101
6.1.1	Comparaison des descripteurs utilisés pour les lignes et pour les tranchées	101
6.1.2	Analyse JMONSEL pour les tranchées	104
6.2	Approximation d'un motif dense par une succession de motifs tranchées	111
6.3	Conclusion	113
7	Analyses statistique	115
7.1	Impact du paramètre d'ajustement de la spline, sur le placement des descripteurs	115
7.2	Impact du nombre d'électrons envoyés par pixel dans JMONSEL, sur le placement des descripteurs	118
7.3	Étude de la propagation de l'incertitude de placement des descripteurs sur l'estimation des grandeurs d'intérêts	120
7.3.1	Calcul empirique des erreurs, sur des résultats de simulation	120
7.3.2	Calcul théorique des erreurs	122
7.3.3	Vérification des calculs théoriques sur les résultats empiriques	123

8	Conclusions et perspectives	125
8.1	Conclusions	125
8.2	Perspectives	127
8.2.1	Propositions d'amélioration	127
8.2.2	Applications spécifiques	129
8.3	Remerciements	130
	Appendices	133

Notations

Dans ce manuscrit de thèse des notations et abréviations seront utilisées par souci de rapidité et/ou de problème de traduction.

Soit la figure suivante, représentant une vue en coupe d'une ligne (bleu), les notations utilisées (noir) et des lignes de construction (vert) :

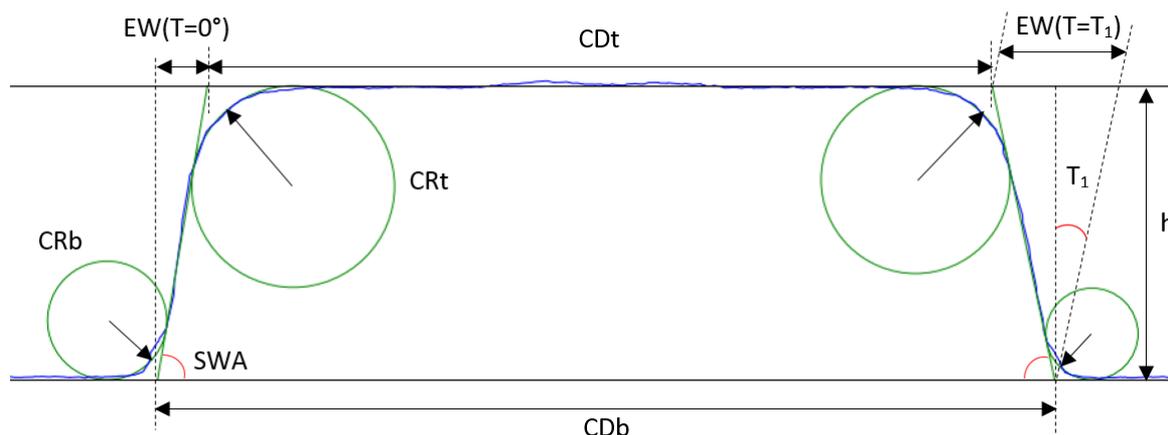


FIGURE 1 – Représentation des grandeurs d'intérêt sur un motif de type "ligne". Les abréviations sont explicitées ci-dessous.

h : Hauteur du motif. La hauteur peut être négative pour les motifs de type tranchée.
 CD : Dimension caractéristique du motif observé (Critical Dimension) soit en haut du motif (CDt) soit en bas du motif (CDb), soit à une hauteur quelconque : CD(h).
 CRt : Dimension du congé en haut du flanc du motif (Corner Rounding Top).
 CRb : Dimension du congé en bas du flanc du motif (Corner Rounding Bottom).
 SWA : Pente du flanc du motif (SideWall Angle). Dans mon travail j'ai donné une définition mathématique à cette grandeur (voir partie 3.1). Avec la convention utilisée dans ce travail un motif à angle droit a un SWA de 90°, mais on voit parfois la convention SWA = 0° pour des angles droits dans la littérature.
 T : Angle d'observation du motif (Tilt).
 EW : Largeur du flanc du motif (Edge Width). Dans mon travail j'ai fait le choix que cette grandeur ne soit pas intrinsèque au motif mais dépendante de l'angle d'observation du motif :

$$EW(T) = \frac{h}{\tan(SWA)} + h \times \tan(T) \quad (1)$$

On pourra éventuellement rajouter un "l" ou un "r" pour indiquer une grandeur du flanc de gauche (left) ou de droite (right) : CRtl, SWAr...

Wafer : Plaquette de silicium mono-cristallin sur laquelle les circuits imprimés sont réalisés. Dans notre étude, on a travaillé avec des wafers de 300 mm de diamètre.

SEM : Microscope électronique à balayage (Scanning Electron Microscope)

CD-SEM : SEM spécialement conçu pour la mesure de masse de CD dans l'industrie.

FOV : Zone bi-dimensionnelle balayée par la machine d'observation (Field Of View).

Spot size : Taille du faisceau électronique à la surface du wafer. La littérature s'accorde à dire

que la probabilité de présence d'un électron envoyé en un point (x_0, y_0) est une gaussienne 3D centrée en (x_0, y_0) [1] [2]. En général dans la littérature, la valeur du spot size est le diamètre du cercle dans lequel la majorité des électrons seront présents (souvent sigma : 68.3%), c'est également la convention que l'on prendra dans ce travail.

TEM : Microscope électronique en transmission (Transmission Electron Microscope)

AFM : Microscope à force atomique (Atomic Force Microscope)

FIB : Sonde ionique focalisée (Focus Ion Beam)

Introduction générale

Les dispositifs de microélectronique sont de plus en plus petits et présentent des architectures de plus en plus complexes. Dans ces conditions, l'industrie de la microélectronique a recours à de la métrologie 3D pour contrôler et améliorer ses procédés. Cependant la métrologie locale actuelle souffre d'une lenteur inhérente (AFM) et est potentiellement destructive (FIB/SEM/TEM). Une métrologie 3D rapide, reproductible, non destructive et précise est donc un enjeu majeur pour l'industrie de la microélectronique.

Le but du travail de thèse qui va être présenté dans ce manuscrit est le développement d'une méthodologie permettant l'obtention de mesures à partir d'images de microscopie électronique, afin d'obtenir des informations tri-dimensionnelles sur les structures observées. Pour ce faire, un SEM dont le faisceau électronique est inclinable sera utilisé, de manière à obtenir des images à différents angles de vue. A partir de ces images une méthode de reconstruction basée sur un modèle linéaire permettra l'estimation des grandeurs d'intérêt. Pour le développement de cette méthodologie, un motif typique de la microélectronique a été étudié : une ligne isolée gravée dans le silicium. La généralisation à d'autres types de structures sera étudiée.

Dans le premier chapitre, on introduira les différentes techniques de métrologie permettant l'obtention de mesures topographiques, qui ont été utilisées dans ce travail, soit comme mesure de référence (AFM, FIB/SEM/TEM), soit pour la prise d'image SEM à différents angles de vue.

Dans la seconde partie, on présentera l'état de l'art de la reconstruction tri-dimensionnelle, dans le cas général, c'est à dire à partir d'images classiques, puis dans le cas spécifique d'images de microscopie électronique. Enfin on présentera les méthodes actuelles de simulations d'images SEM par résolution physique et par modèles compacts, ainsi que les méthodes de résolution du problème inverse.

Dans la troisième partie, on introduira deux outils qui seront la base de toutes les analyses présentées dans ce manuscrit. D'une part on présentera la paramétrisation de la géométrie d'un motif typique de la microélectronique et d'autre part, la méthode de transformation d'une images SEM en une série de descripteurs.

Dans la quatrième partie sera présentée l'analyse des différentes simulations physiques, qui ont permis de repérer les tendances d'évolution des descripteurs en fonction de la topographie du motif observé et des conditions de prise d'image.

Dans la cinquième partie, la création d'un modèle linéaire sera présentée. La calibration expérimentale pour la reconstruction de lignes isolées en silicium sera ensuite étudiée, ainsi que des résultats de reconstruction sur des structures réelles.

Dans la sixième partie, la même méthodologie sera utilisée pour créer un modèle linéaire pour la reconstruction de tranchées en silicium. On verra cependant les limitations rencontrées sur ce type de motif. L'adaptation du modèle tranchée pour des motifs denses sera également étudiée.

Dans la septième partie nous présenterons quelques analyses statistiques permettant de mieux comprendre l'impact de certains paramètres d'entrée de nos analyses, comme le nombre

d'électrons par pixel des simulations ou le niveau d'ajustement des splines approximant les profils SEM.

Enfin nous concluons et présenterons les possibilités d'amélioration du travail présenté ainsi que les applications potentielles pour les besoins de l'industrie de la microélectronique.

Chapitre 1

Introduction aux techniques de métrologie

1.1 La microélectronique

Depuis l'invention du transistor en 1947, brique élémentaire des circuits électroniques, l'industrie des semi-conducteurs n'a cessé de croître, permettant une avancée considérable des technologies numériques. Pour pouvoir répondre aux attentes de rapidité de calcul et d'encombrement des machines, la taille de la grille des transistors a été réduite en suivant la loi de Moore. Cependant depuis les années 2015-2016 la course à la miniaturisation des transistors a fortement ralenti car le rapport entre le coût pour passer à la génération de taille de grille suivante par rapport aux bénéfices d'une telle technologie est devenu trop important. Désormais pour pouvoir encore améliorer les performances des transistors ce n'est donc plus uniquement sur la taille que l'on va jouer mais aussi sur leur architecture [3][4].

La première avancée dans ce domaine a eu lieu en 1978 avec l'apparition de la technologie SOI permettant de limiter les courants de fuite [5] grâce à l'ajout d'une couche isolante de SiO₂ sous le transistor et entre les transistors, puis avec sa démocratisation à partir des années 1990 grâce à l'invention du procédé Smart Cut [6], ainsi que l'utilisation de matériaux isolants à forte constante diélectrique : les high-k. Depuis on peut également ajouter une contrainte mécanique sur le canal ; En effet lorsque que l'on applique une contrainte sur le canal on peut réduire ou augmenter le paramètre de maille du Si cristallin ce qui va modifier la structure de bande et permettre d'améliorer la mobilité des trous (compression) ou celle des électrons (tension) [7].

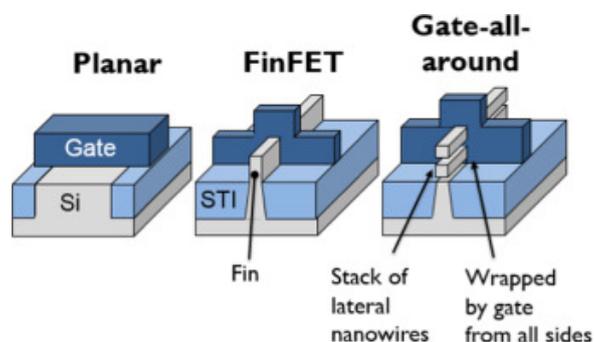


FIGURE 1.1 – Evolution du CMOS [8].

Désormais l'une des voies d'améliorations des transistors est de réaliser des architectures non planaires : Fin FET, Fin FET Tri-gate, Gate-all-around... (Illustré dans la Fig 1.1) [4]. Si jusque-là le comportement des transistors était pratiquement entièrement lié à la taille de la grille (géométriquement parlant) appelée CD (pour Critical Dimension), désormais le comportement

des transistors 3D pourra aussi être relié à leur hauteur ou à l'angle des flancs de leurs motifs ou encore à l'empilement des matériaux utilisés [9][4]. C'est pourquoi le développement de ces nouvelles architectures nanométriques ne peut se faire qu'en parallèle d'une métrologie nanométrique permettant la visualisation et le dimensionnement dans les trois dimensions de l'espace [4][10][11].

De plus les applications développées par l'industrie des micro/nanotechnologies se diversifient de plus en plus : imageurs (Fig 1.2), MEMS, mémoire type 3D NAND [12], photonique, ordinateur quantique... Ces nouvelles applications donnent lieu à des structures très hétérogènes et dont l'architecture 3D peut être compliquée [9].

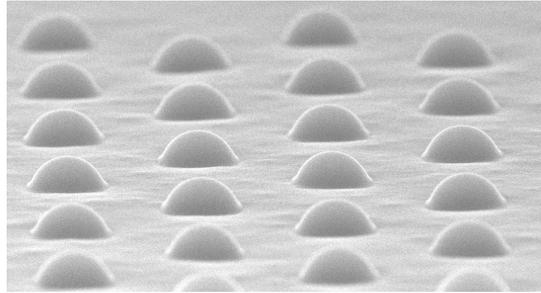


FIGURE 1.2 – Microlentilles pour des applications imageur (CEA LETI).

C'est pour toutes ces raisons que cette thèse se propose de répondre à une demande du milieu industriel quant à la création d'une méthode de métrologie tridimensionnelle rapide et non destructive. Dans le milieu industriel la rapidité de la mesure est évidemment une caractéristique très importante pour pouvoir réaliser des mesures en masse sans ralentir la production, ce qui permet par exemple d'effectuer des études de variabilité sur un wafer requérant beaucoup de mesures. La non destructivité permet quant à elle le retour de la plaque observée dans la chaîne de production permettant de réaliser les mesures directement sur les produits. La microscopie électronique se positionne comme une technique adaptée pour ce type de contraintes. Cependant, il est pour le moment difficile d'utiliser la microscopie électronique pour obtenir des informations tri-dimensionnelles précises.

1.2 Microscopie électronique à balayage

Un microscope électronique à balayage (Scanning Electron Microscope : SEM) est un instrument de mesure permettant l'observation d'objets de très petite taille. Il a été créé pour repousser la limite de résolution des microscopes optiques qui sont limités par la longueur d'onde de la lumière. En effet le critère Rayleigh propose une limite à partir de laquelle une onde électromagnétique de longueur d'onde λ ne peut plus permettre de créer une image dont on peut discerner deux points séparés par une distance δ appelée limite de résolution : $\delta = 0,61\lambda/NA$, avec NA l'ouverture numérique de l'instrument.

A la longueur d'onde de la lumière visible il est donc difficile de faire une image "correcte" de motifs dont la taille est inférieure à 300 nm environ.

Dans le cas du microscope électronique la particule utilisée pour réaliser une image n'est plus un photon mais un électron, qui, lorsqu'il est accéléré et utilisé en faisceau, peut être assimilé à une onde électromagnétique de longueur d'onde :

$$\lambda = \frac{h}{\sqrt{2m_0eU}} \times \frac{1}{\sqrt{1 + \frac{eU}{2m_0c^2}}}, \quad (1.1)$$

avec h la constante de Planck ($J.s$), m_0 la masse d'un électron (kg), e la charge d'un électron (C), U le potentiel d'accélération appliqué sur les électrons (V) et c la célérité de la lumière dans le vide ($m.s^{-1}$).

Le premier terme vient de l'équation de De Broglie et le second terme de la prise en compte de la relativité lorsque la vitesse des électrons n'est plus négligeable devant celle de la lumière.

D'après cette équation on devrait pouvoir obtenir la limite de résolution que l'on souhaite en augmentant la tension d'accélération des électrons dans la colonne électronique. Cependant dans un microscope électronique ce n'est plus la longueur d'onde qui devient le facteur limitant sur la résolution maximale atteignable, mais la taille de la zone des interactions électrons-matière, voire la taille du faisceau électronique à la surface de l'échantillon (spot size) dans le cas d'électrons peu accélérés; en effet lorsqu'on augmente l'accélération des électrons ceux-ci vont pénétrer plus profondément dans le matériau et de ce fait l'information sera de moins en moins locale. Cependant, si les électrons ne sont pas suffisamment accélérés il sera difficile de les focaliser en une zone réduite (car ils se repoussent entre eux).

La tension d'accélération des électrons va donc être utilisée, non pas pour augmenter ou diminuer la résolution, mais plutôt en fonction de l'utilisation souhaitée; par exemple si on s'intéresse à la topographie d'un matériau, on va utiliser une tension d'accélération très faible pour que les électrons pénètrent peu profondément dans la matière. Au contraire si on souhaite connaître l'empilement des matériaux de l'échantillon observé, on va utiliser une forte accélération pour pénétrer profondément dans les différentes couches du matériau.

Les électrons qui sont produits par le canon à électrons puis accélérés et envoyés sur l'échantillon sont nommés les électrons primaires. Lorsque ces électrons primaires entrent dans le matériau, ils peuvent subir différentes interactions physiques. Un détecteur d'électrons positionné dans la chambre de l'échantillon va collecter les électrons ressortant du matériau. On peut catégoriser les électrons ressortant du matériau en deux catégories, ceux ayant une énergie faible (typiquement moins de 50 eV) et ceux ayant une énergie proche de l'énergie des électrons primaires (Fig 1.3).

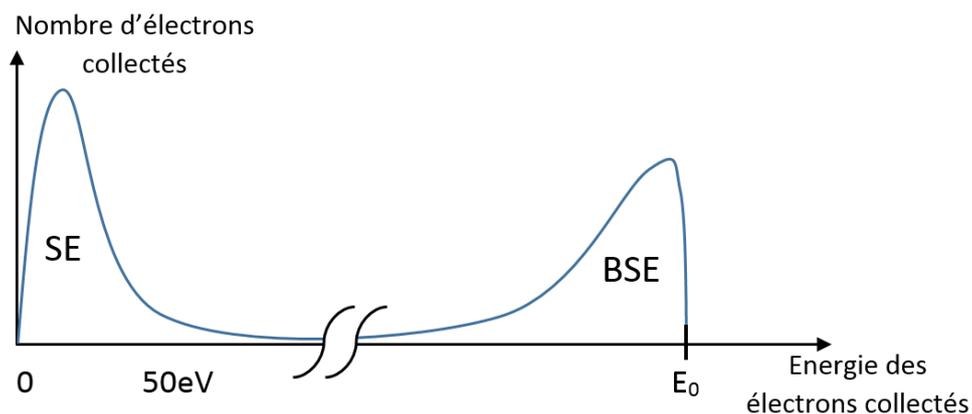


FIGURE 1.3 – Représentation schématisée simplifiée de la répartition énergétique des électrons sortant d'un matériau qui a été bombardé par un faisceau d'électrons primaires d'énergie E_0 .

Les électrons collectés dont l'énergie est proche de celle des électrons primaires, sont des électrons primaires dont la direction a été modifiée par une interaction électrostatique avec le noyau des atomes du matériau sondé (Fig 1.4a). Ces électrons sont appelés électrons rétro-diffusés (ou BSE pour "Back Scattered Electron"). Comme ces électrons ont une énergie importante, n'ayant subi que des interactions quasi-élastiques, l'information qu'ils transportent

peut provenir d'une profondeur plus ou moins importante en fonction de l'énergie des électrons primaires. Les BSE ont interagi avec les noyaux atomiques, la réponse électronique dépendra donc (non-exhaustivement) du matériau de l'échantillon dans la zone sondée.

La majorité des électrons collectés dont l'énergie est faible (typiquement moins de 50 eV), sont des électrons qui ont été éjectés de leur orbital (généralement peu liés dans la bande de conduction) par une collision avec un électron primaire (ionisation, Fig 1.4b). Ces électrons sont appelés électrons secondaires (ou SE pour "secondary electron"). La faible énergie de ces électrons fait qu'ils ne peuvent ressortir du matériau et donc être détectés, que si la collision a lieu suffisamment proche de la surface du matériau. L'information portée par les SE est donc bien plus locale que celle des BSE, c'est pourquoi ils sont souvent utilisés pour les études de topographie. C'est donc ce type d'électrons qui va nous intéresser dans ce travail.

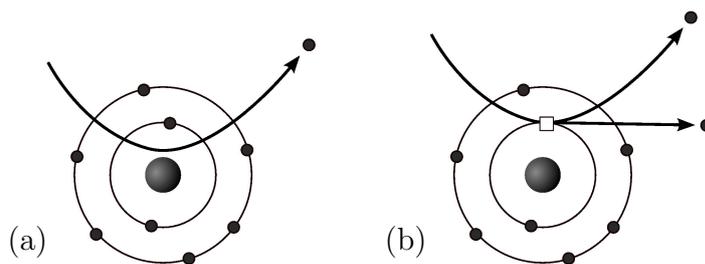


FIGURE 1.4 – Représentation schématique de la création d'un électron rétro-diffusé (a) et d'un électron secondaire (b) [13].

1.3 Métrologie 3D

Actuellement il existe trois principaux équipements permettant d'obtenir des informations de topographie à la surface d'un wafer.

Le microscope à force atomique qui sonde la surface de manière mécanique grâce à une pointe suivant la topographie de la surface, les méthodes liées à une découpe du matériau dans l'axe de la hauteur pour imager en coupe (cross-section, FIB SEM, FIB TEM) et la scattérométrie qui est une méthode de mesure indirecte qui consiste à étudier la réponse optique d'une surface. Cependant cette technique n'est utilisable que sur des motifs répétés sur une "large" zone (réseau de lignes), c'est pourquoi le résultat obtenu sera la topographie moyenne des motifs de la zone. On ne s'intéressera donc pas à cette technique dans cette thèse car nous cherchons à obtenir des informations de topographie locale.

1.3.1 Microscope à force atomique (AFM)

Le microscope à force atomique est un appareil de mesure en champ proche, c'est-à-dire que pour fonctionner, le capteur (qui est une pointe en matériau carboné dur de type diamant) doit se trouver dans le potentiel électrostatique créé par la surface que l'on désire mesurer. Son principe de fonctionnement repose sur le balayage d'une nano-sonde à la surface d'un échantillon pour pouvoir en déterminer la topographie. La pointe est mécaniquement rapprochée de la surface jusqu'à ce que le système mécanique détecte la force répulsive de la surface à analyser ; Cette force provient à la fois de la répulsion coulombienne entre les nuages électroniques des atomes de la pointe et de la surface ainsi que de l'effet quantique qui empêche deux fermions (particule de spin demi-entier comme l'électron) de se trouver simultanément dans le même état quantique : le principe d'exclusion de Pauli. En pratique il existe différentes descriptions empiriques des forces attractives et répulsives qui se créent lorsque l'on rapproche deux objets

entre eux. La description la plus couramment admise est le potentiel de Lennard Jones (Fig 1.5).

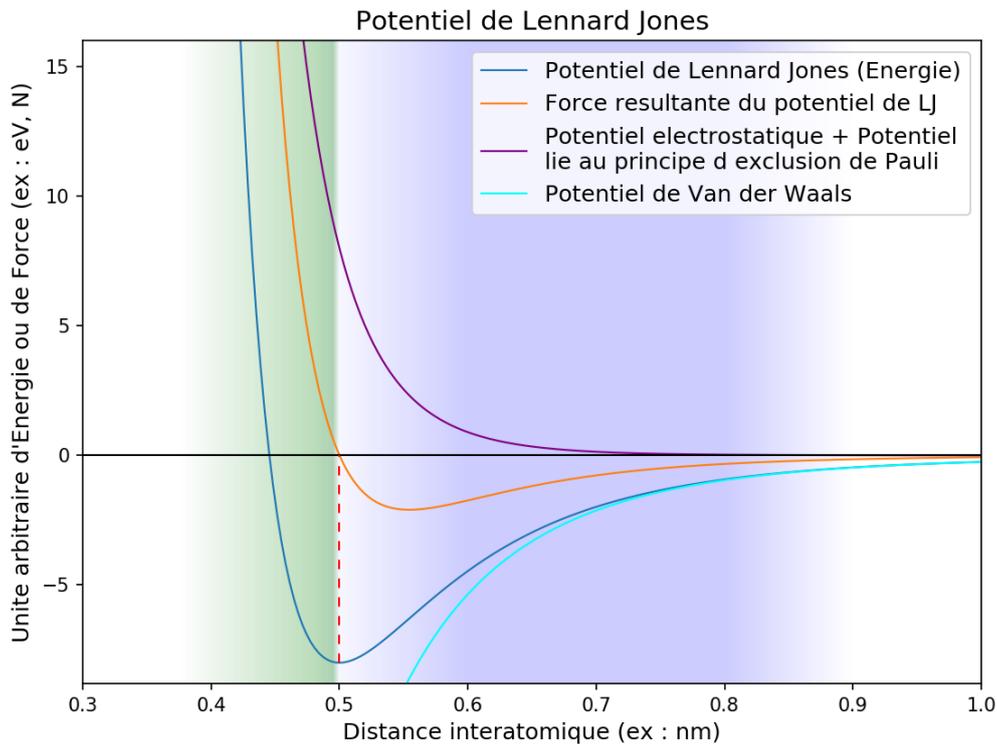


FIGURE 1.5 – Explication physique de : "Qu'est-ce que toucher?". Potentiel de Lennard Jones et la force résultante associée ($F = -grad(U)$) en unités arbitraires. Les forces positives correspondent à une force de répulsion et les forces négatives à une force d'attraction. La zone verte correspond au domaine du mode contact, la zone bleue correspond au domaine du mode tapping de l'AFM.

Il existe deux modes d'utilisation d'un AFM :

Le mode contact : La pointe suit l'échantillon en continu en appliquant, grâce à un asservissement, une force constante sur l'échantillon ; La topographie de surface est alors repérée grâce à la mesure de l'inflexion de la pointe. Cette inflexion est généralement mesurée grâce à la réflexion d'un laser sur le micro-levier vers un cadran de photo-diode. Les avantages de ce mode sont la rapidité et la simplicité de mise en œuvre, mais, en contrepartie, elle a l'inconvénient d'user rapidement les pointes utilisées, voire même l'échantillon à cause de la pression mise en jeu (rapport force sur la surface d'application).

Le mode tapping : Le micro-levier est mis en vibration à sa fréquence de résonance propre (cad sa fréquence de résonance loin de la surface). Lorsque la pointe est rapprochée de la surface sa fréquence de résonance va être modifiée par la perturbation électrostatique principalement provoquée par le potentiel de Van der Waals créé par la surface, permettant ainsi sa détection.

L'AFM est considéré comme une excellente technique pour obtenir avec précision la hauteur d'une structure, en revanche l'information sur la topographie des flancs peut être difficile voire impossible à mesurer à cause des effets liés à la géométrie de la pointe et la pente du flanc [14] (Fig 1.6). C'est d'ailleurs l'une des difficultés de cette méthode, car on devra toujours mesurer la forme de la pointe pour réaliser des mesures consistantes entre elles. De plus, comme les

pointes AFM sont relativement fragiles, elles peuvent s'abîmer voire se casser lors de la mesure. De ce fait, il est difficile de faire des mesures de masse reproductibles, car on doit souvent changer de pointe et la forme de la pointe peut même changer en cours de mesure.

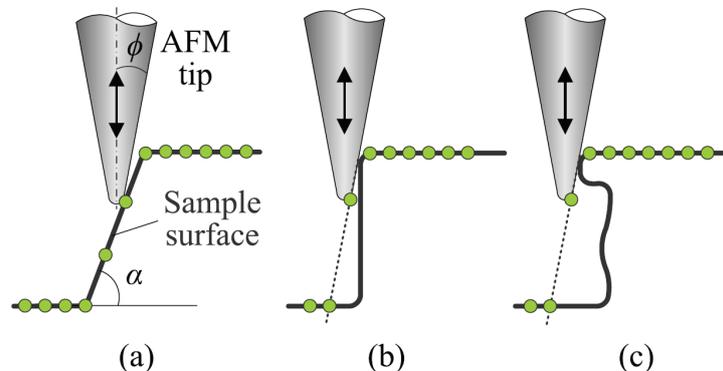


FIGURE 1.6 – Limitation de la mesure AFM en fonction de la forme de la pointe par rapport à la topographie. Dans cet exemple (non-exhaustif), lorsque $90 - \alpha > \phi$ la topographie est mesurée (a), lorsque $90 - \alpha < \phi$ c'est la forme de la pointe qui est mesurée sur la partie en pointillé (b), on ne peut alors plus conclure sur la forme du motif sur cette zone (c). [14]

Pour être capable d'analyser des profils ré-entrants, des techniques spéciales d'AFM souvent appelées AFM-3D ont été développées, soit en inclinant la pointe et l'échantillon relativement l'un à l'autre, soit grâce à des pointes avec des formes en "T" inversés dont la vibration n'est pas uniquement verticale mais possède également une composante latérale pour repérer les concavités [14].

1.3.2 Sonde ionique focalisée (FIB)

Le Focus Ion Beam ou FIB n'est pas une technique de mesure en elle-même, mais une méthode de préparation que l'on pourra combiner avec une observation au microscope électronique à balayage ou en transmission de manière à obtenir une vue en coupe des motifs observés. Le FIB est une technique de micro-usinage permettant de graver la matière par bombardement d'ions à forte énergie (généralement du Gallium) ; il s'agit donc d'une technique localement destructive.

La préparation va être différente en fonction du type d'analyse que l'on souhaite réaliser : Si on veut faire une observation au SEM on va creuser sur la hauteur du motif que l'on souhaite observer, de manière à pouvoir imaginer ce motif en coupe. Comme le flanc à observer est vertical il va falloir incliner l'échantillon pour que le faisceau d'électrons du SEM soit en face de la zone à imaginer ; il est de ce fait non-adapté d'utiliser cette technique sur un wafer entier. Si on veut également avoir une information de topographie sur la longueur du motif on peut faire des gravures successives de manière à imaginer des tranches sur la longueur du motif, cette technique est appelée tomographie.

Si on veut faire une observation au TEM il faut en revanche réaliser une lame mince, c'est à dire graver de part et d'autre de la zone que l'on souhaite observer (Fig 1.7). Pour cela on doit tout d'abord déposer une couche de tungstène qui va protéger les motifs ainsi qu'apporter une meilleure tenue mécanique à la future lame mince. Une fois amincies les lames sont prélevées du wafer grâce à une pointe qui est soudée sur la couche de tungstène, précédemment déposée

puis extraite pour être déposée sur le porte échantillon du TEM.

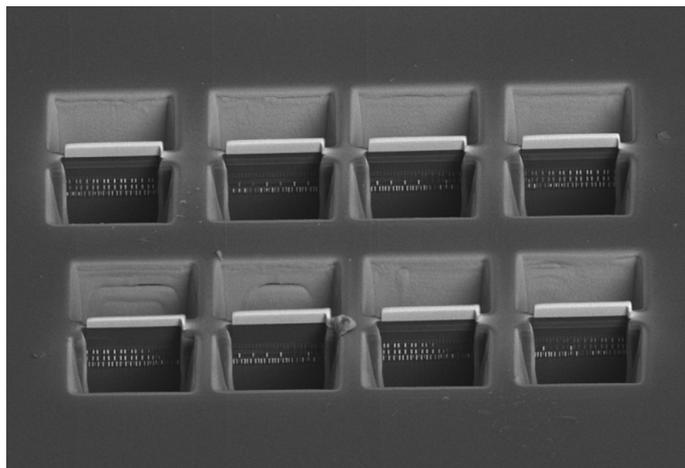


FIGURE 1.7 – Vue au microscope électronique à balayage de plusieurs coupes FIB préparées en lames minces pour l’observation TEM (FOV = $90\mu\text{m}$). Image utilisée avec l’accord de ”Orsay Physics”, prise à l’aide du FIB dual-beam ”Orage” [15].

Pour pouvoir réaliser une mesure directement sur le wafer une méthode de coupe FIB ”inclinée” a été développée [16], permettant ainsi une observation SEM en topview (Fig 1.8), devant ensuite être corrigée par le facteur lié à l’inclinaison de la surface observée.

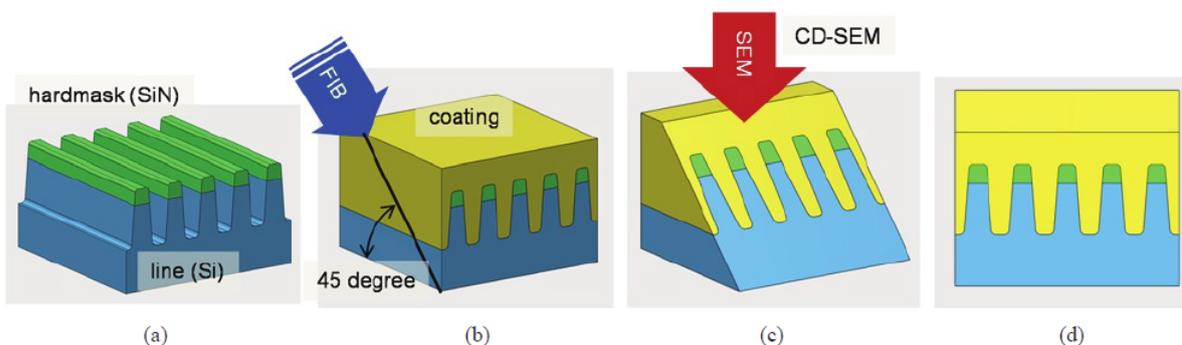


FIGURE 1.8 – Etapes d’une mesure type ”FIB to CDSEM” ; (a) Dispositif 3D avec des lignes et un hardmask, (b) revêtement et coupe FIB à un angle de 45° , (c) Prise d’image CD-SEM en topview et (d) son résultat pour reconstruction 3D [16].

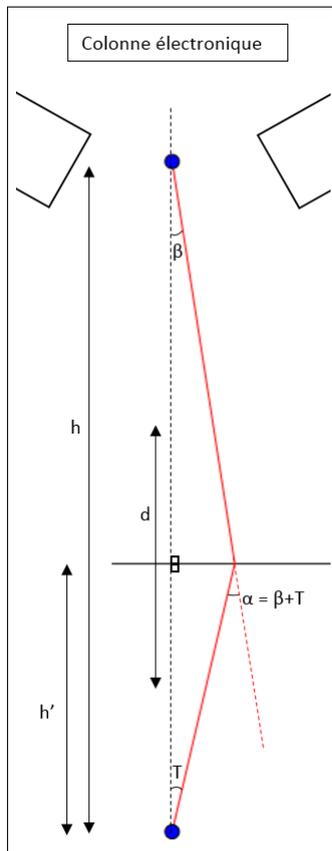
1.4 SEM à faisceau électronique inclinable

Dans un SEM classique, l’échantillon est analysé par un faisceau électronique arrivant perpendiculairement à sa surface. On réalise donc une image dite ”Topview” c’est-à-dire en vue du dessus.

Lorsqu’on veut prendre des images avec un angle de vue différent, la première technique est de changer l’inclinaison de l’échantillon par rapport au faisceau électronique sans changer le fonctionnement du SEM ; La limite de cette technique pour l’industrie est qu’il est trop compliqué d’incliner un wafer de 300mm entier à l’intérieur d’un microscope tout en gardant la calibration du faisceau électronique. La deuxième solution est d’utiliser un microscope dont la colonne électronique est inclinable, cependant la mécanique importante liée à un tel appareillage

fait qu'il est compliqué de concevoir un microscope de précision ; il s'agit donc souvent de microscope d'observation permettant de réaliser des images à forte inclinaison, mais sans mesure précise de CD.

Une nouvelle solution a donc été envisagée pour répondre aux besoins de l'industrie : un microscope dont le faisceau électronique est inclinable. La déviation du faisceau est réalisée par un champ magnétique créé par des bobines en sortie de la colonne juste au dessus du wafer comme indiqué dans la Fig 1.9.



h : Distance entre la sortie de la colonne électronique du SEM et l'échantillon.

h' : Hauteur de la déviation du faisceau (Flèche maximale).

T : Angle avec lequel l'échantillon est observé.

β : Angle de sortie de la colonne électronique. Il doit changer en fonction de l'angle T voulu et de la géométrie de manière à toujours arriver sur l'échantillon dans l'axe de la colonne, tel que :

$$\beta = \arctan \left(\frac{h' \times \tan(T)}{h - h'} \right)$$

α : Angle de déviation qui doit être appliqué sur le faisceau pour arriver sur l'échantillon avec un angle T , grâce au champ magnétique créé par des bobines.

d : Distance nécessaire au faisceau pour être dévié (non représenté ici).

FIGURE 1.9 – Schéma du principe de fonctionnement d'un SEM à faisceau inclinable.

Grâce à cet instrument on peut obtenir des images "tiltées" sans perturbations liées à des mouvements mécaniques. On peut noter que la présence d'un champ magnétique transverse au dessus du wafer entraine une déviation des électrons secondaires en sortie du matériau. De ce fait on est obligé, lorsque l'on veut réaliser des images avec le faisceau dévié, d'utiliser le mode "High Aspect Ratio", qui polarise négativement l'échantillon pour que les électrons soient rejetés fortement et puissent ainsi être captés par les détecteurs.

Lors de cette thèse j'ai utilisé un équipement de Applied Materials : le Verity 4i+ qui permet de prendre des images jusqu'à un angle de 12° , comme montré en Fig 1.10. On peut noter que contrairement aux microscopes dont l'angle est réalisé mécaniquement, on a ici accès à un angle relativement faible, ce qui va demander des efforts de reconstruction supplémentaires.

1.5 Imagerie photonique vs imagerie électronique

Nous sommes très habitués à l'imagerie photonique, c'est à dire à la formation d'une image par bombardement de photons car nous la pratiquons tous les jours par l'utilisation de nos yeux. Cette habitude est tellement grande que nous ne faisons pas forcément attention à tous les effets liés à l'imagerie photonique, car notre cerveau sait très bien analyser puis déconvoluer

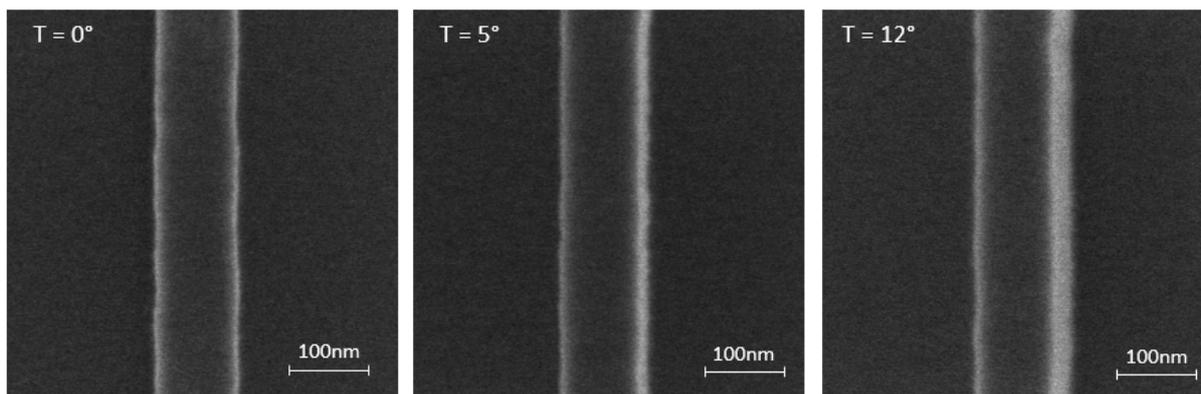


FIGURE 1.10 – Images prises grâce au Verity4i+ de AMAT. Motif en silicium gravé de 100nm de hauteur et de CD, imagé en Topview, 5° et 12°.

un signal photonique de manière à nous faire comprendre ce que l'on voit. Pourtant lorsque l'on veut qu'un ordinateur réalise le même genre d'analyse, on s'aperçoit vite qu'il y a beaucoup de paramètres à prendre en compte pour pouvoir déconvoluer tous les effets liés à la lumière et ses interactions dans un environnement complexe.

De manière analogue lorsque l'on crée une image électronique, elle est également dépendante de beaucoup d'effets électroniques. Cependant, dans ce cas il nous est beaucoup plus difficile de faire l'interprétation directe car notre cerveau n'est, non seulement pas habitué à déconvoluer des effets électroniques, mais on est de plus parfois tenté de déconvoluer une image électronique par des effets photoniques ce qui peut, dans certains cas donner une approximation correcte, mais dans d'autres conduire à une approximation fautive.

On va donc dans cette partie essayer de faire la comparaison entre les différents éléments rentrant en jeu dans la formation d'une image électronique et d'une image photonique en essayant de faire des parallèles et analogies entre les deux. On se limitera ici à la lumière visible, notamment pour éviter une partie interaction rayonnement-matière trop détaillée. On se limitera également à l'imagerie en réflexion, pour éviter le cas plus complexe du microscope électronique à transmission et ainsi nous limiter à l'étude du microscope électronique à balayage.

Dans les deux cas le principe de formation d'une image provient de l'interaction entre un rayonnement et de la matière. On va donc s'intéresser aux deux acteurs principaux : la source du rayonnement et le matériau de l'objet observé.

1.5.1 Source lumineuse, source électronique et formation de l'image

Dans le cas de la formation d'une image photonique en lumière visible la source est dite "lumineuse", c'est à dire qu'elle crée des photons dont les longueurs d'onde appartiennent (au moins en partie) au domaine visible entre 400nm et 800nm. Ces sources sont caractérisées par différentes grandeurs physiques :

- L'intensité lumineuse, en candela (cd).
- la longueur d'onde d'émission ou le spectre discret ou continu de longueurs d'onde d'émission.

Les photons peuvent avoir plusieurs processus de formation :

- Par l'échauffement d'un matériau (généralement un métal qui peut résister à de plus hautes températures) : c'est le cas de l'ampoule à incandescence. Les longueurs d'onde des photons émis par cette technique suivent la "loi des corps noirs" et forment ainsi un spectre continu. Le soleil est un cas particulier naturel de ce type d'émission.
- Par désexcitation d'un électron dans un matériau semi-conducteur ou un gaz : LED, laser, lampe à gaz. Les longueurs d'onde des photons émis par ces techniques sont discrétisées, voire uniques, car les électrons passent d'un niveau d'énergie donné à un autre ce qui crée des longueurs d'onde définies suivant la relation de Planck Einstein : $\Delta E = h.\nu = h.c/\lambda$ où ΔE est le gap d'énergie entre les deux niveaux considérés (eV), λ la longueur d'onde émise (m), h la constante de Planck (eV.s) et c la vitesse de la lumière dans le vide ($m.s^{-1}$).

On peut différencier les sources primaires de lumière qui créent la lumière, des sources de lumières secondaires qui réfléchissent une partie de la lumière qui les atteint après N réflexions (ex : La lune réfléchit la lumière du soleil ; tout objet éclairé devient une source secondaire).

Dans le cas de la formation d'une image par un microscope électronique à balayage, la source est un canon à électrons, c'est à dire un matériau conducteur dont on va extraire des électrons, soit par effet thermique, soit grâce à un champ électromagnétique, soit par la combinaison des deux. Ces électrons arrachés sont ensuite accélérés et focalisés grâce à une succession de champs électromagnétiques créés par des bobines. Une fois accélérés et utilisés en faisceau, ces électrons peuvent être assimilés à une onde électromagnétique de longueur d'onde donnée par l'équation 1.1.

La grande différence entre la formation d'une image photonique et une image électronique (dans le cas du microscope électronique à balayage) est que la source lumineuse va éclairer l'objet entièrement, puis que le capteur (que ce soit l'œil ou le capteur d'un appareil photographique) va récolter de manière continue tout un angle solide de lumière, discrétisée par la taille des pixels du capteur ou par la densité des cônes de l'œil. On peut également noter qu'il existe trois types de cônes, sensibles à différentes longueurs d'onde, ce qui nous permet de discerner la couleur des objets par la proportion de signal réparti entre les couleurs : Rouge, vert, bleu (RGB). Sur un capteur CCD ou CMOS un filtre de couleur va être rajouté devant les cellules photosensibles pour réaliser cette fonction.

Dans le cas de l'image électronique le faisceau incident va être focalisé en un point, envoyer un certain nombre d'électrons, puis se déplacer vers le point suivant et répéter l'opération jusqu'à avoir couvert toute la surface à analyser (Fig 1.11). En chaque point les électrons vont être récupérés via un capteur d'électrons. Bien que l'imagerie électronique ne soit pas sensible à la longueur d'onde, le capteur d'électron est muni d'une grille (appelée collecteur) qui va pouvoir être polarisée positivement ou négativement, de manière à attirer les électrons de faible énergie (comme les SE) ou au contraire les repousser, favorisant alors la captation des électrons

de plus haute énergie (notamment les BSE). En sortie du capteur d'électrons, un ordinateur va simplement construire une cartographie en niveau de gris de la zone observée en fonction du nombre d'électrons récupérés en chaque point. C'est l'origine d'une image provenant d'un microscope électronique à balayage.

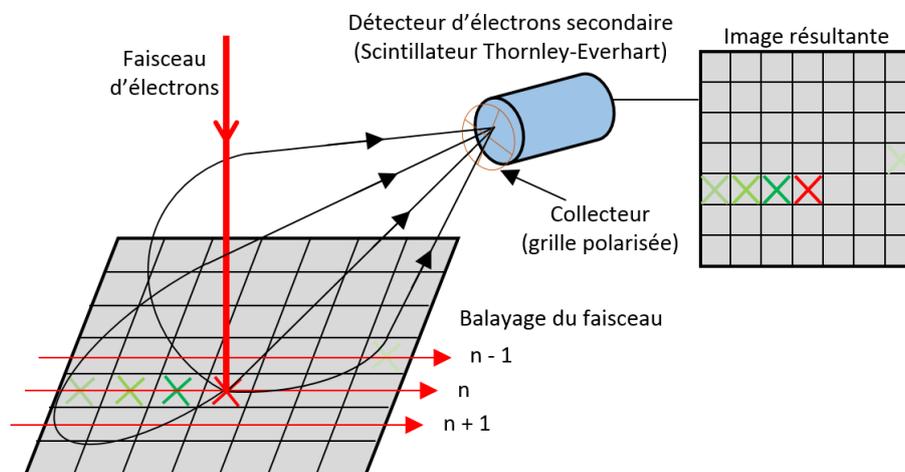


FIGURE 1.11 – Formation d'une image SEM par électrons secondaires. Les électrons secondaires ayant une énergie faible en sortie du matériau sont attirés vers le capteur grâce au champ électrique de quelques centaines de volts créé par le collecteur.

Il y a donc une sorte de symétrie "miroir" entre la formation d'une image photonique et une image électronique de microscope électronique à balayage : dans le premier cas on envoie un rayonnement sur toute la surface d'un objet puis on récupère le rayonnement réfléchi dans un petit angle solide et c'est la somme de différents angles solides qui permet la formation de l'image, dans le deuxième cas on envoie le rayonnement sur un tout petit angle solide, puis on récupère tout le rayonnement sortant et c'est l'envoi répété de rayonnements en différents points qui permet la formation de l'image.

1.5.2 Matériau de l'objet observé

Le matériau de l'objet observé va jouer un rôle très important dans la formation de l'image car ce sont les interactions entre celui-ci et les photons qui l'atteignent qui vont créer l'image obtenue.

On va pouvoir séparer les effets en deux catégories :

- Les effets liés à l'état de surface du matériau.

Si le matériau est très rugueux en surface, les rayons lumineux vont être diffusés dans des directions aléatoires en fonction de leur arrivée sur le matériau. On parlera alors de réflexion diffuse (Fig 1.12).

A l'inverse si le matériau est très lisse en surface, les rayons lumineux arrivant avec un angle d'incidence α par rapport à la normale au plan, repartiront tous avec un angle de réflexion α par rapport à cette même normale, on parle alors de réflexion spéculaire.

- Les effets liés à la composition chimique du matériau.

Certaines longueurs d'ondes vont être préférentiellement absorbées en raison de l'interaction avec certains composés chimiques. C'est l'explication physique de la couleur d'un objet (dans le cas d'une illumination par lumière blanche) : s'il est composé d'un matériau qui absorbe préférentiellement les longueurs d'onde associées à la couleur rouge alors les longueurs d'onde associées à la couleur bleu seront préférentiellement

réfléchies et l'objet nous apparaîtra comme bleu.

De plus si un matériau est diélectrique, alors l'onde évanescente va s'étendre sur une certaine distance dans le matériau alors que s'il est métallique la profondeur de peau sera très faible. De ce fait un matériau diélectrique aura toujours une partie de réflexion diffuse importante, alors que la réflexion d'un métal pourra être parfaitement spéculaire (cas du miroir).

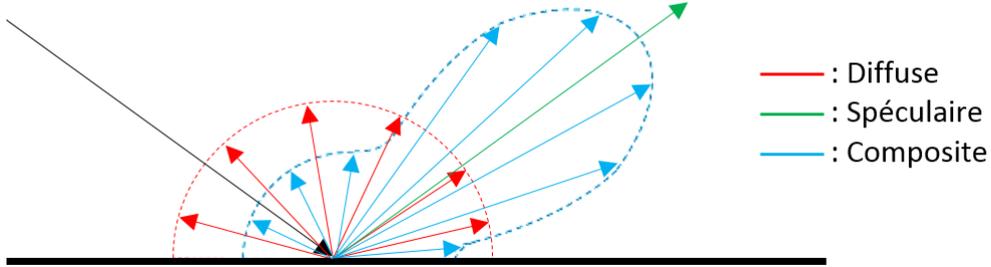


FIGURE 1.12 – Les différents types de réflexions : probabilité angulaire de réflexion des rayons lumineux incidents.

Pour étudier la quantité de lumière réfléchiée dans une direction en fonction de la lumière incidente en un point de façon mathématique on peut utiliser la fonction de distribution de la réflectivité bidirectionnelle (BRDF en anglais) qui dépend de l'angle et l'azimut d'incidence de la source lumineuse (ω_i) et de l'angle d'incidence et l'azimut du capteur (ω_o) tel que :

$$f_r(\omega_i, \omega_o) = \frac{dL_o(\omega_o)}{dE(\omega_i)} = \frac{dL_o(\omega_o)}{L_i(\omega_i) \cdot \cos(\theta_i) \cdot d\omega_i} \quad (1.2)$$

avec $L(\omega)$ la radiance (unité de puissance par aire perpendiculaire à la direction de la lumière par unité d'angle solide ($W/m^2/sr$)) et $E(\omega)$ l'irradiance (unité de puissance par unité de surface (W/m^2))

Pour être physiquement cohérente cette fonction doit respecter trois propriétés :

- Comme c'est une probabilité :

$$f_r(\omega_i, \omega_o) \geq 0 \quad (1.3)$$

- Le principe de réciprocité de Stokes-Helmholtz basé sur la linéarité des phénomènes physiques mis en jeu, implique que si on échange la position de la source et du capteur la même quantité de lumière sera captée, tel que :

$$f_r(\omega_i \rightarrow \omega_o) = f_r(\omega_o \rightarrow \omega_i) \quad (1.4)$$

- La conservation de l'énergie implique que l'énergie réfléchiée est égale à l'énergie incidente (voir inférieure si une partie est absorbée), telle que :

$$\int_{2\pi} f_r(\omega_i, \omega_o) \cdot \cos \theta_i \cdot d\omega_i \leq 1 \quad (1.5)$$

Dans le cas de l'imagerie électronique le matériau de l'objet observé est également important ; En effet les électrons vont interagir avec le matériau en fonction du numéro atomique des atomes rencontrés, de sa conductivité ainsi que de sa topographie de surface.

Le numéro atomique de l'échantillon observé va avoir deux conséquences sur l'intensité électronique collectée ; Premièrement, il est plus difficile pour les électrons de pénétrer dans

un matériau dont le numéro atomique est élevé et donc plus l'information donnée par l'image (à accélération constante) sera superficielle. Deuxièmement, plus le numéro atomique est important plus les électrons ont de chance d'être rétrodiffusés, c'est à dire de subir un changement de trajectoire élastique dû au noyau.

Le nombre d'électrons récupérés dépend également de la topographie de l'échantillon observé. En effet, plus l'angle solide extérieur est important, plus le nombre d'électrons qui arrivent à s'échapper du matériau est grand (Fig 1.13).

D'autre part, plus l'angle entre le rayonnement incident et la surface de l'échantillon observé s'éloigne de la normalité plus les électrons peuvent s'échapper facilement. On peut donc relier le nombre d'électrons sortant du matériau comme une fonction $1/\cos(\text{pente})$ en première approximation, lorsque l'on se place dans l'hypothèse de surface infinie à une certaine pente (des modèles plus complets ont été proposés [17]).

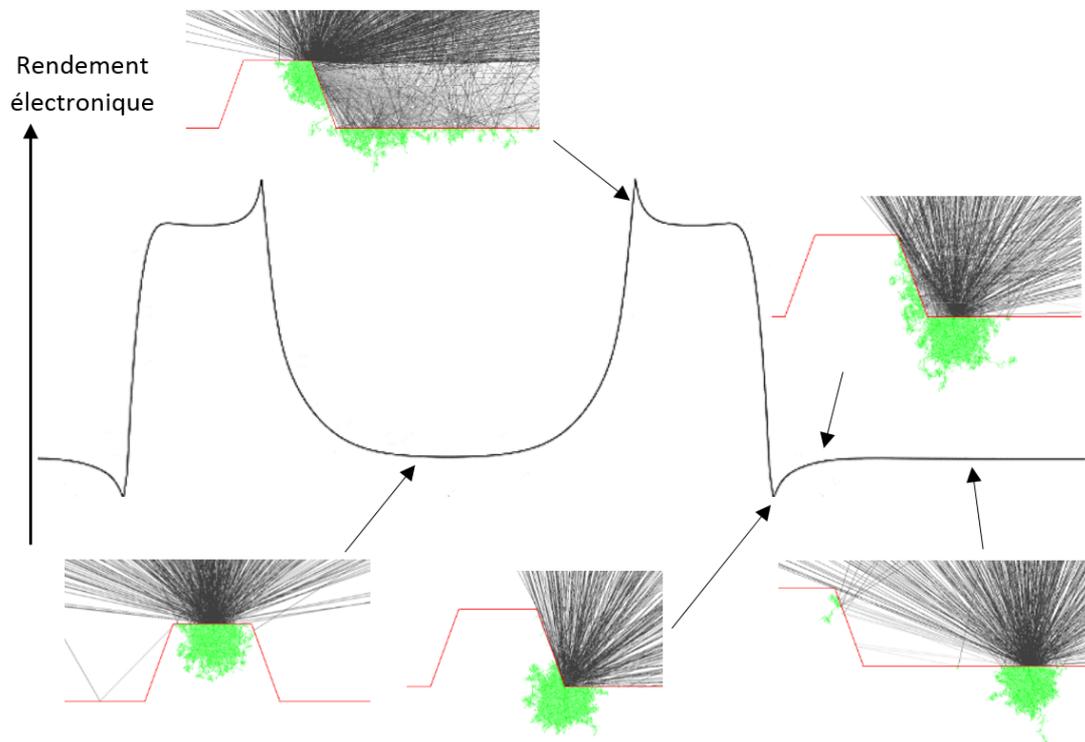


FIGURE 1.13 – Rendement de sortie des électrons en fonction de la topographie locale d'un motif. Les images représentent une vue en coupe du motif observé (en rouge) et la trajectoire des électrons dans le matériau (en vert) et en dehors du matériau (en noir), en fonction de la position du faisceau électronique incident sur le motif.

De plus, dans un matériau conducteur, les électrons vont pouvoir se déplacer plus librement permettant de libérer facilement les charges alors que dans un matériau diélectrique les charges vont avoir plus de difficultés à s'échapper créant des phénomènes de charging sur les images SEM, c'est à dire un phénomène de distorsion de l'image dû aux interactions électrostatiques entre les électrons excédentaires bloqués dans le matériau et les nouveaux électrons du faisceau électronique.

Le but du travail qui va être présenté étant d'obtenir des informations topographiques à partir d'images SEM, on va voir dans le chapitre suivant les différentes possibilités d'utilisation des méthodes de métrologie présentées pour la reconstruction tri-dimensionnelle.

Chapitre 2

Etat de l'art : Reconstruction tridimensionnelle

Le but de ce chapitre est de présenter les différentes techniques de reconstruction tridimensionnelle. On expliquera tout d'abord le principe général de chaque technique puis on verra comment adapter ces techniques à la problématique spécifique des reconstructions topographiques à partir d'images de microscopie électronique. Enfin on présentera les méthodes de simulation d'images SEM par simulation physique ou modèle compact. Grâce à ces méthodes on étudiera les possibilités de résolution du problème inverse, c'est à dire de détermination de la topographie du motif à partir des images SEM.

2.1 Introduction

Le domaine de la reconstruction tridimensionnelle consiste à retrouver la topographie d'un objet. On peut distinguer différents types de reconstructions dans l'espace.

- Les méthodes qui utilisent directement la mesure d'une grandeur physique.
- Les méthodes basées sur une analyse d'images.

Dans les méthodes physiques on peut retrouver :

- La reconstruction par laser qui va sonder une surface grâce à de la lumière permettant ainsi une mesure de distance locale, soit par l'analyse du temps de trajet aller-retour de la lumière, soit par l'analyse de la différence de phase de l'onde.
- La déflectométrie, méthode utilisée pour la reconstruction mais également la détection de défauts sur des pièces mécaniques dont la courbure évolue lentement et dont la taille est importante (aile d'avion, carrosserie de voiture...) par l'étude de la déformation de la projection d'une lumière structurée sur un objet.
- La reconstruction par profilométrie qui consiste à balayer mécaniquement une surface avec une sonde pointue pour en déterminer sa topographie.
- L'AFM dont on a parlé dans la partie "Métrologie 3D".
- Le microscope à effet tunnel (STM : Scanning Tunneling Microscope) qui suit la surface grâce au courant traversant la barrière de potentiel du vide par effet tunnel.

Les méthodes de reconstruction par analyse d'images en deux dimensions consistent à retrouver l'information tridimensionnelle des objets, perdue lors de l'acquisition d'images, c'est à dire lors de la projection de la scène sur un plan.

Il existe différentes méthodes de reconstruction d'un objet dans les trois dimensions de l'espace à partir de la projection bidimensionnelle de ce même objet sur une ou plusieurs images. Nous allons détailler deux de ces techniques, l'une géométrique et l'autre basée sur la réflexion de la lumière sur un objet orienté. On essaiera ensuite de voir comment adapter ces techniques pour

s'en servir à partir d'images de microscopie électronique, ce qui nous amènera à une technique spécifique à l'imagerie électronique : la résolution du problème inverse de la formation d'une image SEM par la résolution des phénomènes physiques ou par modèle compact.

2.2 Reconstruction géométrique par appariement de points et triangulation

La reconstruction par appariement de points (aussi appelée Multi-View Stereo MVS) est une méthode de reconstruction géométrique basée sur l'étude de deux images (ou plus) d'une même scène prise à des angles de vue différents.

Pour pouvoir réaliser une reconstruction correcte il est tout d'abord important de connaître avec précision les caractéristiques des capteurs ainsi que leurs positions et orientations. De plus, pour qu'un point de l'espace puisse être retrouvé, il doit forcément apparaître sur au moins deux images.

La seconde étape d'une reconstruction géométrique est de savoir quels éléments de l'image 1 correspondent aux mêmes éléments de l'image 2. En fonction du type d'image et de la finalité de la reconstruction on va pouvoir essayer d'appairer, soit des points particuliers, soit des segments, voire même des régions [18]. La correspondance entre les images peut être réalisée à l'aide de différentes techniques.

La méthode classique consiste à analyser les images grâce à un produit de corrélation pour repérer les zones similaires ; plus le produit de corrélation est grand, plus la ressemblance est grande. Il existe également des techniques consistant à réduire itérativement la résolution d'une image, ce qui permet d'isoler les zones d'intérêt. Enfin la reconnaissance par machine learning est de plus en plus utilisée pour ce genre de problèmes complexes.

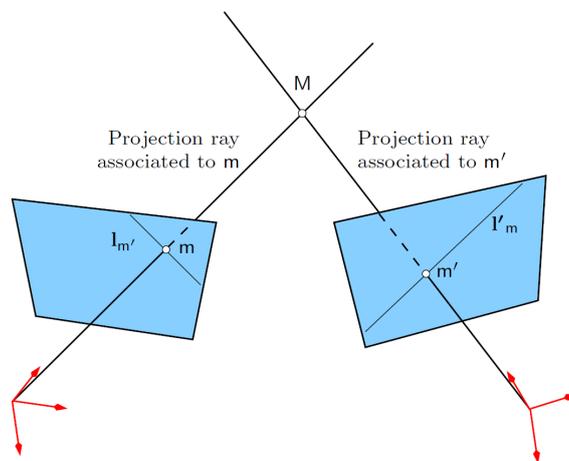


FIGURE 2.1 – Géométrie épipolaire : La ligne épipolaire $l_{m'}$ (respectivement l_m) est l'image de la droite de projection associée à m' (respectivement m), où m et m' sont les représentations du même point M dans chacune des images [19].

L'appariement de point est considéré comme l'étape la plus délicate et la plus critique pour obtenir une reconstruction correcte. Pour améliorer et accélérer cette étape on utilise notamment les lignes épipolaires pour la recherche de points équivalents [19] [20]. Une ligne épipolaire correspond à la projection dans l'image 1 de la droite passant par le point à reconstruire et sa projection dans l'image 2 (Fig 2.1). Pour déterminer la matrice fondamentale permettant de passer en géométrie épipolaire il faut repérer au minimum huit points caractéristiques facilement discernables dans les différentes images (comme un coin

géométrique). Une fois la matrice fondamentale trouvée, on peut l'utiliser, d'une part pour retrouver les positions et calibrations des caméras, mais également pour la recherche de points à appairer moins évidents que ceux choisis pour créer la matrice ; en effet au lieu de chercher dans toute l'image 2 la correspondance à un point de l'image 1, on pourra simplement le chercher sur la droite épipolaire, ce qui réduit fortement le temps de calcul et les risques d'erreur.

Dans le cas où l'objet à reconstruire est une forme peu anguleuse et de couleur uniforme, il est très difficile de repérer des points particuliers à appairer. On peut alors projeter une lumière structurée sur l'objet qui sera déformée en fonction de la topographie de l'objet [19] et ainsi créer artificiellement des points de repère pour faciliter le repérage des points identiques sur les différentes images.

Une fois ces points caractéristiques repérés, la reconstruction peut être résolue par triangulation, de manière triviale en l'absence de bruit :

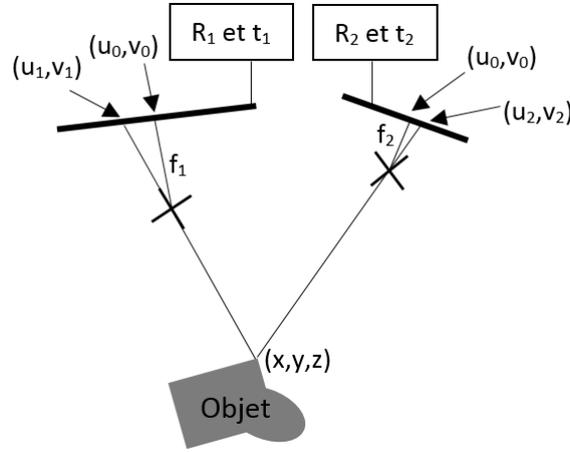


FIGURE 2.2 – Principe de triangulation par la prise de deux images à angle de vue différents, représenté en 2D pour l'espace et 1D pour les images

Si on note (voir Figure 2.2) :

- (x, y, z) les coordonnées du point de la scène dans le repère absolu,
- (u_1, v_1) et (u_2, v_2) les coordonnées du point dans l'image 1 et 2,
- (X_1, Y_1, Z_1) et (X_2, Y_2, Z_2) les coordonnées du point dans le repère de la caméra 1 et 2,
- K_1 et K_2 la matrice de calibration interne de la caméra 1 et 2, correspondant à la fonction

de transfert de l'espace physique au repère de la caméra, tel que $K = \begin{pmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix}$,

où f est la distance focale et (u_0, v_0) les coordonnées du point correspondant à la position de l'intersection de l'axe optique et du plan image. Si on suppose les caméras parfaites, ce point est le centre de l'image.

- R_1 et R_2 la matrice 3x3 d'orientation de la caméra 1 et 2 dans le repère absolu,
- \vec{t}_1 et \vec{t}_2 le vecteur de position de la caméra 1 et 2 dans le repère absolu.

Par construction on a :

$$K_1 \times \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} = \begin{pmatrix} Z_1 u_1 \\ Z_1 v_1 \\ Z_1 \end{pmatrix}, \quad K_2 \times \begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = \begin{pmatrix} Z_2 u_2 \\ Z_2 v_2 \\ Z_2 \end{pmatrix} \quad (2.1)$$

Grâce à chacune de ces équations on va trouver une droite sur laquelle le point (x,y,z) peut se trouver, il ne reste plus alors qu'à trouver le point d'intersection de ces deux droites en résolvant le système suivant :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_1 \times \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} + t\vec{1} = R_2 \times \begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} + t\vec{2} \quad (2.2)$$

Ce qui permet d'obtenir les coordonnées du point observé.

Dans un cas réel les deux droites visant le même point sur les deux images ne seront pas sécantes à cause des incertitudes et de la non perfection des instruments utilisés, on utilisera alors des méthodes plus ou moins complexes de minimisation de l'erreur (point milieu, moindres carrés sur plusieurs couples de points...) [21] [22].

De manière générale plus la distance entre les deux prises de vue (et donc la différence d'angle de vue si on pointe vers le même point) est grande plus la reconstruction sera précise. En revanche il sera plus compliqué de repérer les points équivalents sur des images trop différentes, il faut donc trouver un compromis.

Ces méthodes géométriques peuvent être utilisées de manière similaire sur des images de microscopie électronique à balayage à condition qu'il soit possible de détecter les couples de points d'appariement entre les images. Ces images doivent de ce fait être peu bruitées et peu affectées par des artefacts électroniques, ce qui implique souvent que ces images soient prises sur des objets dont les dimensions ne sont pas trop proches de la limite de résolution du SEM. Différents logiciels commerciaux sont capables de réaliser ce genre de reconstruction (Mex, Mountains Map, Scandium height...). Les capacités de certains de ces logiciels ont été mises à l'épreuve sur des motifs dont les tailles correspondent aux besoins de la microélectronique, à partir d'images simulées par le logiciel de simulation physique JMONSEL [23] [24]. Cette étude a montré que les logiciels de reconstruction ont besoin de beaucoup de détails fins pour pouvoir faire une reconstruction stéréoscopique optimale. En effet sur des motifs lisses l'erreur d'estimation de la hauteur est de 50% malgré la très grande qualité des images (car provenant de simulation). On a vu précédemment que le point critique de la reconstruction stéréoscopique est la recherche de points de correspondance entre les deux images et on voit ici le résultat d'un manque de point de repère lorsque l'on passe par une méthode classique de reconstruction stéréoscopique.

Les auteurs soulignent également que cette méthode de reconstruction manque de robustesse du fait de sa sensibilité à beaucoup de facteurs comme : les erreurs de placement mécanique de l'échantillon, la calibration des instruments de mesure, la vibration, translation ou les effets de charge durant la prise d'image. Enfin, un point important, la réponse électronique d'une même zone prise à des angles de vues différents peut ne pas être similaire due à la sensibilité à l'angle d'incidence des électrons primaires sur les interactions électron-matière, origine de la formation de l'image SEM.

Cette méthode classique de triangulation à partir d'images SEM peut se révéler particulièrement adaptée pour la reconstruction topographique d'échantillons biologiques car leur taille (typiquement au dessus de la dizaine de μm) permet aux images SEM de ne pas être trop perturbées par des effets de l'imagerie électronique et l'abondance de détails permet un repérage des points d'appariement facile, conduisant à une reconstruction efficace [20]. La Fig 2.3 tirée de [25] montre un exemple de reconstruction faite avec des logiciels commerciaux, à partir de 40 images SEM prises à 9° d'intervalle, dont la taille de pixel est de 31nm sur une zone (FOV) de $65 \times 65 \mu\text{m}^2$ environ (a et b). Sur les images c et d (FOV $\approx 15 \times 15 \mu\text{m}^2$) on voit que la reconstruction devient difficile en dessous du μm (Attention à ne pas être trompé par la colorisation de la reconstruction qui ne reflète pas forcément un résultat topographique, objet

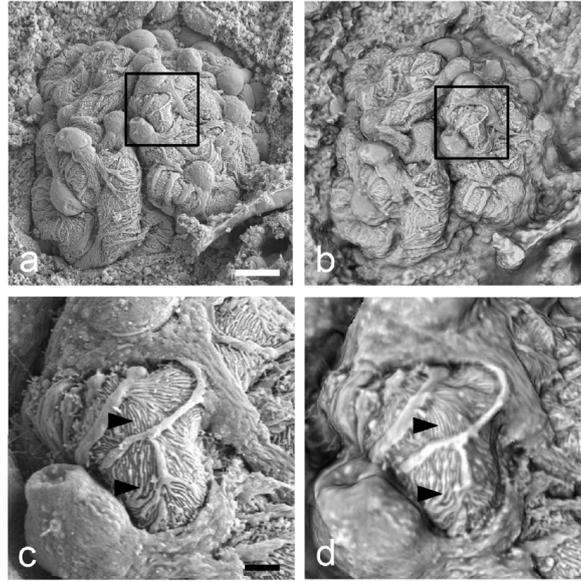


FIGURE 2.3 – Comparaison entre les images SEM (a et c) et la reconstruction (b et d). [25]

de l'étude). L'auteur précise que plusieurs étapes ont besoin d'être optimisées manuellement : l'ajustement de la luminosité et du contraste des images, certaines images jugées trop floues sont supprimées, l'aide au choix de points d'appariement sur les logiciels le permettant. Si on veut une grande précision, on voit donc que l'on va être confronté à des problèmes de reproductibilité et de limitation dimensionnelle.

Dans [20], l'auteur explique également les problèmes liés à la profondeur de champ limitée d'un SEM et à la difficulté de conserver la mise au point (focus) lors de la rotation de l'objet observé.

Dans le cas de la reconstruction topographique pour l'industrie des semi-conducteurs, la connaissance de certains paramètres d'entrée peut faciliter le travail de reconstruction. Par exemple lorsque l'on prend une image SEM sur un wafer, l'emplacement de la prise d'image correspond à une zone choisie du plan du circuit (layout) ; on sait donc quel type de structure doit se trouver à cet emplacement. On peut également savoir quel est le type de matériau du motif ou des couches en dessous du motif, ce qui peut par exemple permettre de réaliser une reconstruction adaptée à la réponse des interactions électron-matière du matériau concerné. De ce fait il n'est pas approprié de faire une reconstruction avec comme unique paramètre d'entrée l'image brute. A la place on peut partir de notre connaissance de la forme des motifs (même si la reconnaissance est faite par un algorithme) puis chercher des points bien particuliers pour nous permettre de reconstruire l'image en trois dimensions avec plus de facilité et de précision.

En revanche le résultat attendu sera plus exigeant qu'une reconstruction donnant la topographie générale de la zone d'observation. En effet on veut obtenir des résultats quantitatifs, comme par exemple être capable de savoir si la hauteur du motif est bien de 100nm comme prévu par les graveurs, ou s'il a été sur-gravé et qu'il ne fait plus que 80nm, menant éventuellement à des problèmes dans la suite du procédé. On va ainsi chercher des grandeurs d'intérêt spécifiques aux besoins de l'industrie de la microélectronique et résumées dans la Figure 1 (Notations page 3). Ces grandeurs seront trouvées notamment grâce à l'étude du profil de niveau de gris de l'image SEM ainsi qu'à l'étude de sa dérivée [26].

Cette méthode est de plus en plus utilisée notamment grâce à l'utilisation de SEM dont le faisceau électronique peut être incliné, ce qui permet de réaliser des mesures précises de la taille du flanc (EW). En effet grâce à des équations géométriques simples, il est possible de remonter aux informations de hauteur et de SWA d'un motif grâce à la mesure de l'EW d'un motif imagé à deux angles de tilt différents (Figure 2.4).

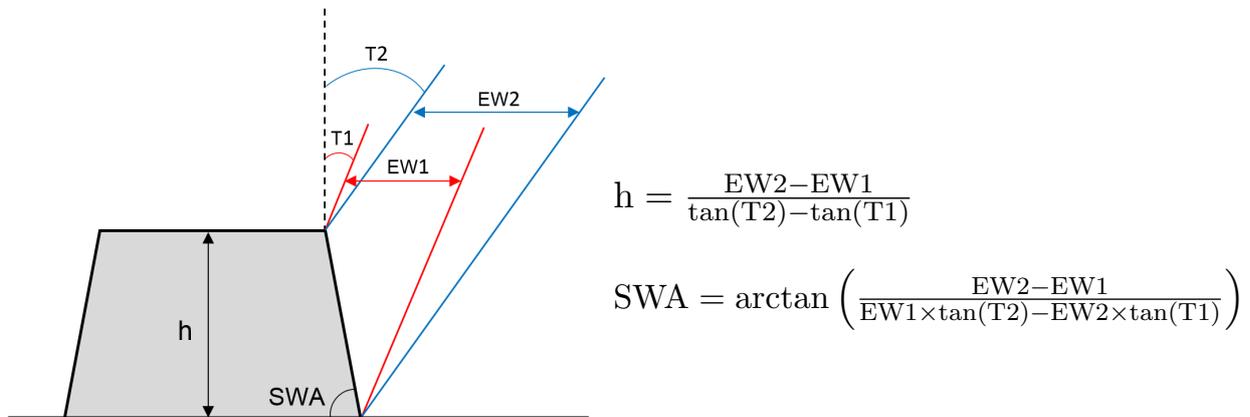


FIGURE 2.4 – Méthode de calcul de la hauteur et du SWA d'un motif à partir de la mesure des deux EW provenant d'images prises à deux angles de vue différents T1 et T2. [27]

Cette méthode a été utilisée sur des motifs, principalement de type ligne, soit en silicium gravé [27] soit en résine [28] [29], mais les résultats sont souvent décevants et demandent des corrections spécifiques à l'application ce qui rend le procédé peu reproductible.

La littérature s'accorde à l'unanimité sur l'amélioration de la précision de reconstruction tridimensionnelle avec l'augmentation du parallaxe (Figure 1 dans [30]), c'est à dire la différence d'angle entre les deux prises de vue, ainsi que la perte critique de précision lorsque la taille de l'EW diminue jusqu'à être de l'ordre de grandeur de la taille du diamètre du faisceau électronique, voire inférieure. De plus certaines topographies sont très difficiles à reconstruire comme : les CR top et CR bot (parfois appelé "footing"), les pentes négatives (dévers local, flanc arqué (bowing) caractéristiques de certains procédés microélectronique, pente globale ...) [11] [30] [28].

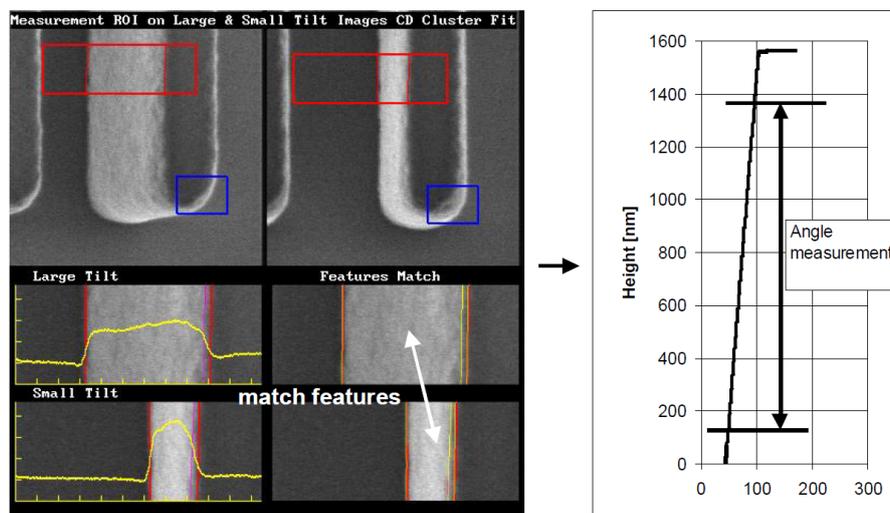


FIGURE 2.5 – Reconstruction d'un motif via deux images de tilts différents. [28]

On peut tout de même noter des résultats intéressants sur des motifs de $1,5\mu\text{m}$ de hauteur et environ 88° de SWA dans la publication [28], qui montre également des reconstructions sur de la résine 193nm. Il est précisé que la reconstruction n'est pas affectée par des problèmes de rétractation de la résine (shrinkage) lors de la mesure. (Figure 2.5)

2.3 Reconstruction photométrique

La photométrie est la mesure d'un rayonnement lumineux tel qu'il est perçu par l'œil humain, c'est à dire dans le spectre du visible et pondéré par la sensibilité de l'œil humain en fonction de la longueur d'onde. La reconstruction photométrique (aussi appelée "shape from shading") est une technique de visualisation par ordinateur permettant de simuler la réponse des normales des surfaces d'un objet en observant celui-ci sous différentes conditions d'éclairage ou bien sous une seule condition d'éclairage mais à des angles de vue différents. Elle est basée sur le fait que la quantité de lumière réfléchiée par une surface dépend de l'orientation de la surface par rapport à la source lumineuse et à l'observateur (pour un matériau donné). Cette technique est de fait très dépendante du matériau sur lequel la lumière est réfléchiée; on doit donc distinguer deux cas :

- Le cas dit "Lambertien", où l'on considère que la réflexion du matériau est purement diffuse, c'est à dire que l'intensité lumineuse est réfléchiée dans toutes les directions de l'espace. Cette hypothèse implique que si les positions de la source lumineuse et/ou du capteur restent sur le même axe de rotation autour de la normale de la surface alors l'intensité détectée sera toujours la même. Dans ce cas seuls les angles solides (source vers surface et surface vers capteur) ont de l'importance.
- Le cas contraire dit "non Lambertien", traite tous les cas dans lesquels une partie de la réflexion est spéculaire, c'est à dire qu'elle est préférentiellement réfléchiée suivant un angle de réflexion égal à l'angle d'incidence (par rapport à la normale à la surface).

La reconstruction photométrique consiste donc à réaliser le problème inverse de la BRDF (partie 1.5.2) ou d'une autre fonction similaire, à partir d'une hypothèse sur les caractéristiques du matériau dans le but de retrouver la topographie.

Il est possible de réaliser une reconstitution à partir d'une seule image, cependant l'une des principales difficultés de cette méthode est l'unicité de son résultat; en effet plusieurs études ont montré qu'il était impossible de trouver une unique topographie à partir d'une seule image 2D en niveau de gris. Pour pouvoir résoudre le problème il faut alors rajouter des contraintes (conditions aux limites, continuité, convexité ...) qui sont souvent liées à une connaissance à priori du résultat [31] [32].

On peut d'ailleurs noter que nous ne sommes pas toujours capables de savoir exactement ce que l'on voit à partir d'une unique image. On peut prendre pour exemple le cas du bas-relief : tant qu'on le regarde selon un seul angle on a une impression de profondeur. Pour s'apercevoir de l'erreur de reconstruction que notre cerveau a commis il faut tourner autour de la sculpture, c'est à dire analyser des images à des angles de vue différents [33].

La reconstruction photométrique peut également être utilisée sur des images dont le contraste est électronique, en effet le nombre d'électrons collectés par un capteur d'électrons secondaires va être sensible à l'orientation de la surface par rapport au capteur. Dans une publication de Drzazga [34], s'appuyant sur les équations de Slowko [35], une équation est proposée pour décrire l'altitude d'une surface en tout point à partir de l'intensité électronique collectée par quatre capteurs placés comme indiqué dans la Figure 2.6.

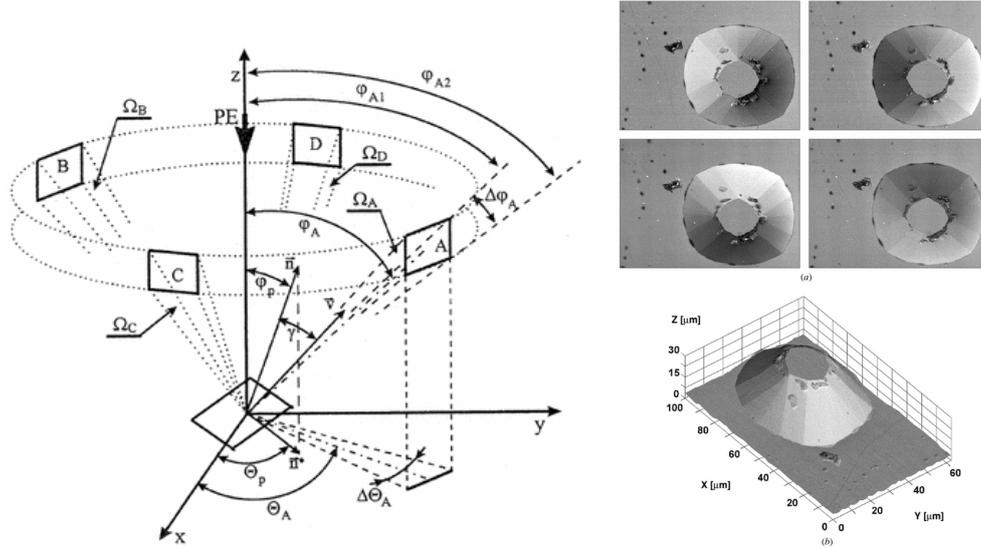


FIGURE 2.6 – Gauche : Architecture du SEM dit ”quatre cadrants” ; droite : Image résultante pour chaque capteur puis reconstruction tridimensionnelle. [34]

L’altitude en tout point est alors donnée par :

$$z(x, y) = a_x \int_x \left[\frac{I_A - I_B}{I_A + I_B} \right] dx + a_y \int_y \left[\frac{I_C - I_D}{I_C + I_D} \right] dy + C_0 \quad (2.3)$$

où I_X est l’intensité perçue par le capteur X, a_x et a_y sont des coefficients dépendant principalement de la géométrie des capteurs et C_0 est l’altitude initiale.

Cette méthode de prise d’image a également été utilisée pour la reconstruction de motifs dont les dimensions correspondent aux besoins de l’industrie de la microélectronique [36] [37].

Dans [36], Fukaya nous présente d’abord une méthode pour déterminer la hauteur, puis une méthode permettant de déterminer le SWA connaissant la hauteur. Dans son travail il utilise beaucoup les dimensions récupérées sur les profils horizontaux de niveau de gris, méthode que je vais également largement employer (Fig 2.7). La méthode de mesure de hauteur est peu détaillée et semble peu robuste car dépendante de beaucoup de paramètres autres que la hauteur (SWA, corner rounding, rugosité de ligne, accélération des électrons, matériau ...). Cela dit, le principe d’utilisation de la zone d’ombrage en bas d’un motif comme zone de recherche sur l’information de la hauteur semble justifié, il faudrait donc déconvoluer le résultat en isolant la part de responsabilité de la hauteur ce qui représente un travail conséquent mais pouvant possiblement donner un résultat intéressant.

L’auteur présente ensuite l’estimation du SWA connaissant la hauteur. Son approche est une mesure de la taille de l’EW. La difficulté de la mesure d’EW consiste à savoir où se situe le pied et le sommet du flanc. Pour cela il propose une méthode qui consiste à soustraire le profil d’une image provenant de détecteurs positionnés de part et d’autre d’un motif, de manière à obtenir une transition plus franche. Cette méthode reprend l’idée de l’équation 2.3 ($I_A - I_B$ avec A et B des capteurs diamétralement opposés par rapport à l’échantillon). Cette analyse donne des résultats intéressants pour la mesure de SWA mais a cependant plusieurs limites (annoncées par l’auteur) :

- Lorsque le SWA tend vers la verticale (90°) la mesure perd fortement en précision, car

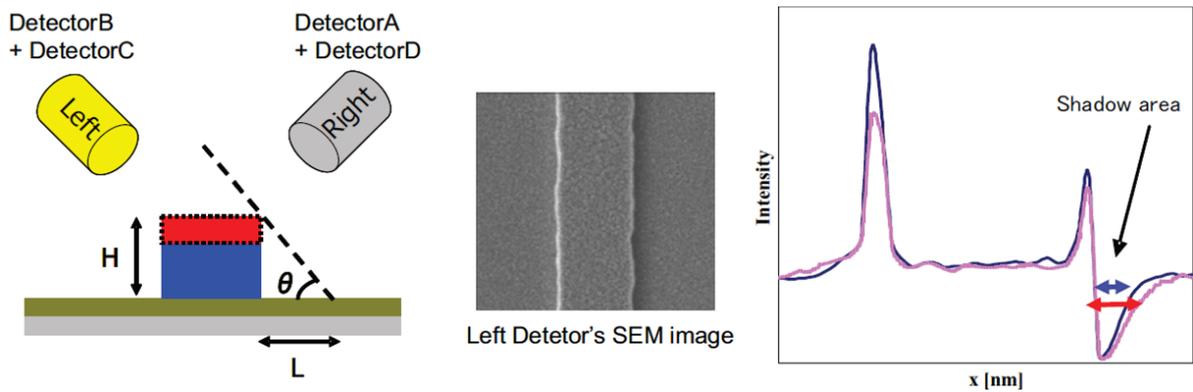


FIGURE 2.7 – Méthode de récupération de la hauteur d’un motif à partir d’une image : $H = L \times \tan(\theta)$. Assimilation de la longueur expérimentale de la zone d’ombrage à la longueur L . [36]

la taille de l’EW devient faible.

- L’influence des congés (CR) n’est pas ou peu prise en compte, menant à des erreurs de reconstruction.
- La rugosité de ligne influence la mesure de l’EW (pour cette dernière il propose une correction par détection de contours).

Des résultats de cette méthode sont également présentés dans la publication [37], qui précise que la hauteur a été calibrée grâce à de l’AFM. L’objectif de cette publication est principalement de démontrer que la reconstruction par images SEM est plus pratique, plus rapide voire plus précise sur certains aspects que l’AFM (pas de problème lié à la taille de la pointe), comme montré en Fig 2.8.

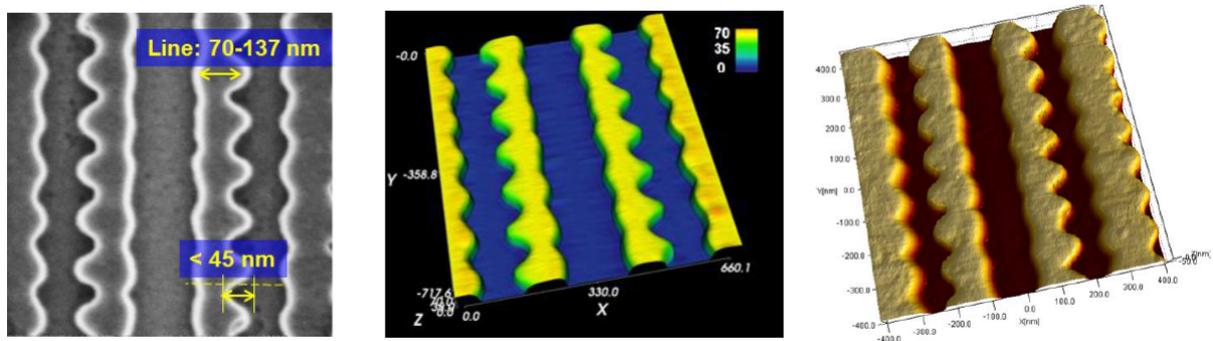


FIGURE 2.8 – Comparaison sur un motif ”peigne” d’une image SEM (gauche), du résultat AFM (droite) et de la reconstruction par la méthode quatre détecteurs (milieu). [37]

2.4 Méthode de reconstruction hybride

On a vu dans les deux parties précédentes que l’on pouvait classer les reconstructions en deux grandes catégories : géométrique et photométrique. Cependant, il est possible de décliner ces deux familles de différentes manières, voire d’utiliser ces techniques de manière couplées pour en utiliser les forces respectives [38]. Par exemple [39] utilise une méthode géométrique pour déterminer des contraintes suffisantes pour ensuite utiliser la reconstruction photométrique.

Dans la publication [40], Wu nous rappelle que la faiblesse de la reconstruction géométrique provient de la difficulté à trouver des points de correspondance entre les images, obligeant la reconstruction à interpoler entre les points caractéristiques. C'est pourquoi il propose de réaliser une reconstruction géométrique pour obtenir une reconstruction de la forme générale de l'objet (basses fréquences), puis d'utiliser une méthode photométrique précontrainte avec la forme générale trouvée de manière à extraire les détails de l'objet (hautes fréquences). Il propose ainsi une méthode hybride dont il nous présente des résultats de précision comparable à une reconstruction par méthode laser (Fig 2.9)

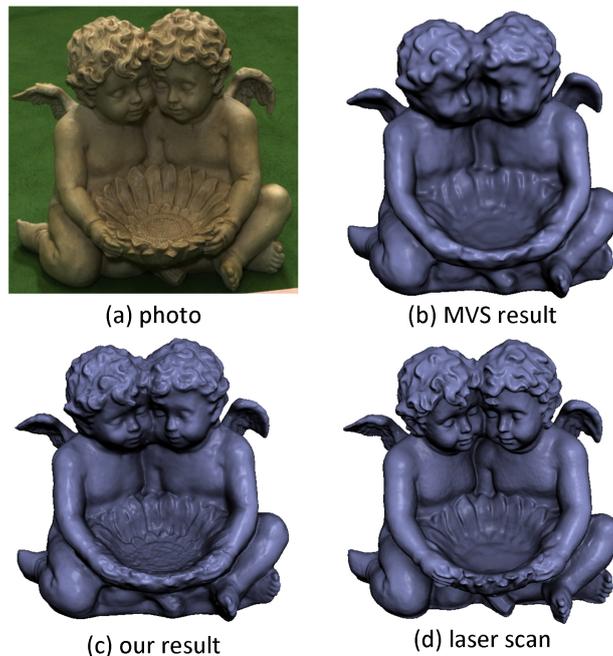


FIGURE 2.9 – Résultat de la reconstruction hybride [40] présentée (c) comparé à une méthode laser (d) et à une méthode de reconstruction géométrique seule (b).

Les méthodes hybrides sont également utilisables pour les besoins de la microélectronique. On peut noter les bons résultats obtenus par Morokuma [30], qui commence par une méthode photométrique puis affine le résultat avec une méthode géométrique. Il commence par déterminer le SWA des motifs en utilisant une version améliorée de l'équation de Lambert ($I = 1/\cos(\text{pente})$) pour la pente et une intégration sur la tangente de la pente sur la distance de l'EW pour la hauteur comme suit :

$$\text{SWA}(x) = \cos^{-1} \left(\frac{kI_0}{I_S(x)} \right)^{1/n} \quad (2.4)$$

$$h(x) = \int_0^x \tan(\text{SWA}(x)) dx \quad (2.5)$$

avec $I_S(x)$ le niveau de gris de l'image en x , I_0 le niveau de gris d'une surface plate. k et n sont des coefficients qui seront estimés grâce à une librairie de pentes pré-calculées en fonction de divers profils. Ces coefficients sont certainement le point faible de cette technique : Y a-t-il une unique topographie par couple (n,k) ? Est-ce efficace quand le SWA tend vers 90° ?

Dans un second temps, il va utiliser une approche géométrique pour effectuer des corrections sur les CR ainsi que le SWA. Les résultats proposés grâce à cette approche semblent très bons malgré l'absence d'échelle sur les images.

2.5 Simulation de la formation d'une image SEM

Comme on l'a vu, l'utilisation directe de méthodes classiques de reconstruction permet de retrouver la topographie de motifs variés, tant que les images SEM ne sont pas trop entachées d'artefacts électroniques. Cependant, lorsque l'on veut utiliser ces méthodes sur des motifs de très petites tailles (qui sont aujourd'hui la norme en microélectronique), on ne peut plus se soustraire à la compréhension de ces artefacts de manière à les prendre en compte lors de la reconstruction. Pour cela des modèles physiques basés sur des méthodes dites "Monte Carlo" (ainsi nommées à cause de l'utilisation de techniques probabilistes pour la résolution du problème) ont été développés dans le but de pouvoir simuler une image SEM.

Le modèle physique JMONSEL

Il existe différents programmes permettant de réaliser des simulations Monte Carlo (Casino, Electron Flight Simulation, Win X-Ray ...), mais le plus connu dans le milieu de la microélectronique est celui développé notamment par J.Villarrubia du NIST (National Institute of Standards and Technology) : JMONSEL (Java MONte Carlo simulator for Secondary Electrons) [41] [42].

Pour simuler la formation d'une image SEM, il faut simuler le parcours des électrons dans la matière. JMONSEL ne fait pas une résolution exacte du problème car les électrons sont traités indépendamment les uns des autres, il n'y a donc pas d'effets liés au temps comme par exemple les effets d'accumulation de charges pouvant conduire à des distorsions dans les images SEM à cause de la répulsion coulombienne entre particules de même charge.

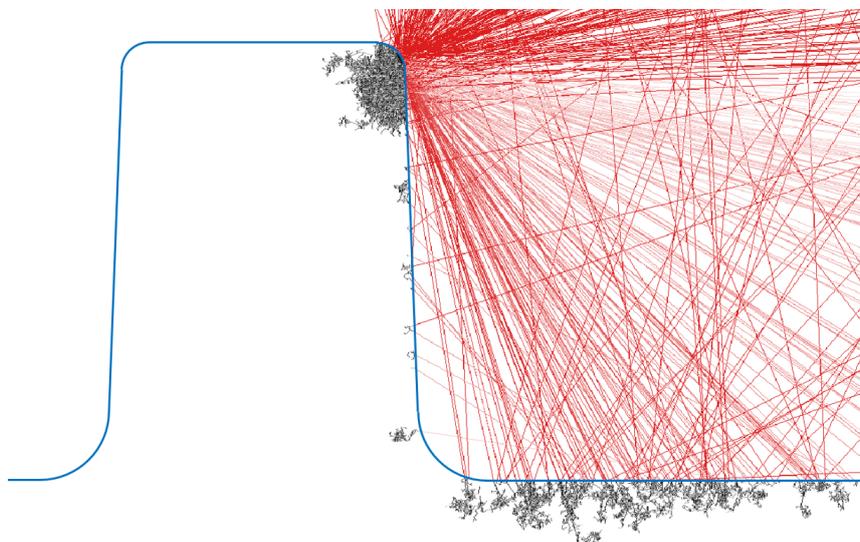


FIGURE 2.10 – Résultat JMONSEL pour la trajectoire de 200 électrons envoyés en un point d'un motif en silicium de type ligne de 150nm de hauteur et 88° de SWA (en bleu) avec une énergie incidente de 800eV et un spot size de 0,5nm. Les électrons du faisceau incident ne sont pas représentés. Les lignes rouges correspondent aux trajets des électrons en dehors du matériau (soit vers le détecteur, soit ré-entrants dans le matériau). Les lignes noires correspondent aux trajets des électrons dans le matériau.

Le modèle va calculer pour chaque électron son déplacement sans interaction dans la matière tel que la probabilité de non diffusion évolue en $e^{-\gamma x}$, avec x la distance parcourue par l'électron et γ le taux de diffusion (dépendant du matériau et de la section efficace elle même

dépendante de l'énergie des électrons). La distance moyenne parcourue sans interaction par l'électron, appelée le libre parcours moyen est égale à $1/\gamma$ [41].

Entre chacun de ces déplacements, l'électron subit l'une des interactions physiques possibles en fonction de la probabilité d'occurrence de cette interaction, jusqu'à ce qu'il ait perdu toute son énergie (piégé dans la matière) ou qu'il se soit échappé de la matière.

Les interactions inélastiques engendrent une perte d'énergie de l'électron, mais la prise en compte de toutes les interactions inélastiques possibles étant impossible (notamment celles dont le transfert d'énergie est faible), cette perte d'énergie est ramenée à une fonction de perte d'énergie continue en fonction de la distance parcourue [43] [42]. La perte d'énergie liée à une interaction inélastique peut produire des électrons secondaires dont le parcours sera également étudié.

Il existe différents modèles physiques utilisables avec JMONSEL dont les principales différences proviennent des approximations du traitement des sections efficaces, de la génération des électrons secondaires, des propriétés des matériaux, du traitement des phénomènes de diffusion aux interfaces... [42].

Pour fonctionner ce programme va prendre en entrée :

- La taille de la zone d'analyse (x,y) ainsi que la topographie et les matériaux de la zone à analyser.
- Les informations sur la physique des matériaux ainsi que les phénomènes physiques (et leur modèle associé) que l'on va prendre en compte pour ce matériau. Exemple : Pour des simulations à faible accélération des électrons on préfère utiliser le modèle physique de Mott pour le calcul des sections efficaces, qui est plus précis dans ces conditions que le modèle classique de Rutherford, car prenant en compte des effets quantiques.
- La taille d'un pixel et le nombre d'électrons envoyés par pixel.
- L'énergie des électrons envoyés et éventuellement (ce sera important pour mon travail) l'angle d'incidence des électrons sur l'échantillon.
- Le diamètre du faisceau d'électrons au niveau de l'échantillon (spot size).
- La surface de récupération des électrons s'échappant du matériau. Pour simuler un détecteur d'électrons secondaires, on prendra ici en général une demi-sphère au dessus de l'échantillon pour simuler le champ électrique qui ramène la majorité des électrons secondaires vers le détecteur.

A partir de ces paramètres d'entrée, l'analyse permet de trouver le nombre d'électrons collectés par le détecteur en chaque pixel analysé, permettant de tracer en chaque pixel le rendement électronique (rapport du nombre d'électrons envoyés sur le nombre d'électrons collectés) comme présenté en Fig 2.11. La courbe présentée correspond au niveau de gris (en relatif) de l'image SEM qui sera obtenue si on analyse ce motif dans les conditions précisées en paramètres d'entrée. Lorsque l'on fait une analyse d'un motif dont la forme ne varie pas dans une direction (comme une ligne invariante selon l'axe y tel que présentée en Fig 2.11), il est équivalent de simuler A lignes en envoyant B électrons par pixel que de simuler une ligne en envoyant $A \times B$ électrons par pixel. De manière générale, plus un motif (voire même le procédé en général) présentera de symétries plus on pourra simplifier l'analyse. Par exemple sur la Fig 2.11, la ligne possède un axe de symétrie selon l'axe y ce qui nous aurait permis de simuler seulement la moitié du profil et donc diviser par deux le temps de calcul.

JMONSEL est un logiciel très puissant mais son principal désavantage est sa lenteur ; en effet pour tracer une ligne avec des conditions suffisantes de précision, il faut simuler un grand nombre de trajectoires d'électrons ce qui entraîne un temps de calcul important, devenant critique si l'on souhaite simuler une zone 2D entière, d'où l'utilité de simplifier le problème en fonction des symétries.

Pour accélérer le processus on peut utiliser des outils de calcul plus performants, comme par exemple réaliser les calculs en parallèle sur plusieurs cœurs (multi-thread), ou encore comme

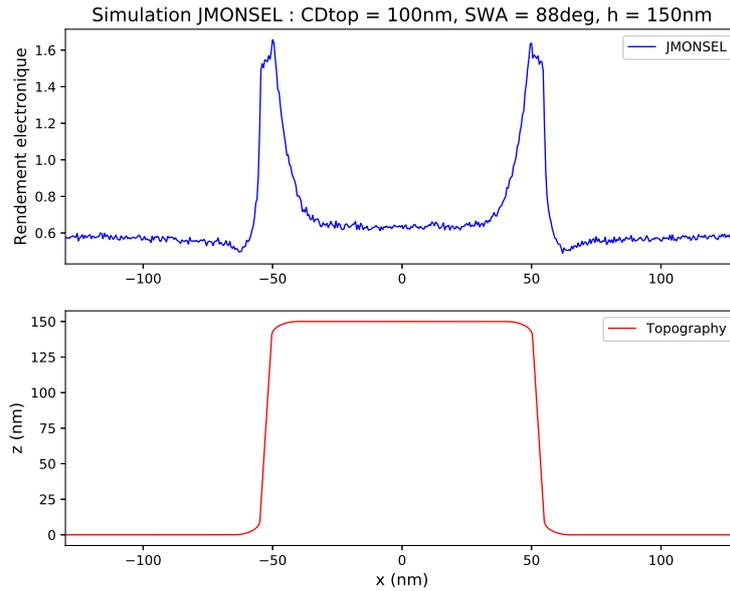


FIGURE 2.11 – Résultat JMONSEL de la simulation d’une ligne (en haut) dont la vue en coupe est représentée dans le cadre inférieur (Attention : Proportion (x,z) non conservée pour pouvoir utiliser le même axe x). On observe facilement l’effet de flanc, c’est à dire la forte probabilité de sortie des électrons due aux flancs du motif.

présenté dans [44] où est proposée une résolution par utilisation de la puissance de calcul des GPU.

2.5.1 Les modèles compacts

Une alternative est d’utiliser d’autres modèles qui traiteront le problème à un niveau plus ”haut”, c’est à dire ne pas résoudre le problème électron par électron mais simuler le comportement général de la poire d’interaction électronique. C’est notamment la méthode qu’avait proposé Nyssonen [45] grâce au calcul d’une fonction de densité de probabilité.

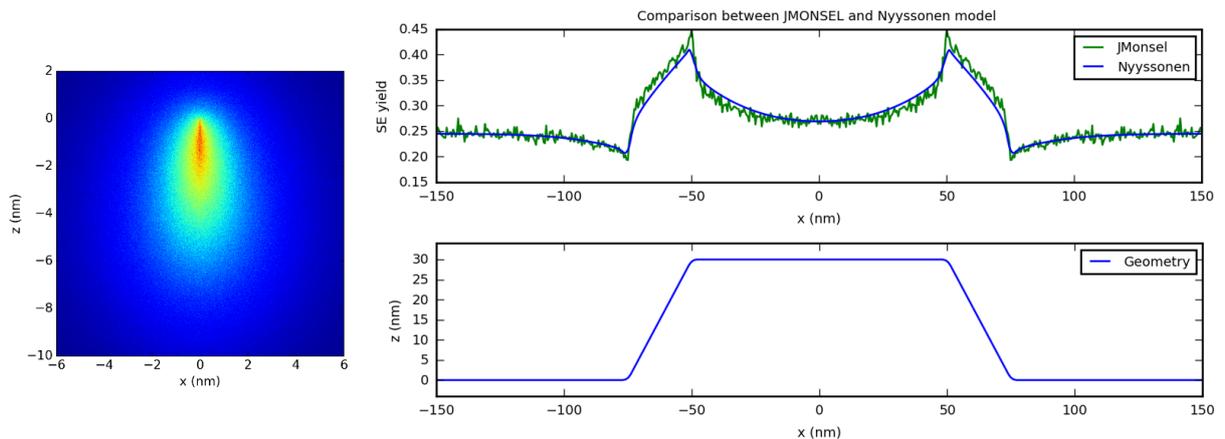


FIGURE 2.12 – Gauche : Fonction de densité de probabilité calculée par JMONSEL pour une énergie de 500eV. Droite : Comparaison entre le résultat JMONSEL et la méthode de Nyssonen sur une topographie pour une énergie de 2500eV [46].

Cette méthode a d’ailleurs été reprise [46] mais en utilisant une fonction de densité de

probabilité trouvée grâce à JMONSEL sur un très grand nombre d'électrons pour une énergie et un matériau donné (gauche de la Fig 2.12). Il a alors été montré que l'application de cette fonction de densité de probabilité appliquée sur la topographie d'un motif permettait d'obtenir un résultat cohérent avec JMONSEL (droite de la Fig 2.12). Les courbes obtenues par cette méthode sont très lisses comparées à celles obtenues par JMONSEL car elles se comportent comme si un nombre très important d'électrons ayant toujours exactement les mêmes interactions avait été envoyé en tout point de la topographie.

Seeger [47] propose quant à lui une méthode intermédiaire pour accélérer la méthode JMONSEL. Il pré-calculé des trajectoires d'électrons dans un matériau massif qui seront ensuite "collées" sur la topographie pour une énergie et un matériau donné également. Il peut ainsi à la fois obtenir la rapidité tout en gardant une part d'aléatoire car il pourra par exemple utiliser 10000 trajectoires choisies aléatoirement parmi le million pré-calculé en chaque point.

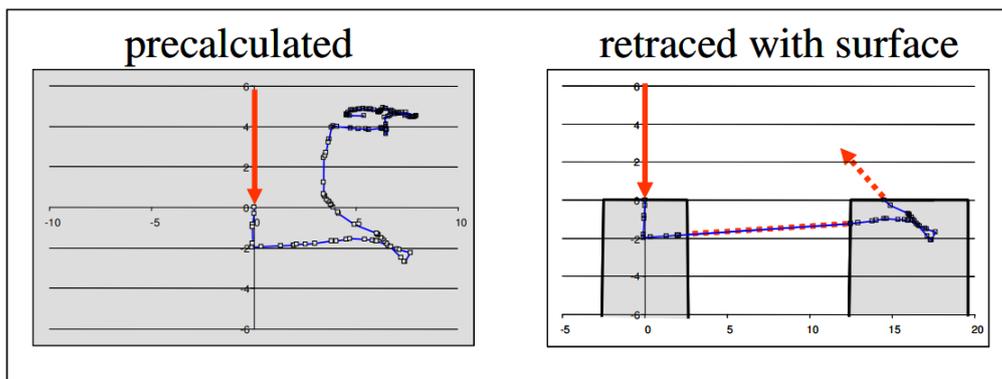


FIGURE 2.13 – Exemple de trajectoire pré-calculée (gauche) dans un volume infini de matériau (en gris) et sa trajectoire correspondante retracée sur une topographie particulière (droite) [47].

L'autre avantage des modèles compacts, c'est qu'il est possible à partir de cette résolution du problème dit "direct" d'essayer de résoudre le problème dit "inverse", c'est à dire de partir de l'image (ou des images) SEM pour en déduire la topographie.

2.6 Résolution du problème inverse

Un problème inverse est considéré comme bien posé si il a une solution unique et stable (c'est à dire que de petites fluctuations dans la mesure ne conduisent pas à une variation excessive du résultat) [48]. L'unicité peut éventuellement être obtenue par l'ajout de contraintes supplémentaires, comme par exemple dans le cas de la reconstruction topographique par la connaissance de la polarité d'un motif.

Pour réussir à obtenir la taille des CD du haut et du bas du motif à partir d'une image SEM, Frase [49] propose d'approximer le signal de niveau de gris obtenu par des sections de fonctions exponentielles (Fig 2.14). Les coefficients de ces fonctions sont calculés pour ajuster au mieux sur le signal d'entrée. Une fois ces coefficients optimisés il est alors possible d'obtenir des grandeurs d'intérêt. Cependant la résolution n'est pas simple en raison du nombre important de paramètres. C'est pourquoi il résout le problème en sous parties dans lesquelles il trouve la solution de quelques uns des paramètres en faisant des hypothèses sur les autres.

Chris Mack, tout comme Frase propose de décomposer le signal SEM comme des fonctions exponentielles, mais de manière plus détaillée grâce à l'ajout de nombreux paramètres permettant de mieux prendre en compte la présence de congés (CR) en haut et en bas des motifs. [50].

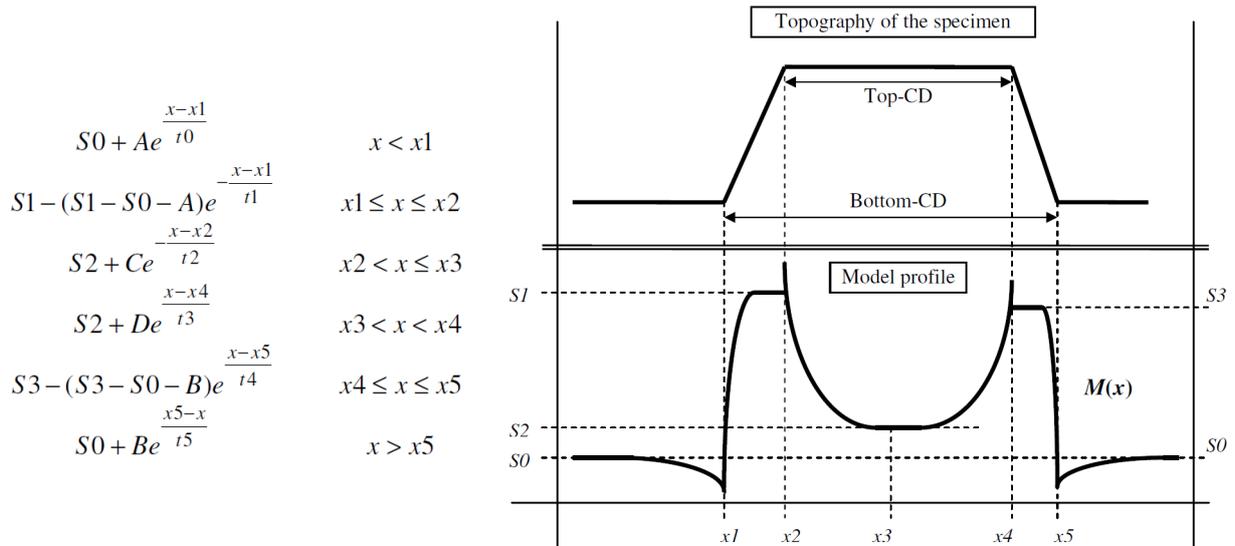


FIGURE 2.14 – Gauche : Approximation par parties d’un profil SEM par des fonctions exponentielles. Droite : Correspondance entre la topographie et le profil SEM avec les paramètres associés [49].

Seeger a également proposé une méthode de reconstruction basée sur le shape-from-shading [51], permettant à la fois de réaliser une simulation de surface à partir d’une topographie ainsi que la résolution du problème inverse. Pour cela il utilise une banque de filtres créée à partir de simulations JMONSEL, qui va être utilisée pour convoluer une topographie avec différents filtres permettant d’estimer le signal SEM qui sera obtenu, par la sommation de ces différentes convolutions (Fig 2.15)

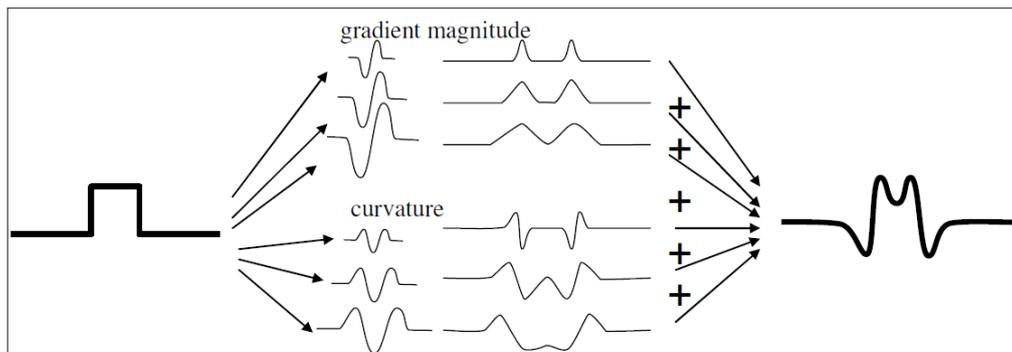


FIGURE 2.15 – La topographie de gauche est convoluée avec différents filtres de convolution de pente et de courbure pour obtenir des images intermédiaires qui seront ensuite sommées pour obtenir l’estimation de l’image SEM [51].

Cette technique, en plus d’être bien plus rapide qu’une méthode de résolution physique, permet d’être inversée. En effet il est possible de partir de l’image SEM et de déconvoluer grâce à la même banque de filtres de manière à obtenir un résultat topographique. Seeger présente des reconstructions très probantes à partir de cette méthode, qui a cependant quelques limitations qu’il explique. Tout d’abord la technique se base sur une fonction de la hauteur, ce qui rend impossible la résolution de motifs dont la pente peut être supérieure à 90° . Il souligne également l’impossibilité d’obtenir la hauteur à partir d’une image seule, en précisant qu’il s’agit d’un problème sous-contraint. Les résultats présentés sont donc calibrés en hauteur

grâce à une mesure AFM.

Enfin je rajouterais que sa méthode basée sur des convolutions a tendance à arrondir les coins anguleux comme on peut facilement s'en rendre compte en regardant ses résultats de reconstructions, comparé à de l'AFM. C'est une limitation que j'avais déjà soulevé dans la partie traitant du shape-from-shading car je pense qu'il s'agit d'une limitation inhérente à la technique. Cette technique sera donc plus adaptée pour les topographies à gradient faible en général.

Enfin Seeger fait une remarque très juste; il dit qu'il est très compliqué d'utiliser le niveau absolu de gris d'une image SEM, car celui-ci peut être modifié de manière automatique par le SEM lors du focus, de l'amélioration de la luminosité et du contraste. Pour que cette information soit utilisable il faudrait donc contrôler de manière très précise la chaîne de traitement du SEM.

2.7 Conclusion

Dans ce chapitre nous avons vu différentes techniques de reconstruction tri-dimensionnelle, ainsi que leur adaptation à la reconstruction à partir d'images de microscopie électronique.

Dans la suite de ce travail, nous avons choisi d'utiliser et de développer la méthode de reconstruction géométrique utilisant des images d'un même motif imagé à plusieurs angles de vues différents, comme présentée dans la Fig 2.4, grâce à l'utilisation d'un microscope à faisceau inclinable (Fig 1.9). Nous pensons que c'est la méthode qui permettra d'obtenir des résultats précis et robustes pour la reconstruction de motifs de petite taille (de la dizaine à la centaine de nanomètres) et à gradient fort (SWA proche de 90°).

De plus, le logiciel JMONSEL sera largement utilisé pour réaliser des simulations, dans le but d'observer les effets de la topographie d'un motif ainsi que des conditions de prise d'image, sur la réponse électronique d'un SEM, afin de créer un modèle de reconstruction topographique.

Dans la partie suivante, nous allons nous arrêter sur les outils permettant une analyse quantitative des grandeurs utiles à la caractérisation de motifs typiques de la microélectronique.

Chapitre 3

Outils pour la reconstruction 3D

Dans ce chapitre on introduira deux outils qui seront ensuite utilisés dans tout le reste de ce travail. Tout d'abord on proposera une approximation d'un motif ligne typique de la microélectronique par une série de quantités que l'on nommera les grandeurs d'intérêt. Dans un second temps on présentera la méthode générale de pré-traitement des images SEM expérimentales, avant la reconstruction 3D.

3.1 Paramétrisation de la géométrie : Approximation "8 paramètres"

Pour pouvoir créer une base commune de travail entre les différents instruments de mesure qui seront utilisés dans la thèse, il est nécessaire de trouver une approximation de la topographie d'un motif typique de l'industrie de la microélectronique. Cette approximation doit être suffisamment précise pour être représentative des grandeurs d'intérêts en évitant les détails trop fins qui ne pourront pas être correctement estimés. Grâce à cette approximation il sera alors possible de comparer les résultats topographiques obtenus par reconstruction SEM avec les mesures de référence (AFM ou FIB/STEM).

On rappelle dans la figure 3.1 ci-dessous les différentes grandeurs d'intérêts (noir) qui seront utiles pour la description des motifs, sur une vue en coupe d'un motif ligne (bleu) avec les lignes de construction (vert, orange et rouge) qui seront explicitées par la suite. Les notations utilisées sont explicitées page 3.

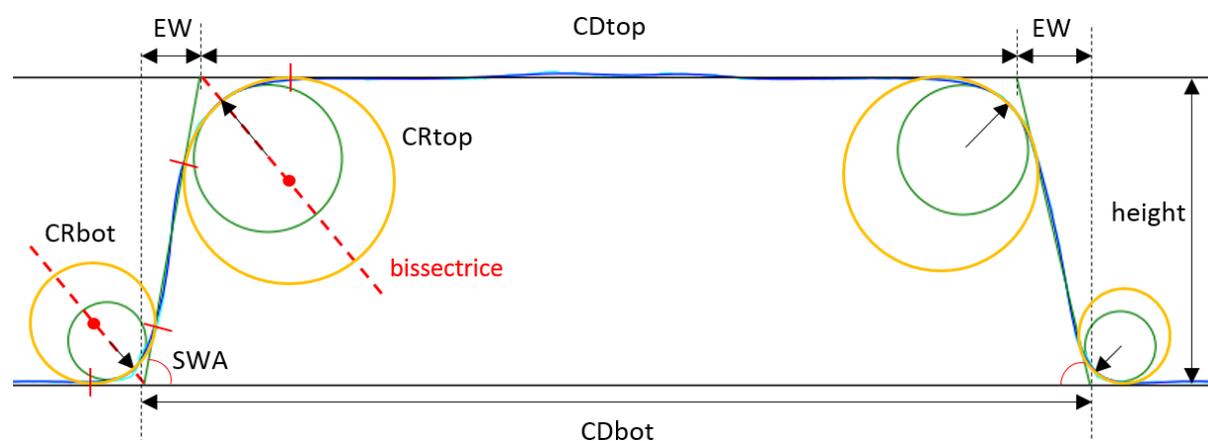


FIGURE 3.1 – Représentation des grandeurs d'intérêts sur un motif de type ligne, vue en coupe.

Pour approximer la topographie du profil, la première étape est de transformer les coor-

données topographiques (x, z) en un objet mathématique que nous allons pouvoir manipuler. Nous avons décidé d'utiliser les splines car elles peuvent être lissées de manière à réduire en partie le bruit de mesure. Les splines sont des fonctions définies par morceau par des fonctions polynômiales. Il faut noter que l'utilisation des splines a des conséquences sur l'analyse des données ; en effet, plus l'erreur tolérée sur l'ajustement de la spline sera grande plus le bruit sera éliminé, mais également plus la précision de l'approximation sera mauvaise ; il faut donc trouver un compromis pour que la spline soit la plus proche des données d'entrée sans toutefois être trop influencée par le bruit de mesure. Cette problématique sera étudiée plus en détail dans la partie 7.1.

3.1.1 Étude de la courbure

Le profil étant désormais un objet mathématique, il est alors possible de le dériver en fonction de la coordonnée x . Grâce à cette dérivée on peut facilement repérer les zones correspondant à des flancs et ainsi déterminer une valeur z moyenne pour le haut du motif et une autre pour le bas du motif et donc obtenir la hauteur h du motif.

On peut ensuite calculer la courbure C de la spline et le rayon de courbure R en tout point donnés par les formules :

$$C = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{3/2}} \text{ et } R = \frac{1}{C} \quad (3.1)$$

On peut noter que dans le cas où il existe au maximum une valeur y pour un x donné (c'est à dire dans le cas où $y = f(x)$ avec f une fonction continue quelconque, c'est à dire quand le motif ne présente pas des flancs ré-entrants ($\text{SWA} > 90^\circ$)), alors l'équation peut être simplifiée :

$$C = \frac{y''}{(1 + y'^2)^{3/2}} \quad (3.2)$$

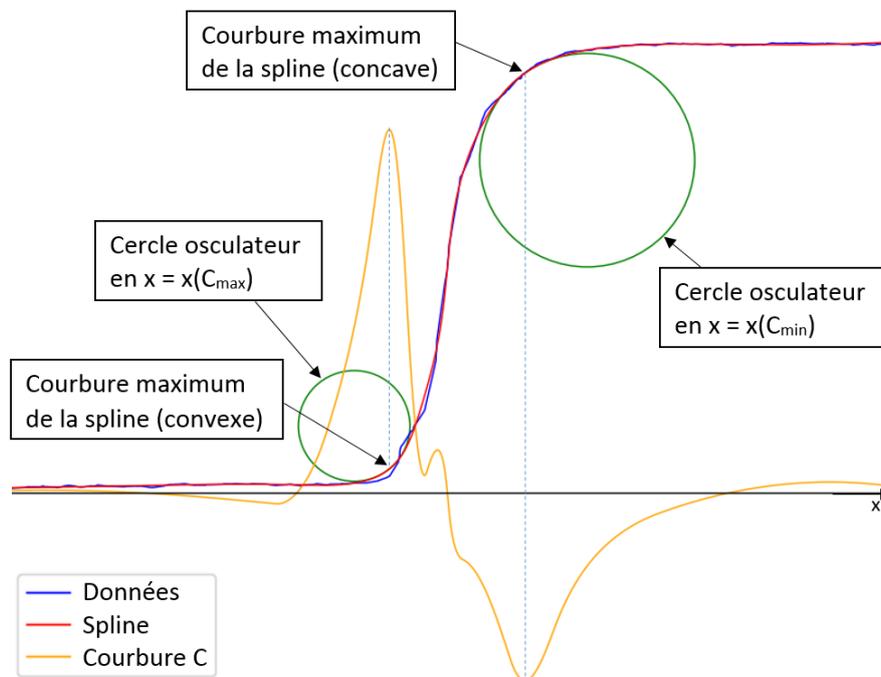


FIGURE 3.2 – Représentation du flanc gauche d'un motif de type ligne. On peut voir que l'approximation par une spline (rouge) bien calibrée permet de lisser les détails (bleu) auxquels le SEM sera de toute façon insensible. La courbure (orange) permet de repérer les CR.

Un cercle osculateur est un objet mathématique permettant la description locale d'une courbe suffisamment régulière en un point donné. Le rayon de courbure de ce cercle est donné en tout point par la formule du rayon de courbure explicitée ci-dessus. Les positions des CR des motifs seront définies par le point de la topographie (donc de la spline) où la courbure est maximale en valeur absolue, c'est à dire où ce rayon de courbure est minimal.

Dans notre cas nous n'aurons pas de problème de définition car la topographie brute ayant été transformée en spline, les points non dérivables auront été préalablement "arrondis" par la spline, plus ou moins en fonction de ses paramètres de fit. Par exemple la pointe d'une fonction valeur absolue sera arrondie par la spline la rendant dérivable en zéro. On peut noter toutefois que la courbure en zéro de la spline de la fonction valeur absolue ne reflète que le niveau d'ajustement de la spline. On pourra se trouver dans ce genre de situation si le motif observé est très anguleux ; les CR tendront alors vers un minimum lié à l'ajustement de la spline.

On va alors pouvoir déterminer pour chacune des "zones de flanc" :

- En bas du flanc le cercle osculateur de la spline en $x = x(C_{max})$. A ce stade, le rayon de ce cercle va définir la valeur du rayon du congé en bas du flanc (CRb).
- En haut du flanc le cercle osculateur de la spline en $x = x(C_{min})$. A ce stade, le rayon de ce cercle va définir la valeur du rayon du congé en haut du flanc (CRt).

Ces cercles sont représentés en vert dans la Fig 3.1 et la Fig 3.2.

3.1.2 Définition du SWA

Il convient ensuite de trouver une définition du SWA qui permette de représenter au mieux la pente des flancs du motif. En effet, en fonction de la définition mathématique donnée au SWA il pourra prendre des valeurs sensiblement différentes. Les différentes définitions du SWA qui ont été testées sont présentées, jusqu'à celle qui a été retenue.

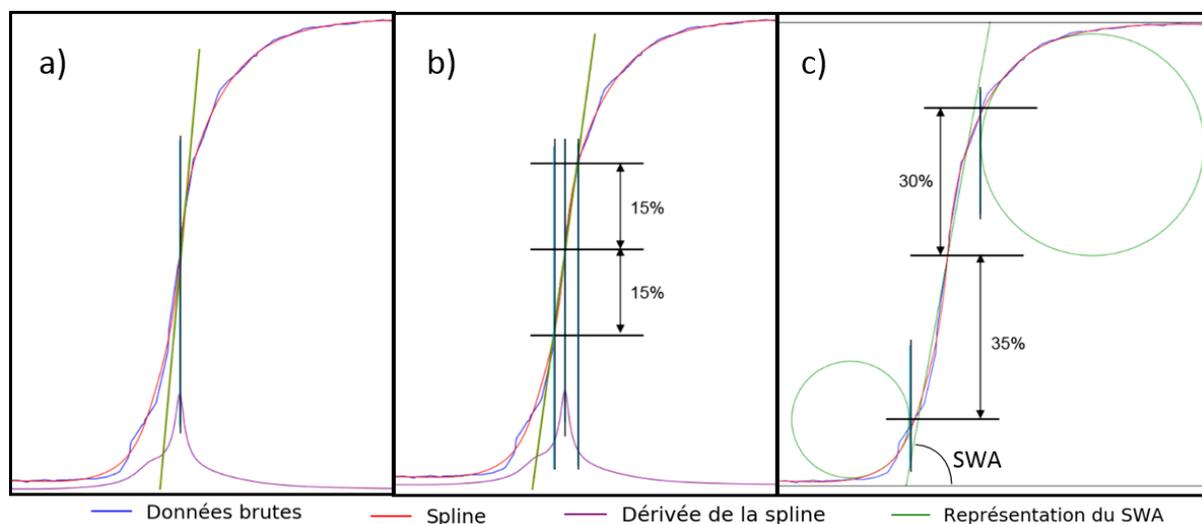


FIGURE 3.3 – Représentation des différentes définitions données au SWA. a) $SWA = \max(\text{dérivée})$, b) $SWA = \text{moyenne}(\text{dérivée}[\text{intervalle}])$, c) $SWA = \text{moyenne}(\text{dérivée}[\text{intervalle entre les CR}])$. Les valeurs présentées sont qualitatives et la dérivée est en unité arbitraire.

L'idée la plus évidente au premier abord a été de prendre la valeur maximale de la dérivée de la spline ; cependant dans la plupart des cas cela conduit à une estimation du SWA qui est trop grande car il y a souvent un point du flanc qui est très pentu et qui représente mal la pente moyenne du flanc (Fig 3.3 a).

De cette observation, il paraît logique d'évoluer vers une représentation du SWA comme une portion de la pente, c'est à dire par l'utilisation de la moyenne de la dérivée sur un certain intervalle centré sur la dérivée maximale. Cette définition fonctionne bien dans le cas où la dérivée de la zone de flanc est symétrique (c'est à dire si le profil présente un point de symétrie à mi-hauteur de la hauteur du flanc, c'est à dire si $CRb \approx CRt$), mais est moins efficace dans le cas d'un motif dissymétrique. Or, il est très courant d'avoir par exemple un grand CRb avec un petit CRt (Fig 3.4) ou inversement (Fig 3.3) à cause du procédé de fabrication utilisé. L'autre problème de cette définition est que la valeur trouvée dépend fortement de la taille de l'intervalle qui a été choisi arbitrairement : plus l'intervalle sera grand plus le SWA sera petit (Fig 3.3 b). La dernière évolution a donc été de choisir automatiquement l'intervalle en fonction des valeurs des rayons de courbures des cercles précédemment définis. De notre expérience, c'est une méthode puissante car elle permet un choix automatique du pourcentage de pente utile, sans risque de sous utilisation ni de dépassement sur les CR (Fig 3.3 c).

3.1.3 Définition des CR

Ensuite pour que l'approximation soit la plus simple possible, nous avons décidé de redéfinir les CR de manière à ce qu'ils soient tangents au SWA ainsi qu'au haut du motif pour le CRt et au bas du motif pour le CRb (Voir cercle orange sur la Fig 3.1 et l'exemple de construction en rouge).

Plus précisément le CRt correspond à l'unique cercle dont le centre se situe sur la bissectrice de l'angle entre le SWA et le haut du motif et touchant le point d'intersection entre la spline et la bissectrice, ainsi que le SWA en un unique point (et de ce fait par construction le haut en un unique point). De même pour le CRb mais avec le bas du motif (Voir les lignes rouges sur la Fig 3.1).

Avec cette définition la portion de cercle qui va servir sera différente en fonction du SWA du motif ; Lorsque $SWA = 90^\circ$ un quart de cercle sera utilisé, si $SWA < 90^\circ$ une portion plus faible sera utilisée (on peut facilement s'en convaincre en prenant le cas limite d'un angle tendant vers la ligne droite, la portion de cercle utilisée tendra alors vers zéro). Lorsque $SWA > 90^\circ$ une portion plus grande du cercle sera utilisée (on peut facilement s'en convaincre en prenant le cas limite d'une pointe dont l'angle tend vers zéro, la portion de cercle utilisée tendra alors vers la moitié).

Cette simplification affecte peu la qualité de l'ajustement ; en revanche cette définition change substantiellement la valeur donnée au CR. Ces cercles sont représentés en orange dans la Fig 3.1 et seront par la suite considérés comme la valeur des CR.

3.1.4 Application du modèle paramétrique sur des données expérimentales

Finalement on obtient une approximation de notre profil constituée uniquement de portions de droites et de portions de cercles. Cette approximation est définie par huit valeurs : quatre CR, deux SWA, une hauteur et une distance entre les deux flancs.

On peut noter que les deux flancs des motifs étant traités séparément, l'approximation fonctionne sur des motifs dissymétriques droite/gauche comme on peut le voir sur la Fig 3.4 à gauche.

Deux exemples d'approximation sur des motifs en silicium mesurés à l'AFM sont présentés en Fig 3.4 :

Ligne (gauche) : $h = 167.9$ nm, $SWAl = 89.3^\circ$, $SWAr = 89.6^\circ$, $CRbl = 34.2$ nm, $CRbr = 29.6$ nm, $CRtl = 3.6$ nm, $CRtr = 3.3$ nm, $CDt = 98.0$ nm, $CDb = 101.2$ nm.

Tranchée (droite) : $h = -126.7$ nm, $SWAl = 84.6^\circ$, $SWAr = 89.6^\circ$, $CRbl = 5.2$ nm, $CRbr = 19.0$ nm, $CRtl = 15.6$ nm, $CRtr = 13.1$ nm, $CDt = 41.7$ nm, $CDb = 28.7$ nm.

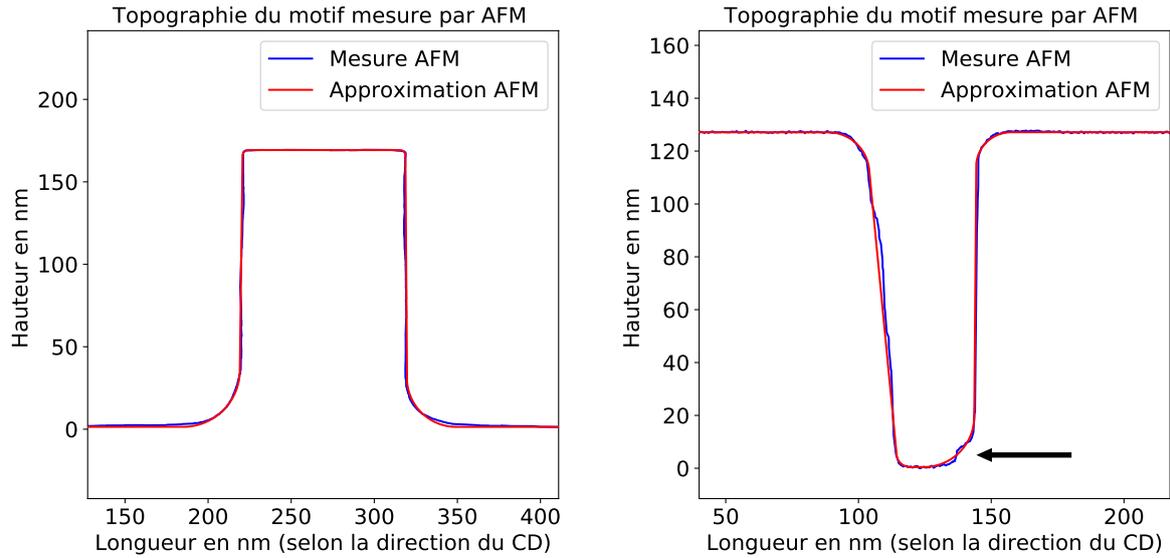


FIGURE 3.4 – Approximation d’une mesure AFM (bleu) par l’approximation 8 paramètres (rouge) sur un motif ”ligne” (gauche) et un motif ”tranchée” (droite).

Sur le motif ”tranchée” on peut voir des limitations de cette approximation. La mesure AFM du CRb de droite présente une anomalie (indiqué par une flèche) ce qui résulte en une surestimation du CR. Comme cette anomalie est présente sur toutes les mesures AFM on peut suspecter que cela vient d’un effet de pointe; il faudrait donc refaire la mesure avec une nouvelle pointe. Dans le cas où il s’agirait de la vraie topographie du motif, on se heurterait alors à une limitation de l’approximation présentée (même si voulue) qui n’est pas censée tenir compte de ce genre de topographie. On peut toutefois noter que le CR s’est adapté pour en tenir compte. Si le motif approximé est trop éloigné de la forme générale attendue, alors l’approximation risque d’échouer. Enfin on peut également remarquer que le flanc gauche est légèrement bombé, ce qui n’est pas pris en compte par l’approximation.

Grâce à la simplicité de cette approximation, nous allons pouvoir créer une référence atteignable et quantitative pour la reconstruction SEM. Le but de la reconstruction SEM ne sera donc pas de trouver une topographie quelconque libre, mais de trouver ces huit paramètres qui permettront de décrire au mieux la topographie du motif observé via l’approximation ”8 paramètres”.

En fonction des besoins, il est envisageable d’améliorer cette approximation pour pouvoir prendre en compte des topographies plus complexes. Par exemple le flanc du motif pourrait être approximé par plus d’une pente pour la reconstruction de motif présentant un ou plusieurs points d’inflexion. Pour les structures présentant des flancs incurvés, les portions de lignes pourraient être remplacées par des courbes...

3.2 Analyse d'une image SEM

L'objectif de cette partie est de présenter comment les images issues de mesures SEM sont analysées de manière à pouvoir être utilisées pour la reconstruction topographique.

Tout d'abord on rappelle qu'une image SEM est une matrice de pixels dont les valeurs peuvent varier entre une borne minimum qui représente un pixel noir (c'est à dire un minimum de rendement électronique) et une borne maximum qui représente un pixel blanc (c'est à dire un maximum de rendement électronique). En fonction de l'appareil utilisé et du format d'enregistrement des images, ces bornes peuvent avoir des valeurs absolues différentes (les plus courantes sont 8 bits : 0 à 255 ou 16 bits : 0 à 65535). Pour pouvoir traiter de façon simple tout type d'entrée on commence par normaliser la valeur absolue du niveau de gris des images SEM pour que la valeur absolue maximum soit égale à 1.

Dans ce travail de thèse nous avons décidé que les images SEM d'entrée devraient toujours être des lignes orientées verticalement. En effet l'inclinaison du faisceau électronique (tilt) provenant de la droite, la meilleure façon de l'exploiter est d'avoir des flancs orthogonaux à la direction du tilt. Des améliorations de la méthode qui va être présentée pourraient permettre d'étendre les possibilités de reconstruction (discuté dans la partie 8.2). On peut noter qu'une grande partie des motifs de la microélectronique sont dit "Manhattan", c'est à dire que ce sont des lignes soit horizontales soit verticales, notamment pour les grilles des transistors, pour des raisons de limite de résolution lorsque les dimensions sont très faibles.

3.2.1 Transformation des images en profil et méthode du ruban

La première étape de traitement des images SEM est la transformation de l'image en un profil horizontal de niveaux de gris, en calculant la valeur moyenne des pixels de chaque colonne de l'image. Cette étape va permettre de réduire le bruit de mesure caractéristique d'une image SEM en fonction du nombre de lignes moyennées. A cette étape de l'analyse nous ne voulons pas tenir compte de la rugosité des flancs des motifs (Line Edge Roughness : LER); c'est pourquoi avant la transformation en profil, une étape d'alignement des flancs des images est réalisée (Fig 3.5).

Pour cela une courbe appelée "ruban" sera ajustée sur chacun des flancs de manière à obtenir la forme de la rugosité du motif sur la hauteur de l'image. A partir de ce ruban, chaque ligne de l'image originale va subir une translation horizontale de manière à rendre la courbe de rugosité verticale et centrée en sa moyenne en x. Ceci est équivalent à transformer le repère Cartésien en un repère curviligne qui suivrait le flanc du motif. On peut noter que ces translations peuvent être des portions de pixel, mais comme on veut tout de même obtenir une image finale dans laquelle les pixels sont les uns au dessus des autres, une étape d'interpolation sera réalisée pour donner la valeur de niveau de gris à chaque pixel en fonction de son pourcentage d'appartenance au pixel directement à sa droite et directement à sa gauche. Ce traitement peut aussi bien être utilisé sur des motifs dont l'image a été prise en vue du dessus qu'avec une inclinaison du faisceau électronique.

On peut noter plusieurs conséquences de cette transformation. Tout d'abord pour éviter les problèmes liés à des irrégularités en bord d'image, les rubans ne sont pris en compte qu'à partir d'une certaine distance des bords; les images redressées seront donc légèrement plus petites (verticalement) que celles d'origine. Les translations vont faire bouger les bords; la taille horizontale de l'image sera donc également réduite, en prenant comme limite gauche la ligne qui a le plus translaté à droite et inversement.

Cette opération va en quelque sorte transposer toute la rugosité d'un flanc vers l'autre flanc. Cette opération sera donc réalisée séparément pour chacun des flancs et les opérations sur les grandeurs du flanc droit seront réalisées sur l'image dont le flanc droit a été redressé et inversement. Cette transformation ne va pas altérer la distance moyenne sur chacune des images

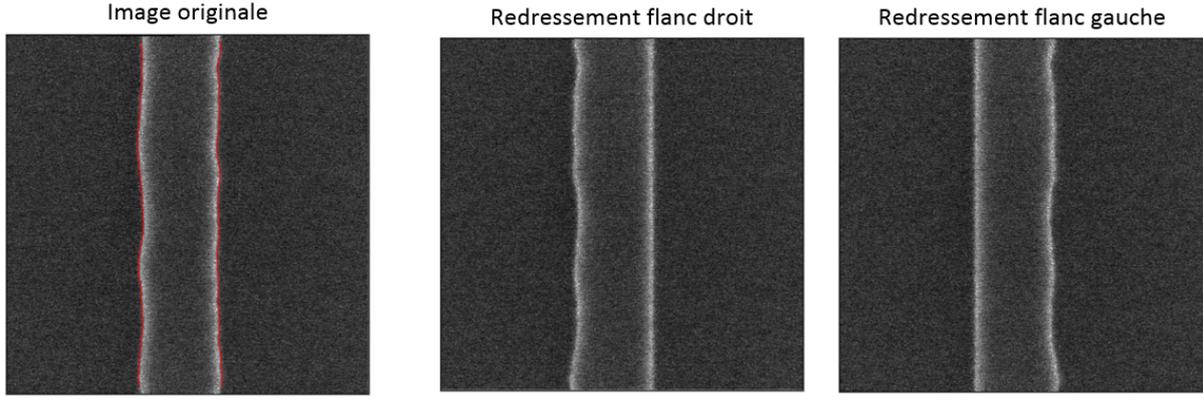


FIGURE 3.5 – Redressement des flancs de manière à éviter l’élargissement artificiel des EW dû à la rugosité de ligne, lors de l’étape de sommation des colonnes. FOV = 500 nm.

entre les deux flancs car chaque ligne de pixels bien que translatée gardera sa distance bord à bord. De plus le redressement est réalisé de manière à garder le centre de gravité des rubans au même x , les distances moyennes entre les deux bords seront donc inchangées même entre l’image redressée à droite et l’image redressée à gauche.

Enfin on peut noter qu’il s’agit de la seule opération réalisée directement sur l’image et par conséquent il s’agit de l’opération la plus longue de tout le procédé.

3.2.2 Limite du ruban

Lorsque le flanc n’est pas parfaitement orthogonal à la direction du tilt, alors l’EW observé est réduit car seule la projection du tilt sur le flanc participe à l’agrandissement de l’EW, on a alors :

$$EW = \frac{h}{\tan(\text{SWA})} + h \times \tan T \times \sin(\theta) \quad (3.3)$$

où θ est l’angle entre l’azimut de l’inclinaison du faisceau électronique (c’est à dire la direction d’où provient le tilt) et la tangente au flanc du motif (Fig 8.1). On a donc une réduction de la taille apparente du flanc à cause de la rugosité ; cependant dans ce travail de thèse ce phénomène a été négligé car lorsque θ s’écarte en moyenne sur l’image de moins de 10° de l’orthogonalité (en positif ou en négatif) alors $\sin(\theta) \approx 1$. Une fonctionnalité a tout de même été ajoutée pour vérifier que l’on est bien dans ce régime de fonctionnement (sur les images utilisées $(90 - \theta) < \pm 5^\circ$ dans la majorité des cas).

L’utilisation du ruban pourra donc également compenser une légère inclinaison et/ou incurvation de la ligne, tant que l’on reste dans le régime précédemment décrit ; au-delà la compensation 3.3 devra être utilisée.

Finalement ces rubans sont conservés en mémoire pour pouvoir ré-introduire la rugosité du motif plus tard dans l’analyse.

3.2.3 Les descripteurs

Une fois ce traitement réalisé, on se retrouve donc non plus avec des images mais avec des profils de niveaux de gris en fonction de l’axe x . Ces profils bruts sont lissés par une convolution avec une gaussienne de manière à réduire le bruit d’un pixel à l’autre. Chaque profil lissé est ensuite transformé en spline dont on calcule la dérivée première et seconde. On place alors des points sur la spline correspondant à des valeurs particulières de sa dérivée. Ces positions selon l’axe x seront nommées des descripteurs.

La technique précise de placement de ces descripteurs est accessible en annexe (Appendices, II.). Chaque flanc est analysé de manière similaire (Fig 3.6). Le descripteur C est positionné au maximum d'intensité du flanc (dérivée nulle), c'est à dire à l'endroit où le rendement électronique est maximal. Les descripteurs B et D sont placés de part et d'autre du point C (vers l'intérieur du motif pour D et l'extérieur pour B) lorsque la dérivée est maximale en valeur absolue, c'est à dire à l'endroit où le gradient de rendement électronique est le plus important. Le descripteur A est positionné après B lorsque la dérivée s'annule. Cependant comme il n'y a pas toujours un effet d'ombrage en bas des motifs (notamment en mode tilté), le point A ne serait pas consistant entre les images si il était placé lorsque la dérivée s'annule. Son placement est donc lié à une valeur de coupure de la dérivée "proche de zéro", la valeur optimale étant empirique; elle a notamment été trouvée grâce aux simulations JMONSEL pour l'analyse des CR que nous verrons plus tard (4.4). Enfin le descripteur E partage la même définition que A de manière à être consistant mais vers l'intérieur du motif.

On peut voir le résultat du placement de ces descripteurs sur des profils issus de vraies images SEM sur la Fig 3.6 ainsi que l'impact du tilt sur les descripteurs. C'est à partir de ces descripteurs que les images SEM vont être analysées de manière à obtenir les informations topographiques recherchées.

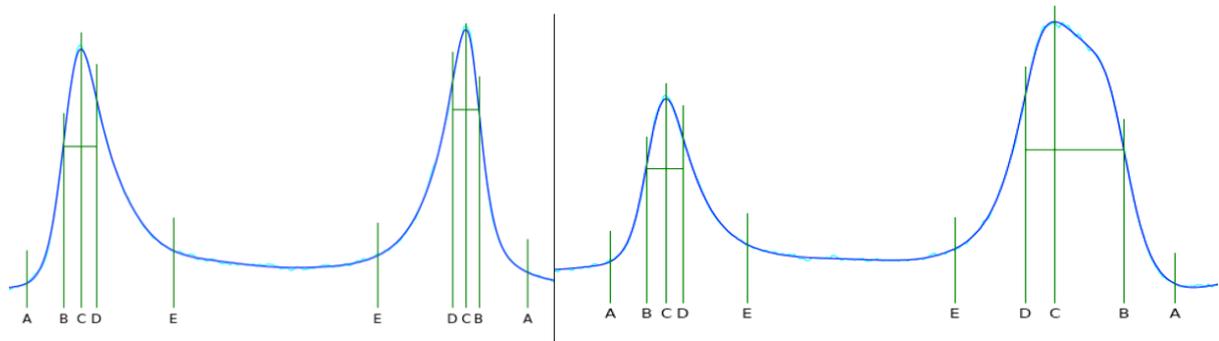


FIGURE 3.6 – Profil de niveaux de gris d'une image prise en vue du dessus (gauche) et avec un tilt de 12° (droite). Les descripteurs A, B, C, D et E correspondant à des valeurs particulières de la dérivée sont présentés.

A partir de ces descripteurs il serait possible de relier de manière directe les distances entre ces descripteurs à des grandeurs d'intérêt. Cependant lorsque l'on utilise ce genre de méthode directe on obtient des erreurs qui deviennent critiques à mesure que la taille du motif diminue. Ces erreurs étaient attendues, car nous ne prenons pas en compte les effets des interactions électron-matière liés à la topographie des motifs observés. Il est facile d'en être convaincu en prenant un cas limite : par exemple un motif de 90° de SWA observé en vue du dessus devrait avoir un EW de 0nm; or ce n'est jamais le cas à cause de la probabilité croissante des électrons de s'échapper du matériau lorsque le faisceau se rapproche du bord. Il existe beaucoup d'autres artefacts plus difficiles à se représenter juste par la pensée et qui seront détaillés dans la partie suivante 4.

Chapitre 4

Détermination des tendances d'évolution des descripteurs, à partir de l'analyse de simulations physiques

L'état de l'art ainsi que nos premières expériences sur la reconstruction tridimensionnelle de motif à partir d'image SEM nous ont permis de déterminer que les artefacts électroniques intrinsèques à l'imagerie SEM entraînaient des difficultés pour obtenir une mesure topographique précise pour différents types de motifs. C'est pourquoi il est nécessaire de comprendre les différents effets liés aux interactions électron-matière à partir de simulations SEM de manière à pouvoir analyser au mieux les images SEM pour en déduire un modèle permettant de remonter aux grandeurs topographiques.

4.1 Intérêt de la simulation SEM

Comme on l'a vu dans l'état de l'art, il est difficile d'obtenir une approximation correcte de la topographie d'un motif typique de la microélectronique ($h < 500$ nm, $CD < 500$ nm, SWA proche de 90° , CR variables) par des mesures directes. Ce travail de thèse se propose donc d'étudier les images SEM d'une grande diversité de motifs typiques de la microélectronique de manière à créer un modèle linéaire permettant de reconstruire la topographie sous forme d'une approximation "8 paramètres". Ce modèle sera détaillé plus tard dans la partie 5, mais il est important de garder en tête que toutes les analyses qui vont être présentées, seront réalisées dans l'optique de créer des grandeurs mesurables permettant de créer ce modèle de reconstruction topographique, c'est à dire un modèle qui permettra de remonter aux huit grandeurs d'intérêt de l'approximation "8 paramètres".

Théoriquement la meilleure solution pour comprendre et analyser ces effets serait de fabriquer des wafers présentant une large variété de topographies. Cependant, comme nous nous appuyons sur l'approximation "8 paramètres", il faudrait faire varier cinq des huit grandeurs d'intérêt (car certaines sont juste dédoublées pour la gauche et la droite) de manière indépendante pour pouvoir isoler les effets de chacune d'elles sur l'image SEM. Il est important de noter que, pour que le modèle puisse optimiser toutes les quantités d'intérêt, celles-ci doivent varier (et si possible de manière indépendante) dans l'ensemble des motifs utilisés pour créer et calibrer le modèle.

En pratique il est difficile de pouvoir réaliser un tel ensemble de motifs car il est difficile de faire varier chacune des grandeurs d'intérêt en conservant toutes les autres ainsi que les conditions de prise d'image du SEM inchangées. De plus dans le cas de l'utilisation de vrais motifs la

connaissance de leur topographie réelle est également un problème car elle doit être mesurée par une méthode de référence. Or comme on l'a vu (partie 1.3), les méthodes de références ne sont pas adaptées pour la mesure de masse comme on aimerait en avoir pour calibrer un modèle (d'où l'intérêt de la mesure topographique par SEM). Enfin les mesures de référence peuvent également présenter des erreurs, d'autant plus que leur lenteur peut questionner sur la reproductibilité des mesures sur un temps long. Or la reproductibilité est tout à fait critique pour la calibration du modèle ; en effet si on mesure toujours avec un biais constant le modèle s'ajustera avec un biais. En revanche, si les mesures de référence évoluent dans le temps (par exemple à cause de l'usure d'une pointe AFM au fur et à mesure des mesures), alors le modèle risque d'être pollué par des artefacts externes aux variations de topographie.

C'est pour toutes ces raisons que nous avons décidé de commencer par la réalisation de simulations dans le but d'identifier les tendances d'évolution des images SEM en fonction des valeurs des grandeurs d'intérêt. Grâce à ces tendances un modèle pourra être créé et calibré à partir de données simulées sur une large plage de valeurs et grandeurs d'intérêt. Dans un second temps des wafers seront fabriqués de manière à ajuster le modèle issu de la simulation sur des images SEM réelles.

4.2 Paramètres du simulateur JMONSEL

Nous avons choisi d'utiliser le logiciel JMONSEL pour réaliser nos analyses car c'est un logiciel reconnu pour la qualité de ses simulations dans le domaine de la microélectronique. Il a été utilisé dans un grand nombre de publications du domaine.

Certains des paramètres d'entrée des simulations JMONSEL ont été pris pour être homogènes avec les conditions de prise d'image du SEM que nous avons utilisées :

- La seule tension d'accélération des électrons accessible pour prendre des images avec le faisceau électronique incliné est de 800 V, c'est pourquoi l'énergie des électrons primaires des simulations a toujours été réglée à 800 eV.
- Après avoir réalisé des tests sur le SEM pour connaître la taille optimale des pixels de l'image (comparaison temps de mesure / qualité de l'image / échantillonnage lié à la taille du faisceau électronique) il a été décidé que les images utilisées pour ce travail de thèse auraient une taille de pixel carré de 0,5 nm de côté.
- les tilts utilisés sont les mêmes que ceux accessibles sur le SEM soit : 0° (topview), 3°, 5° et 12°.
- Les motifs simulés sont entièrement en silicium. En effet pour éviter la complexité des phénomènes d'accumulation de charges liés au bombardement électronique dans un matériau diélectrique, ainsi que la simulation de matériaux non cristallins dont les paramètres physiques sont plus difficiles à connaître avec précision, nous avons choisi de créer un modèle pour la reconstruction de motifs en silicium gravé.

D'autres paramètres ont en revanche été choisis pour optimiser le résultat de la simulation :

- Le nombre d'électrons envoyés par pixel a été réglé sur 10000, ce qui permet d'avoir un bruit acceptable (Fig 4.1) tout en limitant le temps de simulation (avec ce jeu de paramètres la simulation d'un motif dure environ une heure).
- La taille du faisceau électronique à la surface du wafer (spot size) a été réglée sur une valeur nulle. En effet si on veut introduire un flouage dû à la taille de faisceau il est possible de le faire en convoluant le résultat à spot size nul avec une gaussienne. En revanche, si le résultat de la simulation a été fait à spot size non nul il n'est alors pas possible de connaître quel aurait été le résultat à spot size nul, c'est pourquoi

le spot size a toujours été introduit après la simulation et non comme paramètre d'entrée.

Enfin pour simuler le collecteur d'électrons secondaires (définition dans la partie 1.2) du SEM, seuls les électrons dont l'énergie est inférieure à 50 eV en sortie de matériau seront pris en compte pour la formation de l'image SEM simulée.

Différentes analyses de simulations JMONSEL vont donc être présentées sur différents motifs "8 paramètres" dont les grandeurs d'intérêts vont varier indépendamment les unes des autres.

4.3 Influence de la hauteur et du SWA sur le signal SEM

Les premières quantités d'intérêt que nous avons choisi de faire varier sont la hauteur et le SWA car ce sont les deux paramètres les plus importants pour réussir une reconstruction topographique correcte.

Dans cette analyse, 80 motifs différents ont été simulés avec les grandeurs suivantes :

Constante : $CDt = 100 \text{ nm}$, $CRb = CRt = 0 \text{ nm}$, $T = 0^\circ$

Variation : $h = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 400, 500] \text{ nm}$

$SWA = [70, 75, 80, 85, 90]^\circ$

4.3.1 Influence du SWA

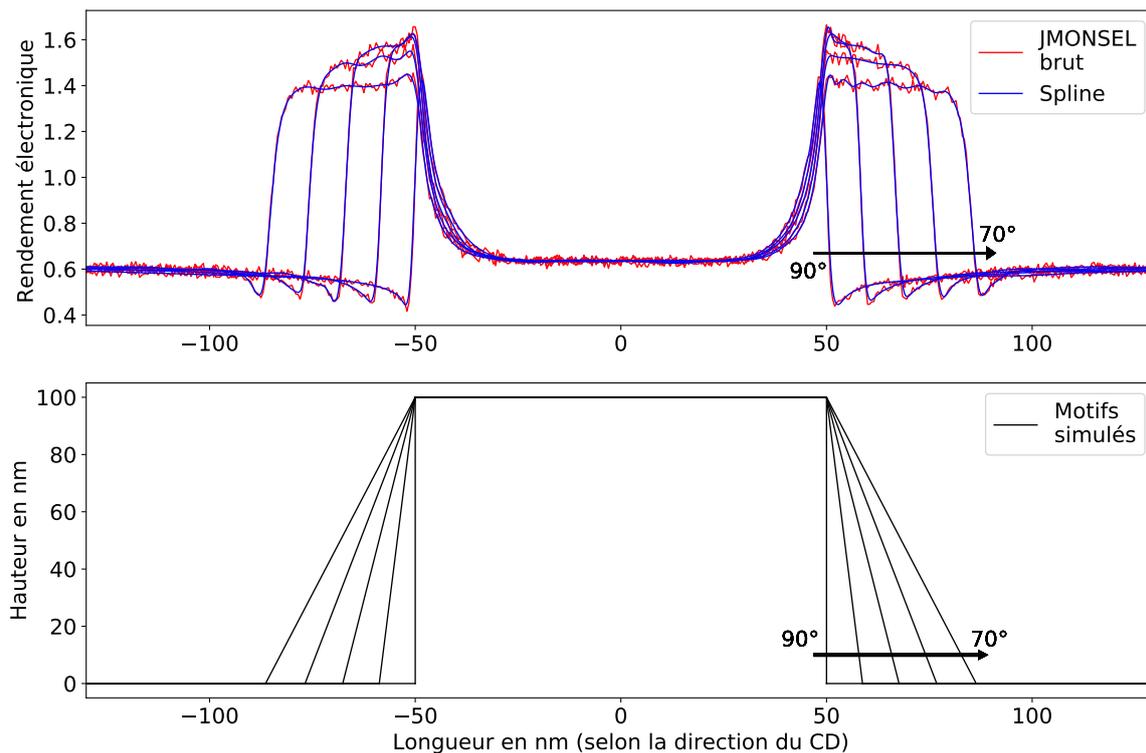


FIGURE 4.1 – Résultat de simulations JMONSEL de cinq motifs de 100nm de haut pour différents SWA allant de 90° à 70° par pas de 5°. En bas sont représentés en coupe les motifs utilisés comme entrée de la simulation. Au dessus sont représentées les courbes calculées par JMONSEL en fonction des motifs d'entrée.

La Fig 4.1 présente les résultats de la simulation JMONSEL de cinq motifs de 100 nm de haut pour différents SWA. Dans cette première figure, le résultat brut de JMONSEL a été

représenté en rouge pour pouvoir apprécier le niveau de bruit résultant de la simulation à 10000 électrons par pixel. Par la suite seule la spline approximant ce résultat brut, tracée en bleu, sera représentée pour plus de clarté. Il faut noter que les images provenant d'un SEM réel seront également moyennées et approximées par une spline, ce qui par conséquent rend cette approximation homogène avec ce qui sera fait pour l'utilisation finale de cette technique pour reconstruire de vrais motifs.

On rappelle que ces courbes correspondent au rapport entre le nombre d'électrons secondaires (moins de 50 eV) sortant du matériau par rapport au nombre d'électrons primaires envoyés, ce qui correspond directement à la valeur relative de niveau de gris de l'image SEM qui en résultera. Le terme "niveau de gris" sera d'ailleurs généralement préféré à "rendement électronique", car dans la méthode finale sur un SEM nous n'aurons pas accès à la valeur absolue du rendement électronique mais à un niveau de gris relatif.

Les tendances que nous cherchons par la simulation vont être repérées par l'analyse des similarités et différences des courbes entre elles lorsque l'une des grandeurs d'intérêt varie indépendamment des autres. Par exemple sur la Fig 4.1 on peut voir l'effet d'une variation de SWA pour certaines valeurs fixes des autres grandeurs d'intérêt.

La différence évidente est l'élargissement de la zone de fort rendement électronique correspondant à la zone de flanc du motif, mais cet élargissement sera traité dans la Fig 4.4 avec l'ensemble des 80 simulations, car il n'est pas spécifique à un changement de SWA uniquement.

La seconde différence est le niveau de gris de cette même zone de flanc. On peut observer une diminution du niveau de gris lorsque le SWA baisse ce qui est en accord avec la tendance de la loi de Lambert dont on a parlé dans l'état de l'art ($I=1/\cos(\text{pente})$). Cependant quand SWA tend vers 90° cette tendance ralentit et même diminue. Ceci est dû à la diminution de la taille de l'EW lorsque le SWA tend vers 90° ; on sort alors du régime stable de la loi de Lambert où l'on se situe sur un plan infini de pente donné.

4.3.2 Influence du SWA sur le CDt

Enfin on peut voir un effet très intéressant lorsque le faisceau électronique est en haut du motif et qu'il se rapproche de l'un des flancs; on peut alors observer l'augmentation de la probabilité de sortie des électrons due au rapprochement du flanc. Cette augmentation est plus ou moins rapide en fonction de la valeur du SWA. Cet effet est important car il modifie une zone de la courbe qui sera utilisée pour l'estimation du CD du haut du motif. Or, à cause de cet effet l'estimation du CDt pourrait être modifiée par la valeur du SWA.

Pour décrire la distance entre les deux flancs, le descripteur "D" de chaque flanc a été utilisé (définition en partie 3.2 sur les descripteurs et Fig 3.6).

Sur la Fig 4.2 sont tracées à gauche les distances entre ces deux descripteurs en fonction du SWA pour toutes les hauteurs simulées (représentées par les différentes couleurs). Quelle que soit la hauteur des motifs on observe toujours cet effet de réduction de la distance DD lorsque le SWA tend vers 90° .

A partir de cette distance DD, on voudrait être capable d'estimer le CDt; en effet il est évident que cette distance DD devrait varier de manière linéaire avec le CDt au moins pour les valeurs de CD "grandes" (on verra la problématique des "petits CD" et les limitations qui y sont associées dans la partie 4.5).

Tout d'abord on peut voir que cette distance DD sous-estime de manière systématique le CDt, on peut donc compenser cette différence systématique par un biais. Il faut ensuite compenser l'effet lié au SWA par une fonction du SWA. Cette fonction a été trouvée de manière théorique. En effet on peut comprendre cet effet par le rapprochement du flanc et de la poire d'interaction des électrons avec le matériau lorsque le SWA tend vers 90° . Pour mieux comprendre l'effet une vue en coupe d'un motif a été représenté en bas de la Fig 4.2b.

Tout d'abord on sait que la probabilité de sortie d'un électron de la matière est proportionnelle

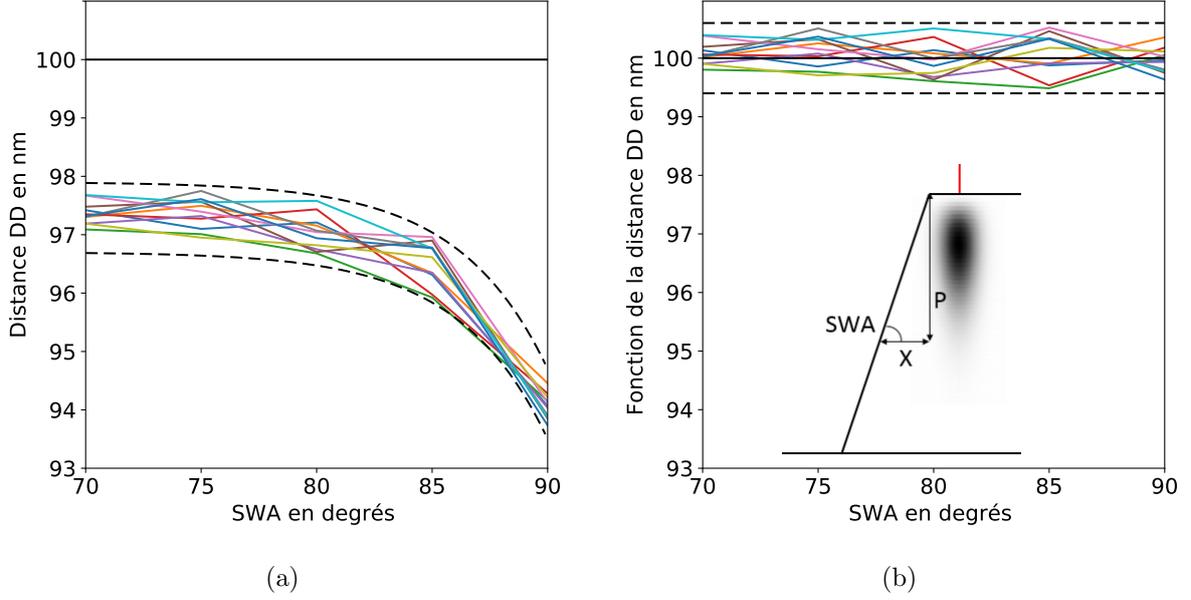


FIGURE 4.2 – (a) : Distance entre les deux descripteurs D pour toutes les hauteurs (couleur) en fonction du SWA. Le trait noir plein représente la valeur du CDt simulé (100 nm). Les pointillés noirs représentent l’enveloppe des courbes. (b) : Valeur du CDt approximée à partir de la mesure de la distance DD et une fonction du SWA.

à une décroissance exponentielle de la distance à la surface (au moins en première approximation) [52] [53]. Ensuite d’après la vue en coupe on peut voir que cette distance est égale à $P \times \tan(90 - \text{SWA})$ (on pourrait se poser la question entre tan et sin mais pour les petits angles $\tan \approx \sin$ et lorsque le SWA va diminuer la distance X va augmenter, résultant en un écrasement de l’effet dans la décroissance exponentielle de X), où P peut être vu comme la profondeur de pénétration des électrons dans la matière. L’effet a donc été compensé en utilisant la fonction :

$$f(\text{SWA}) = A \exp(-X) \quad (4.1)$$

$$f(\text{SWA}) = A \exp(-P \times \tan(90 - \text{SWA})) \quad (4.2)$$

La valeur de P est très probablement liée à l’accélération des électrons ainsi qu’au matériau sondé car ce sont les deux paramètres qui vont influencer sur la profondeur de pénétration des électrons dans la matière. Pour le moment seules des simulations sur du Silicium et à une accélération de 800 V des électrons ont été réalisées, c’est pourquoi on prendra P comme une constante. Cependant si on voulait étendre le domaine de validité à d’autres matériaux et énergies des électrons il faudrait faire des simulations dans ce sens pour trouver la fonction $P(\text{mat}, \text{Vacc})$.

Dans notre cas A et P ont donc été trouvés de manière empirique de façon à obtenir le résultat de la Fig 4.2b. La valeur de P trouvée est de 15 nm, ce qui est cohérent avec la taille des poires d’interactions à 800 eV dans le Silicium. La valeur de A trouvée est de 3.2 ce qui est difficile à interpréter (ce serait une sorte de facteur de longueur de diffusion ou de libre parcours moyen) mais elle doit probablement dépendre principalement du matériau.

Les pointillés noirs de la Fig 4.2 représentent l’enveloppe des courbes. Comme ces courbes ne sont pas ordonnées en fonction de la hauteur du motif, l’enveloppe peut être considérée comme l’erreur à laquelle on peut s’attendre sur la mesure de DD (a) et sur l’estimation du CDt (b). Sur cet exemple l’écart type sur l’estimation du CDt est de 0.27 nm sur une mesure unique du CDt (dans ce cas idéal de simulation sans CR).

Enfin, nous avons également remarqué que l’évolution de la distance DD dépendait de manière non négligeable du niveau de lissage des données d’entrée. Pour l’application aux images SEM il faudra donc faire attention à la taille du spot et adapter cette fonction f.

4.3.3 Influence de la hauteur

Dans cette partie on s'est intéressé à l'impact d'une variation de la hauteur sur le signal SEM pour un SWA constant.

Sur la Fig 4.3 on peut observer que le niveau de gris des flancs est en moyenne le même quelle que soit la hauteur du motif (et donc la taille de l'EW). Ceci nous conforte dans la tendance de type loi de Lambert, à savoir qu'en régime stable, seule la pente influe sur le niveau de gris pour une illumination et un matériau donné. Régime stable veut dire qu'il n'y a pas d'effet de bord, on peut d'ailleurs voir que sur les motifs les plus petits, la valeur maximale du rendement électronique diminue, de la même manière qu'elle diminuait lorsque le SWA tendait vers 90° dans la Fig 4.1. Dans ces deux cas la raison est la même, on sort du régime stable car l'EW est très petit et on a donc des effets de bord : on n'est plus dans le cas d'un plan infini à pente donnée.

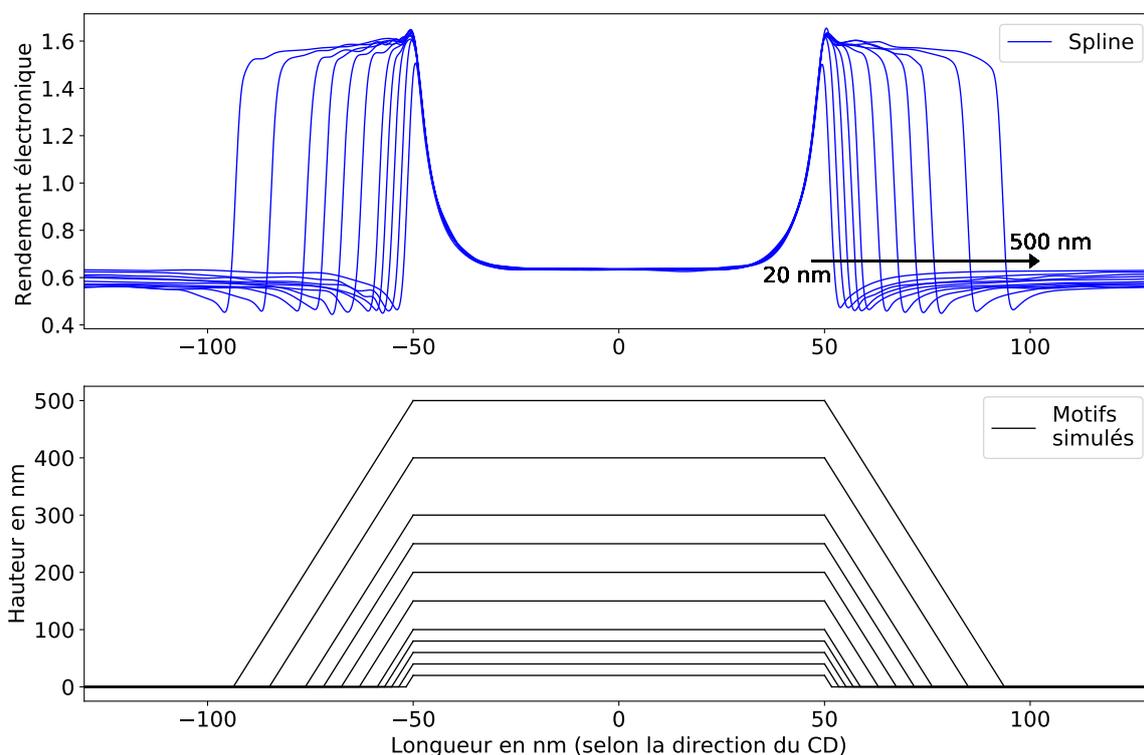


FIGURE 4.3 – Résultat de simulations JMONSEL de onze motifs de 85° de SWA pour différentes hauteurs allant de 20 nm à 500 nm. En bas sont représentés les motifs simulés vus en coupe. Au dessus sont représentées les courbes calculées par JMONSEL en fonction des motifs d'entrée.

Ensuite on voit que les courbes n'ont pas le même niveau de gris en bas des motifs mais cette effet sera traité en détail dans la partie 4.7 sur l'effet d'ombrage.

Finalement, tout comme dans la Fig 4.1 on observe l'évidente augmentation de la taille des EW lorsque la hauteur augmente et/ou que le SWA diminue, c'est l'objet de l'étude suivante.

4.3.4 Influence de la hauteur et du SWA sur l'EW

On sait qu'en absence de tilt, l'EW théorique peut être calculé géométriquement par la formule :

$$EW = \frac{h}{\tan(\text{SWA})}, \text{ pour } T=0^\circ \quad (4.3)$$

La Fig 4.4 propose donc de comparer la distance entre les descripteurs B et D (voir Fig 3.6) à la valeur théorique de l'EW simulé, pour vérifier si ces descripteurs seront pertinents pour l'estimation de l'EW.

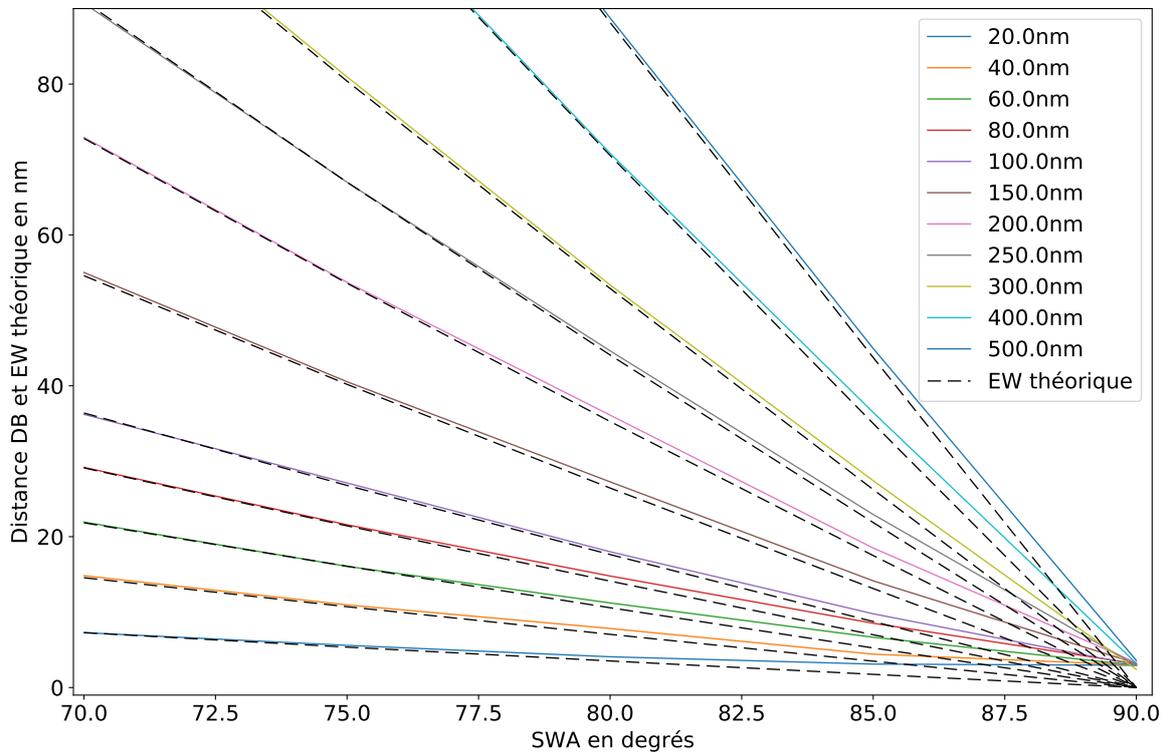


FIGURE 4.4 – Comparaison de la distance BD (traits colorés pleins) et l'EW théorique (pointillés) pour toutes les simulations réalisées, c'est à dire pour différents couples de SWA (axe des abscisses) et hauteur (couleur du trait).

On peut voir que la distance BD est corrélée à la valeur de l'EW, ce qui était attendu. Cependant on voit également que lorsque le SWA tend vers 90° la corrélation diminue pour tendre vers une valeur limite différente de zéro en 90° (environ 3 nm) conformément à l'observation sur l'impossibilité d'avoir des EW nuls sur des images SEM.

Dans la mesure du possible on essaiera donc de ne jamais se trouver dans le cas où les EW tendent vers ce minimum, notamment en utilisant le tilt du SEM, ce sera l'objet de l'étude de la partie 4.6.

4.4 Influence des CR sur le signal SEM

Les secondes quantités d'intérêt qui ont été analysées sont les congés en haut (CRt) et en bas (CRb) des motifs. En effet, la quantité critique pour la bonne estimation de la hauteur et du SWA des motifs est l'EW, or on peut anticiper que la variation des CR va entraîner une variation dans la mesure des EW, c'est pourquoi nous avons décidé de nous intéresser aux CR.

4.4.1 Influence des CR sur le signal SEM global

Avant d'entrer dans l'analyse détaillée, la Fig 4.5 permet de se représenter l'effet que vont avoir les CR sur une image SEM à partir de résultats de simulation JMONSEL représentés sous forme d'images SEM. On peut voir sur la partie centrale un motif sans CRb ni CRt. On observe alors l'augmentation de rendement électronique classique due à l'approche d'un flanc

du motif. Sur sa droite on peut voir que l'ajout d'un CRt entraine une augmentation de la zone de "rendement élevée" vers l'intérieur du motif, alors que sur sa gauche, l'ajout d'un CRb entraine une augmentation de la zone de "rendement élevée" mais vers l'extérieur du motif. Les deux parties extérieures sont identiques et permettent de voir que ces deux phénomènes peuvent s'additionner sans influencer l'un sur l'autre.

Même si la différence peut paraître évidente, elle ne l'est que parce que l'on peut les observer relativement aux autres. Cependant si on donne seulement une des parties de cette simulation à un observateur, il lui sera très difficile de savoir si elle provient d'un motif présentant des CR et encore moins d'estimer leurs valeurs. Si on lui en présente deux sans informations supplémentaire il croira probablement que l'un des motifs a un SWA supérieur à l'autre.

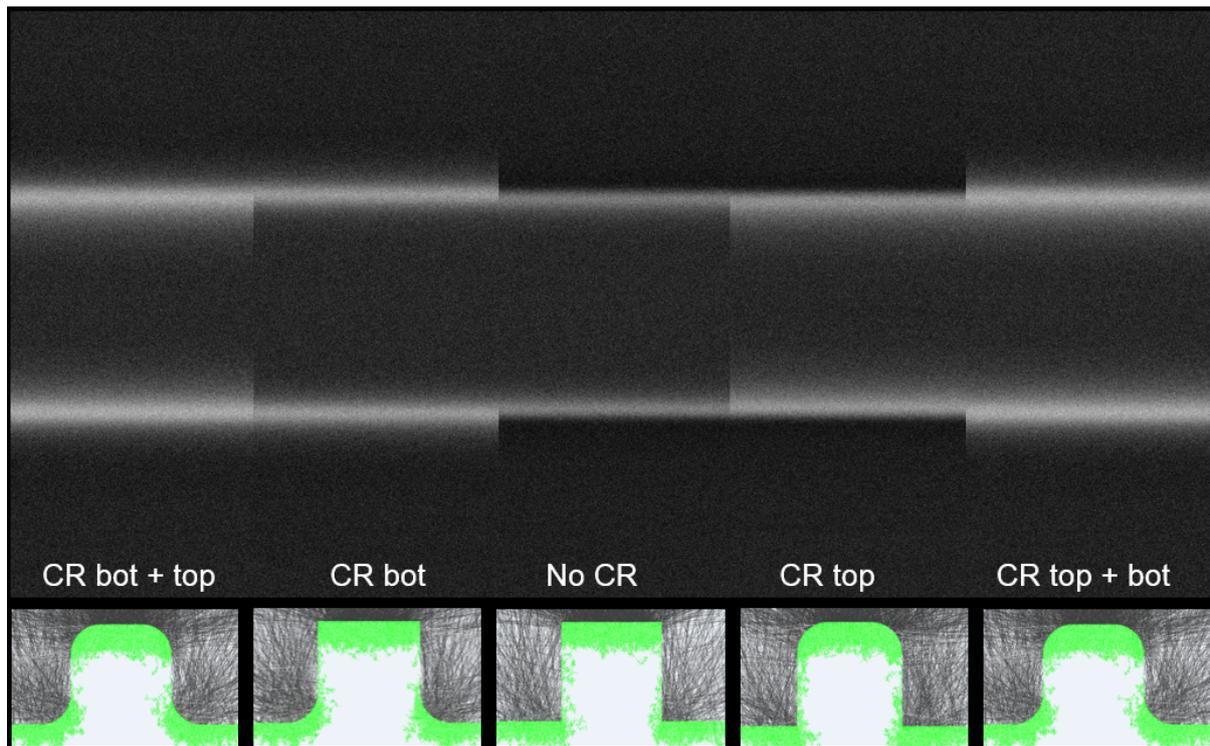


FIGURE 4.5 – Résultats de simulations JMONSEL sur quatre motifs de 100 nm de hauteur et de CD, à un SWA de 90° , affichés sous forme d'une image SEM. Le motif est une ligne horizontale dont les CR sont soit nuls soit égaux à 30 nm par partie (les transitions ne sont pas simulées pour bien voir la différence de niveau de gris). Un lissage équivalent à un spot size d'image SEM a été appliqué (1,5 nm de spot size). En bas sont représentés les motifs simulés vus en coupe avec les zones d'interaction électronique en vert.

Le but de cette analyse est de trouver des descripteurs permettant une estimation des CRb et CRt respectivement, ainsi que d'observer l'influence des CR sur les autres grandeurs d'intérêt et notamment l'EW et le CD.

Dans cette analyse 294 motifs différents ont été simulés avec les grandeurs suivantes :

Constante : $CD_t = 100 \text{ nm}$, $h = 100 \text{ nm}$, $T = 0^\circ$
 Variables : $CR_b = [0, 2, 5, 10, 15, 20, 30] \text{ nm}$
 $CR_t = [0, 2, 5, 10, 15, 20, 30] \text{ nm}$
 $SWA = [75, 80, 82.5, 85, 87.5, 90]^\circ$

Sur la Fig 4.6 sont représentées 49 courbes correspondant à tous les couples possibles de CRb/CRt pour un SWA de 90° . Pourtant on ne peut distinguer que 7 faisceaux de courbes

car le CRb n'influence pas la zone du CRT et inversement le CRT n'influence pas la zone du CRb. C'est déjà une observation très importante : on pourra travailler sur le CRT et le CRb de manière indépendante, car leurs zones d'influence sur l'image SEM ne se superposent pas. Pour en être convaincu, une colorisation particulière a été utilisée : la partie gauche des courbes est colorée en fonction du CRb et on peut voir que dans la zone du CRT les couleurs sont mixées et inversement la partie droite est colorée en fonction du CRT et on peut voir que dans la zone du CRb les couleurs sont mixées.

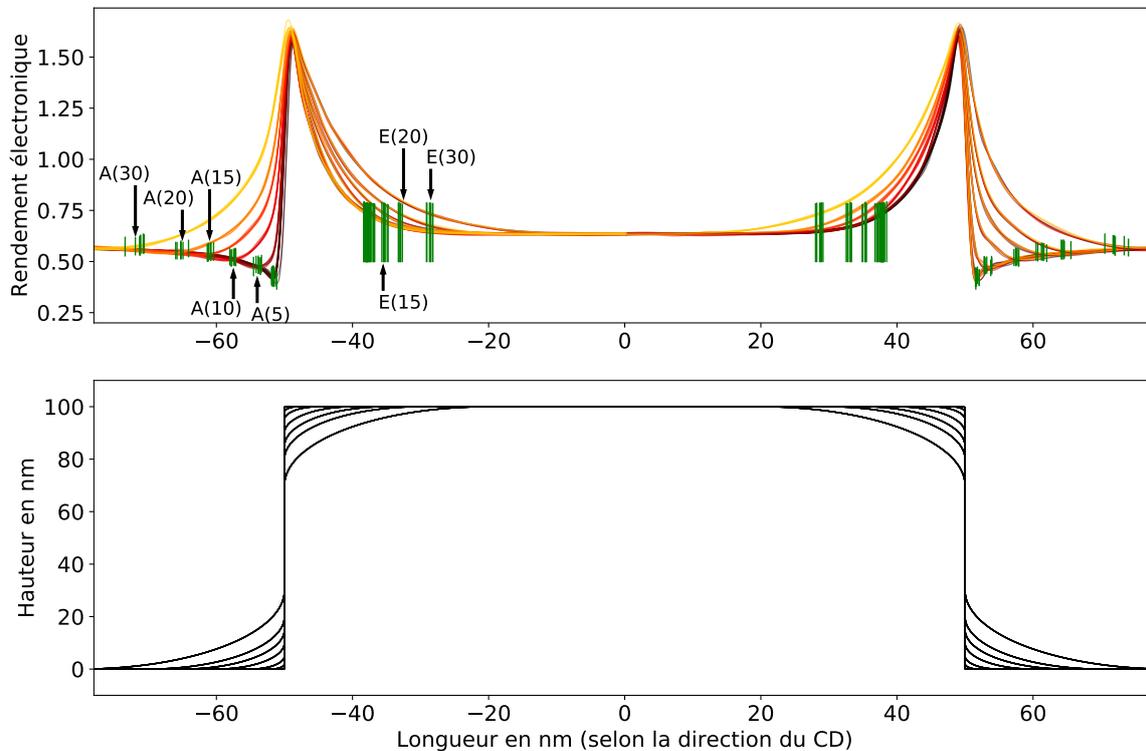


FIGURE 4.6 – Résultat de simulations JMONSEL de 49 motifs de 90° de SWA pour différents CRb et CRT. En bas sont représentés les motifs simulés vus en coupe.

Il est important de noter que la représentation des CR n'est pas circulaire sur la Fig 4.6 car l'échelle de hauteur et de longueur est différente pour pouvoir mettre en regard les vues en coupe avec les résultats JMONSEL, mais ce sont bien des portions de cercles qui sont simulées comme imposé par l'approximation 8 paramètres.

Grâce à cette expérience nous avons pu déterminer que les descripteurs A étaient intéressants pour la caractérisation du CRb, car ils sont bien séparés en fonction du CRb simulé. Cependant on peut également voir que lorsque le CRb est plus petit que 5 nm, il devient alors difficile de différencier les courbes entre elles.

Pour le CRT, c'est un peu plus compliqué car la zone d'influence du CRT est confondue avec l'effet d'augmentation progressive de la probabilité de sortie des électrons à mesure que le faisceau s'approche du bord du motif (ce que l'on va appeler effet de flanc). Cependant on peut voir que le descripteur E peut être utilisé au moins pour repérer les motifs à fort CRT.

Enfin on peut voir que les courbes sont globalement rangées les unes au dessus des autres en fonction de la valeur des CR (que ce soit en haut ou en bas). A partir de cette observation, il peut sembler intéressant d'utiliser l'aire en dessous des courbes comme métrique pour la détermination des CR. Cependant il s'agit d'une métrique plus complexe qu'une distance pour plusieurs raisons. Tout d'abord ici nous avons accès à plusieurs valeurs de CR ce qui nous permet de comparer les courbes entre elles ; nous faisons donc une observation relative aux autres, mais

dans la réalité nous n'aurons accès qu'à une mesure. Il faudrait alors que cette mesure soit absolue. Or, si dans le cas d'une simulation le rendement électronique peut être vu comme une valeur absolue, lorsque l'on utilisera de vraies images SEM la valeur absolue de niveau de gris pourra varier à cause d'un changement de contraste ou de brillance de l'image. Pour ne pas être influencée par la brillance de l'image (décalage constant du niveau de gris pour tout x) l'aire sous la courbe ne sera pas mesurée jusqu'au zéro mais jusqu'au point le plus bas de l'intervalle mesuré. En revanche cette métrique restera influencée par le contraste de l'image.

4.4.2 Influence des CR sur l'EW

La Fig 4.7 est une analyse de l'influence des CR sur la distance DB qui, on l'a vu dans la partie précédente, est la grandeur qui va permettre de remonter à l'information sur l'EW. Différentes informations peuvent être tirées de ce graphique, nous allons les détailler une par une.

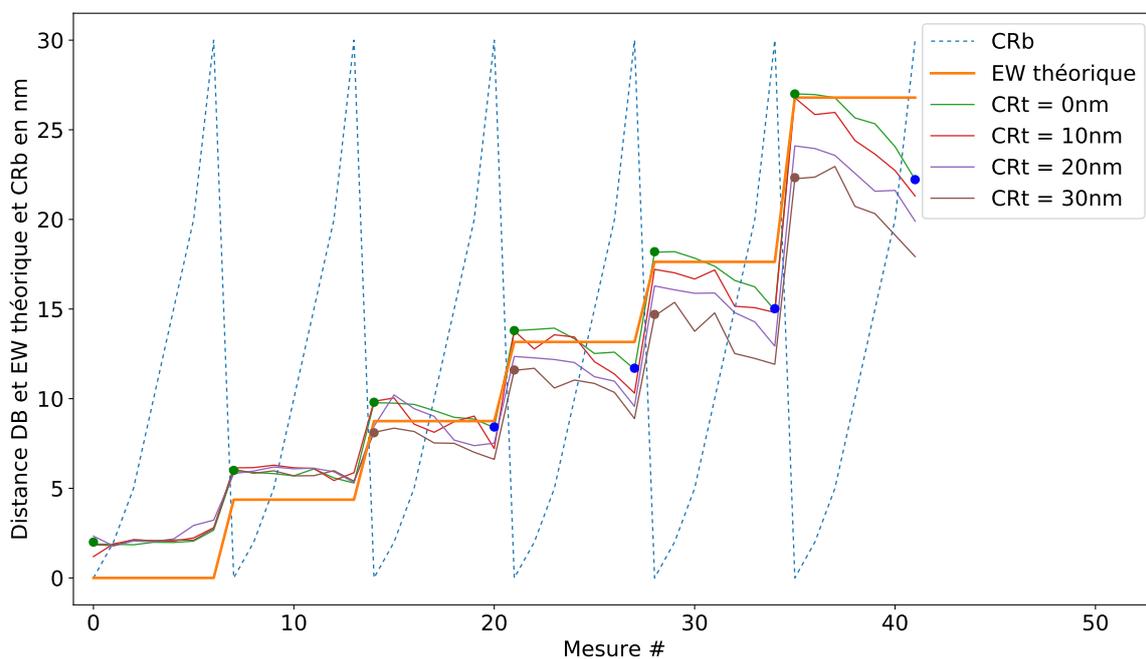


FIGURE 4.7 – Influence du CRb et du CRT sur la distance DB. La courbe en gras orange est indicative pour connaître la valeur d'EW théorique. La courbe en pointillé bleu est indicative pour connaître la valeur des CRb. Les autres courbes représentent les valeurs de distance DB en fonction du CRb (classement par palier) et du CRT (couleur). Les points de couleur sont des points particuliers de ces courbes explicités dans le texte.

En orange est représentée la valeur de l'EW théorique du motif observé. Comme les motifs simulés dans cette partie ont tous une hauteur de 100 nm et qu'il n'y a pas de tilt, la taille de l'EW change avec le changement de SWA. Les mesures ont été rangées de manière à les classer dans l'ordre croissant de taille d'EW théorique, puis pour chaque palier d'EW les mesures ont été classées par valeur croissante de CRb (représenté en pointillé bleu pour information). Les autres courbes représentent la distance DB en fonction du classement précédemment expliqué et en fonction du CRT (couleur).

Si on s'intéresse uniquement à la courbe dont le CRT est nul (verte) et le CRb est nul (représenté par les points verts) on retrouve le résultat de l'analyse h/SWA (Fig 4.4) ; Pour les valeurs élevées d'EW la distance DB est une bonne estimation des EW mais lorsque celui-ci diminue la distance DB tend vers une limite (autour de 2.5 nm).

Ensuite on peut voir sur chaque palier que lorsque le CRb augmente, la distance DB diminue car les courbes sont décroissantes sur la longueur du palier. Cet effet est de plus en plus marqué à mesure que la taille de l'EW augmente. Cela veut dire d'une part qu'il faudra corriger l'estimation de l'EW par la valeur du CRb mais également en fonction de la taille de l'EW elle-même.

Ensuite on peut observer que le CRt a la même influence et dans les mêmes proportions sur la distance DB que le CRb. En effet on observe également une diminution de la distance DB lorsque le CRt augmente et de façon plus marquée lorsque l'EW augmente (écart entre les courbes de couleur). De manière quantitative on peut voir que cette influence est à peu près identique car si l'on compare le dernier point de la courbe verte sur chaque palier (cad CRb = 30 nm et CRt = 0 nm (points bleus)) au premier point de la courbe marron sur chaque palier (cad CRb = 0 nm et CRt = 30 nm (points marrons)) on peut voir qu'ils sont situés à la même hauteur au bruit près. Pour en être convaincu, le même graphique a été tracé en classant chaque palier par CRt avec les couleurs représentant les CRb et on obtient bien un graphique similaire (non présenté ici).

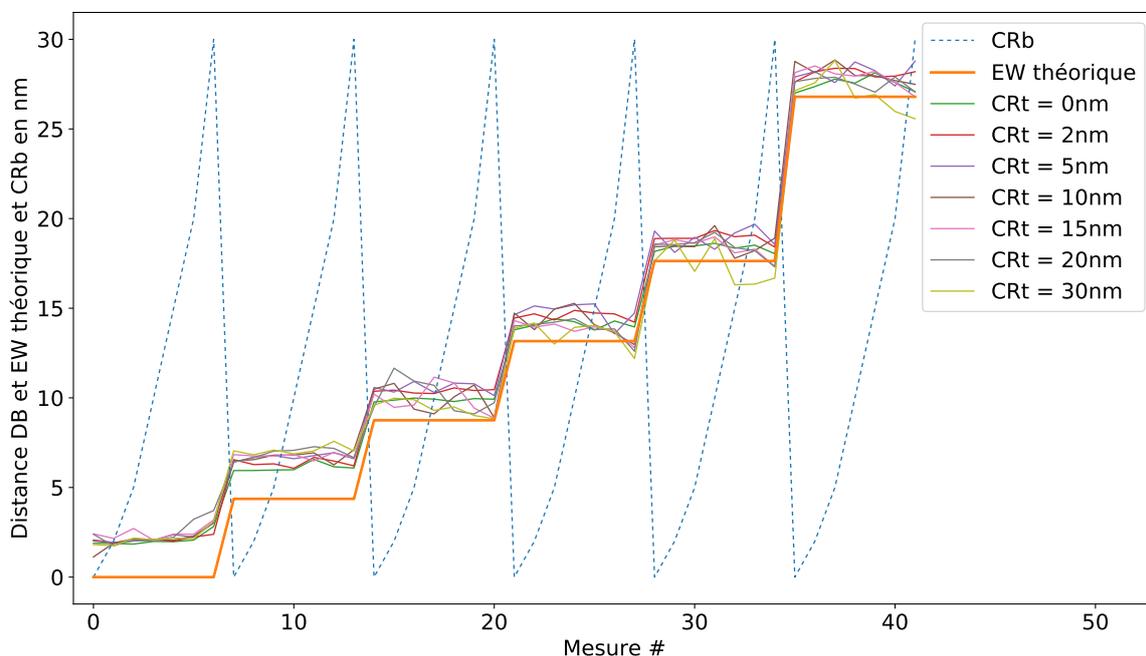


FIGURE 4.8 – Correction de l'influence des CR sur DB. Tous les CRt disponibles ont été affichés pour donner une idée du bruit lié à la simulation et à l'analyse.

D'après les observations faites grâce à la Fig 4.7, si l'on veut ramener la distance DB à CR variable à celle observée à CR nul on peut imaginer une correction du type :

$$DB_{correctionCR} = DB_{mesure} + a \times (DB_{mesure} - b)(CRt + CRb) \quad (4.4)$$

Où "b" est la valeur minimale vers laquelle tend la distance DB quand l'EW tend vers zéro et "a" un coefficient empirique lié à la réduction de DB en fonction du CRb et du CRt. Si on estime qu'il y a une différence suffisamment significative entre l'effet du CRb et du CRt sur la distance DB, on peut séparer les deux avec un coefficient pour chacun :

$$DB_{correctionCR} = DB_{mesure} + (DB_{mesure} - b)(a \times CRb + c \times CRt) \quad (4.5)$$

Si on utilise l'équation 4.4, avec $b = 2$ nm et $a = 0.008$ on obtient le résultat affiché en Fig 4.8. On peut voir que l'effet des CR a été en grande partie compensé (courbes redressées et couleurs rassemblées). On pourrait éventuellement affiner avec un ordre de correction plus élevé.

4.4.3 Influence des CR sur le CDt

La Fig 4.9 est une analyse de l'influence des CR sur la distance DD qui, on l'a vu dans la partie précédente, est la grandeur qui va permettre de remonter à l'information sur la largeur en haut du motif (CDt). Différentes informations peuvent être tirées de ce graphique, nous allons les détailler une par une.

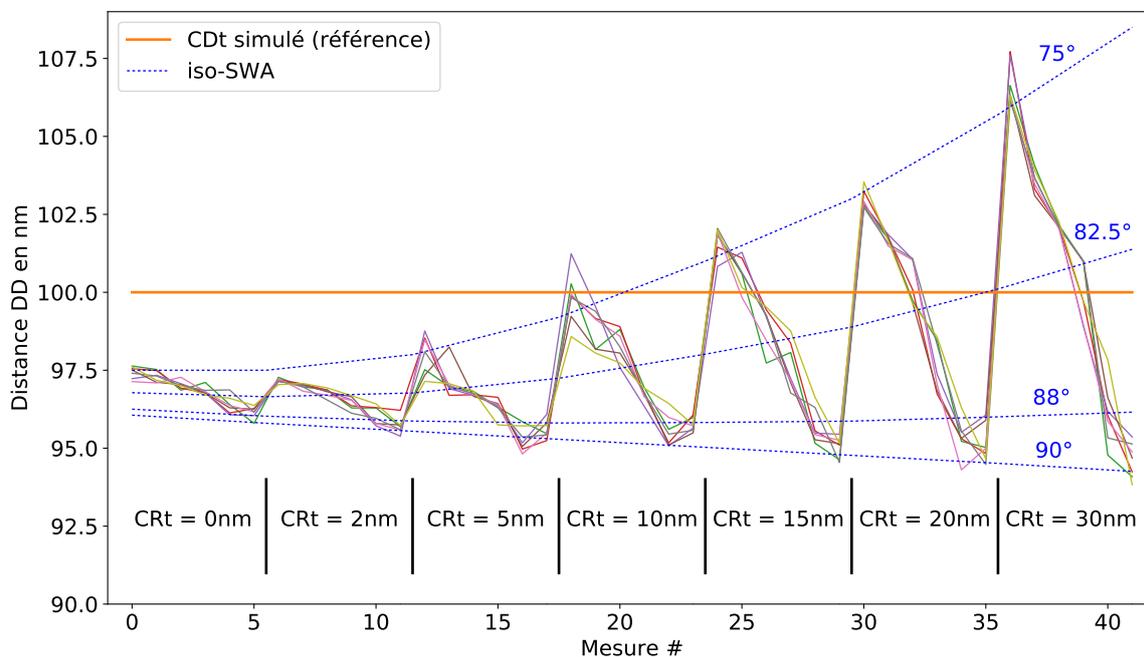


FIGURE 4.9 – Influence du CRb et du CRt sur la distance DD. La courbe en gras orange est indicative pour connaître la valeur du CD d'entrée simulé (100nm). Le faisceau de courbes de couleur correspond aux différentes valeurs de CRb. Les pointillés bleus correspondent à l'évolution de DD en fonction du CRt à un SWA donné, de 90° à 75°.

En orange est représentée la valeur du CDt des motifs simulés, soit 100 nm pour tous les motifs. Les mesures ont été rangées de manière à les classer dans l'ordre croissant de taille de CRt (cases), puis dans chaque case de CRt en fonction de SWA de 75° à 90°. Le faisceau de courbes de couleurs représente la distance DD pour différentes valeurs de CRb. On peut voir que la valeur du CRb n'influe pas sur la distance DD car les courbes ne sont pas ordonnées entre elles, c'est pourquoi le faisceau permet de se rendre compte du niveau de bruit attendu sur la mesure.

Ensuite si l'on s'intéresse seulement à la case CRt = 0 nm ou CRt = 2 nm on obtient le même effet que celui que l'on avait trouvé en Fig 4.2a, à savoir la diminution de la distance DD lorsque le SWA tend vers 90° avec un biais négatif de DD par rapport au CDt. Puis on observe que cet effet est fortement amplifié à mesure que le CRt augmente. Il est alors intéressant d'étudier les courbes iso-SWA, car l'augmentation du CRt va influencer de manière différente sur la distance DD en fonction du SWA du motif. Si l'on s'intéresse à la courbe iso-SWA la plus basse qui correspond à un SWA de 90°, on voit que l'augmentation du CRt va réduire la distance DD. A l'inverse la courbe la plus haute qui correspond à un SWA de 75°, montre une augmentation importante de la distance DD avec le CRt.

La correction du CRt sur la distance DD devra donc dépendre de la valeur du SWA. Lorsque le SWA $\approx 88^\circ$, le changement de CRT influe peu sur la distance DD, donc si on veut corriger la distance DD pour la ramener à sa valeur à CRt nul on devra effectuer une correction du type :

$$DD_{correctionCR} = DD_{correctionSWA} + a \times (88 - DD_{correctionSWA})(CRt - b) \quad (4.6)$$

Où $DD_{correctionSWA}$ est la distance DD dont on a corrigé l'effet du SWA comme vu dans la partie précédente (Eq 4.2), tel que :

$$DD_{correctionSWA} = DD_{mesure} + \text{biais} + a \times \exp(-P \times \tan(90 - SWA)) \quad (4.7)$$

Cependant on peut voir que l'effet exponentiel disparaît rapidement lorsque le CRt augmente, c'est pourquoi cette dernière correction sera finalement transformée en forme linéaire.

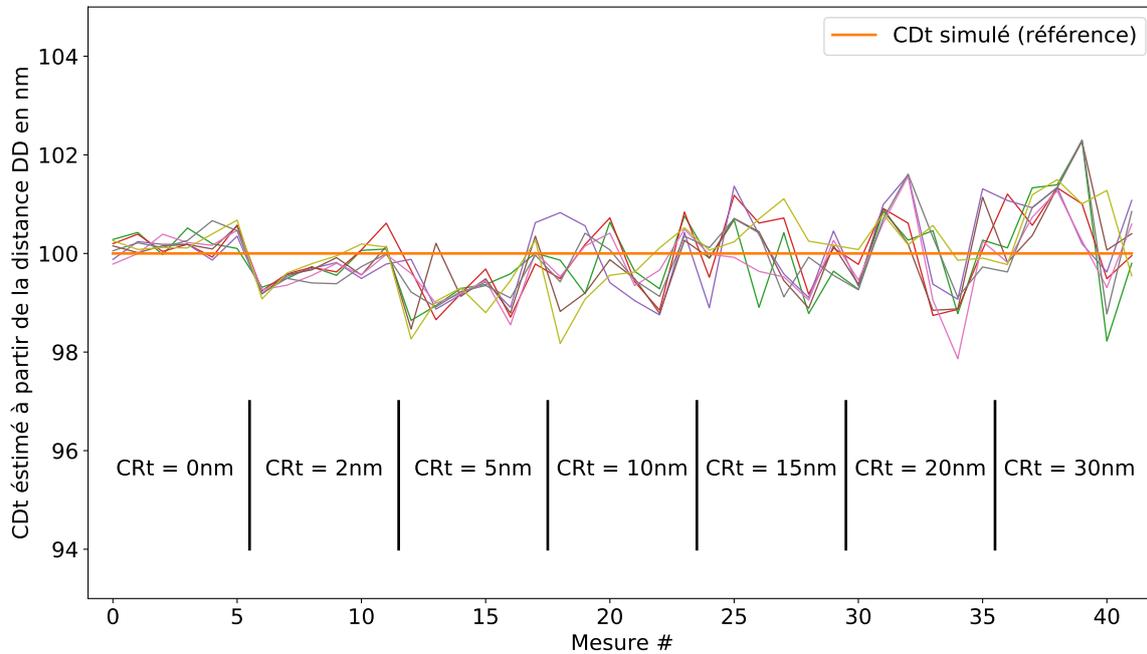


FIGURE 4.10 – Estimation du CDt grâce à une fonction de la distance DD prenant en compte la correction due à l'effet de SWA et de CRt combinés.

La fig 4.10 montre le résultat de cette estimation ; cependant on commence à voir que l'ajout de paramètres successifs rend l'optimisation de notre modèle de plus en plus compliquée. La mise en œuvre d'un algorithme d'optimisation du modèle général va donc devenir rapidement nécessaire. Cependant ces résultats de tendance d'évolution des grandeurs dans des sous-espaces de dimensions inférieures permettent de créer des fonctions de descripteurs linéaires ou non, que nous appellerons descripteurs secondaires et qui rendront le modèle plus fiable et pertinent d'un point de vue physique.

4.5 Limite de la méthode pour les petits CD

Le but de l'analyse "petit CD" est de séparer le régime dit "grand CD" du régime dit "petit CD". En effet pour que notre méthode de reconstruction fonctionne, les deux flancs doivent être suffisamment séparés, de manière à ce que leurs effets sur l'image SEM ne se superposent pas. Le régime "petit CD" sera donc un régime dans lequel la méthode présentée ne fonctionne plus car les flancs sont trop proches, c'est à dire que le CDt est trop faible. On analysera également si la hauteur et le SWA du motif ont une influence sur cette limite.

Dans cette analyse 132 motifs différents ont été simulés avec les grandeurs suivantes :

Constante : $CRt = CRb = 0 \text{ nm}$, $T = 0^\circ$
 Variables : $h = [50, 100, 150, 200] \text{ nm}$
 $SWA = [90, 87.5, 85]^\circ$
 $CDt = [5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100] \text{ nm}$

De manière à détecter ce changement de régime, deux métriques particulières ont été créées (aide visuelle dans la Fig 4.11 à droite dans l’encart) :

- h_{diff} : La différence de niveau de gris entre le milieu du motif et ”loin” du motif. Cette métrique est intéressante car lorsque le CD va tendre vers l’infini elle va tendre vers 0. Le régime ”petit CD” pourra donc être repéré par une augmentation de h_{diff} .
- gap : La longueur au centre du motif sur laquelle le niveau de gris est constant (distance entre les deux descripteurs E). Cette métrique est intéressante car lorsque le CD va tendre vers l’infini elle va tendre vers une asymptote tel que $gap = CD + cst$, où cst est une constante quelconque. Le régime ”petit CD” pourra donc être repéré par un décollement du gap de l’asymptote.

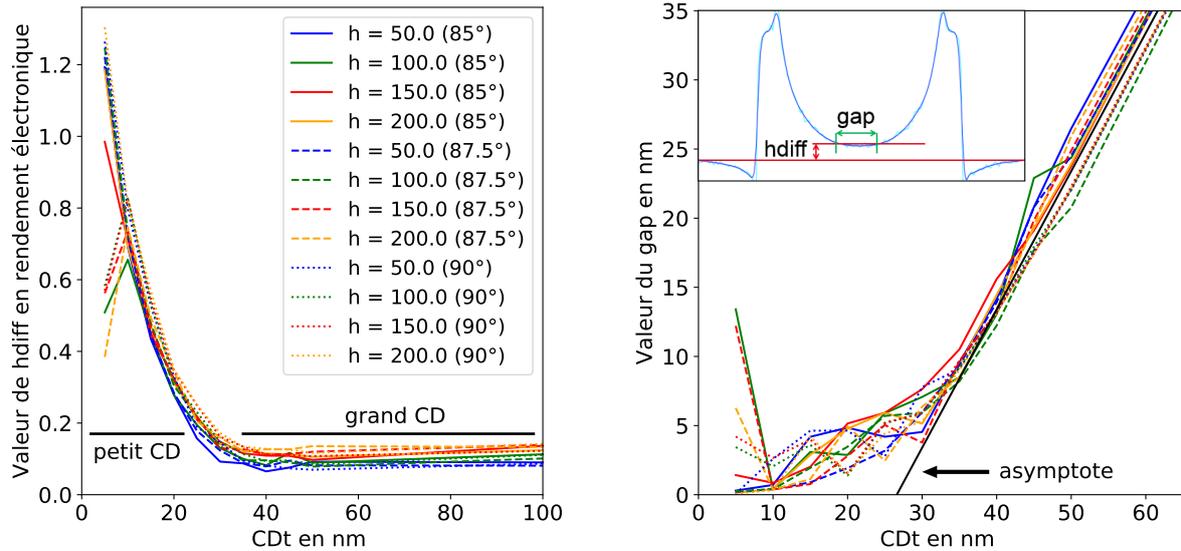


FIGURE 4.11 – Gauche : Evolution de h_{diff} lorsque le CDt diminue. Droite : Evolution du gap lorsque le CDt diminue. Le code couleur et type de ligne est le même sur les deux graphiques : bleu : $h = 50$ nm, vert : $h = 100$ nm, rouge : $h = 150$ nm, orange : $h = 200$ nm ; Trait plein : $SWA = 85^\circ$, tirets : $SWA = 87.5^\circ$, pointillés : $SWA = 90^\circ$.

Sur la Fig 4.11, on peut voir l’évolution de h_{diff} et du gap en fonction du CDt . Comme prévu on observe bien dans les deux cas une transition du comportement des métriques pour les ”grands CD” et pour les ”petit CD”, autour de 25 nm. Il n’y a pas une limite à partir de laquelle la méthode présentée ne va plus fonctionner, en revanche à partir de cette limite la probabilité pour que la méthode échoue dans le placement des descripteurs va être de plus en plus grande. A partir de l’étude de ces métriques (” h_{diff} ” et ” gap ”) et de notre expérience d’utilisation de la méthode, on peut déterminer que le placement des descripteurs n’échouera jamais jusqu’à un CD de 25 nm, puis la probabilité pour que le placement échoue augmentera progressivement jusqu’à devenir critique en dessous de 10-15 nm. Sur la Fig 4.11 on peut d’ailleurs voir que certains points sont désolidarisés du reste des courbes, c’est typiquement le cas lorsque le placement des descripteurs a échoué (ici pour $CDt = 5$ nm).

On n’observe pas de corrélation franche entre la limite ”petit CD” et la hauteur ou le SWA du motif car les courbes ne sont pas ordonnées. On peut tout de même observer que les courbes sont classées les unes au dessus des autres en fonction de la hauteur du motif (colorisation) en régime ”grand CD” sur le graphique de l’évolution de h_{diff} . Cependant ce n’est pas lié à l’effet ”petit CD”, mais à l’effet d’ombrage, du fait que notre point ”loin” n’est pas suffisamment loin. En effet si l’on prenait le point ”loin” du motif à une distance infinie, h_{diff} tendrait bien vers zéro et non vers une valeur légèrement supérieure à zéro comme c’est le cas ici. Notre point ”loin” du motif est donc toujours influencé par le motif et dans cette zone d’influence la hauteur du mo-

tif joue sur le rendement électronique. Cet effet sera analysé avec plus de détail dans la partie 4.7.

On peut donc conclure que cette méthode va être adaptée pour les motifs dont le CDt est supérieur à 20-25 nm. Cependant si l'on voulait pouvoir reconstruire les motifs dont le CDt est inférieur, il faudrait alors trouver des descripteurs différents qui permettraient de mieux décrire un motif de type "petit CD". On pourrait alors étendre la méthode qui détecterait automatiquement que l'on se trouve dans le cas "petit CD" et qui utiliserait donc cet autre jeu de descripteurs.

Enfin on pourrait également faire reculer le point limite en utilisant une énergie des électrons primaires plus faible ; en effet cela diminuerait la taille de la poire d'interaction et donc la taille de la zone d'influence des flancs sur l'image SEM.

4.6 Influence du tilt sur le signal SEM

Le but de l'analyse est de déterminer l'effet du tilt sur les différentes grandeurs d'intérêt. Il est important de noter que dans cette analyse, contrairement à toutes les précédentes, les effets vont être différents entre le côté droit du motif (côté du tilt) et le côté gauche du motif (côté opposé au tilt). Pour les motifs dont le SWA est proche de 90° , les informations du côté gauche vont être très rapidement détériorées puis disparaître, au profit du côté droit. On va donc s'intéresser aux grandeurs d'intérêt du côté du tilt (droit dans notre cas), dans le cas où la direction du tilt est orthogonale au flanc du motif (c'est à dire quand $\theta = 90^\circ$ dans l'Eq 3.3). Dans cette analyse 160 motifs différents ont été simulés avec les grandeurs suivantes :

Constante : CDt = 100 nm, CRt = CRb = 0 nm
 Variables : h = [50, 100, 150, 200] nm
 SWA = [92, 91, 90, 89, 88, 87, 86, 85] $^\circ$
 T = [0, 3, 5, 12, 15] $^\circ$

Les tilts simulés correspondent à ceux qui sont disponibles sur notre SEM, à l'exception du tilt 15° , auquel nous n'avons finalement pas eu accès lors de ce travail de thèse.

Avant de nous pencher plus en détail sur l'analyse des simulations, nous allons expliquer les différents phénomènes qui se produisent lorsque l'on incline le faisceau électronique de manière à ce qu'il arrive sur l'échantillon avec un angle T (Fig 4.12).

4.6.1 Etude théorique

Avec l'hypothèse que les deux flancs ont le même SWA, tant que $T < (90 - SWA)$ la taille global du motif va être inchangée car la diminution de l'EW de gauche (EWl) va être exactement compensée par une augmentation de l'EW de droite (EWr). Lorsque $T = (90 - SWA)$, le SWA du flanc gauche et du faisceau électronique sont parallèles et donc EWl est théoriquement (c'est à dire géométriquement) nul et EWr a doublé par rapport à sa valeur à tilt nul. Au delà, l'EWr va continuer à augmenter en suivant l'équation de l'EW 1, repoussant d'autant l'emplacement du flanc gauche.

La Fig 4.12, représente le cas où $T = 0^\circ$ (pointillés noirs) et $T > (90 - SWA)$ (rouge). Comme le faisceau électronique balaye l'échantillon à une certaine vitesse (notée V_0) en envoyant un flux régulier d'électrons alors, le nombre d'électrons reçus par unité de surface du motif observé sera différent en fonction de la topographie locale. Or on peut facilement comprendre que, plus le nombre d'électrons par mètre est faible, plus le niveau d'information provenant de cette zone sera faible. On propose donc une approximation dans laquelle la vitesse du faisceau électronique projetée sur le motif est proportionnelle à l'inverse du niveau d'information. Cette approximation se justifie à la fois par la réduction du nombre d'électrons déposés par unité de surface, mais

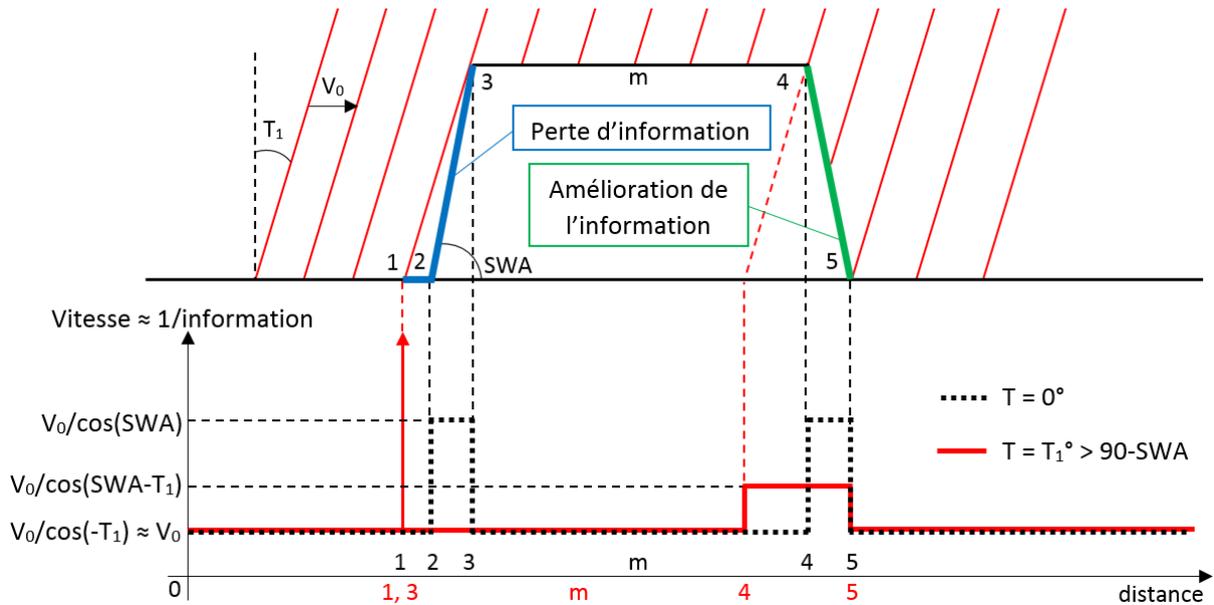


FIGURE 4.12 – Effet du tilt sur le niveau d’information d’une image SEM en fonction de la zone du motif.

(Haut) : Motif représenté en coupe. Les points 1, 2, 3, m, 4 et 5 représentent des emplacements physiques du motif, le point 1 étant le seul qui varie en fonction de T_1 . En rouge sont représentés les faisceaux électroniques tiltés d’un angle T_1 , avançant à une vitesse V_0 .

(Bas) : Courbes de vitesse de balayage projetées sur le motif, en pointillés noirs pour une image en vue du dessus et en trait plein rouge pour un faisceau incliné. La flèche rouge correspond au Dirac de perte d’information entre le point 1 et 3 lorsque $T > (90 - SWA)$. L’emplacement de la représentation des points physiques du motif est donné avec le même code couleur.

également par l’effet de dilatation de la gaussienne du spot size dû à sa projection sur un plan non orthogonal, ce qui va engendrer un niveau d’information plus diffus (c’est à dire moins local). Dans les deux cas il s’agit d’un effet de projection qui évolue en fonction de l’inverse du cosinus du SWA moins le tilt.

Sur un motif à SWA proche de 90° , le tilt va fortement améliorer l’information provenant du flanc droit (car le cosinus varie fortement proche de 90°), l’information du flanc gauche va disparaître (le cosinus négatif au-delà de 90° , nous informant de la perte de sens physique) et l’information sur les zones planes sera très faiblement perturbée (car le cosinus varie faiblement proche de 0°).

4.6.2 Cas pratique de simulation

Pour des raisons de simplification de visualisation, les simulations seront présentées en les ré-alignant sur le flanc gauche (Fig 4.13), ce qui permet de conserver la partie haute du motif au même emplacement. Dans le cas de l’utilisation d’un vrai SEM (non présenté ici), l’algorithme d’alignement essaye en général de centrer le motif dans sa globalité on se retrouve donc dans un état intermédiaire aux deux précédemment cités tel que : centrage = $(x_m + x_m)/2.0$.

Dans la Figure 4.13, on peut donc voir l’impact du tilt sur l’image SEM. Nous avons choisi de présenter un unique motif dont le SWA est de 86° car cela permet d’observer tous les cas précédemment décrits de manière théorique et de voir l’implication pratique sur le rendement électronique.

Partons de la courbe noire qui représente le motif observé en vue du dessus et qui nous est

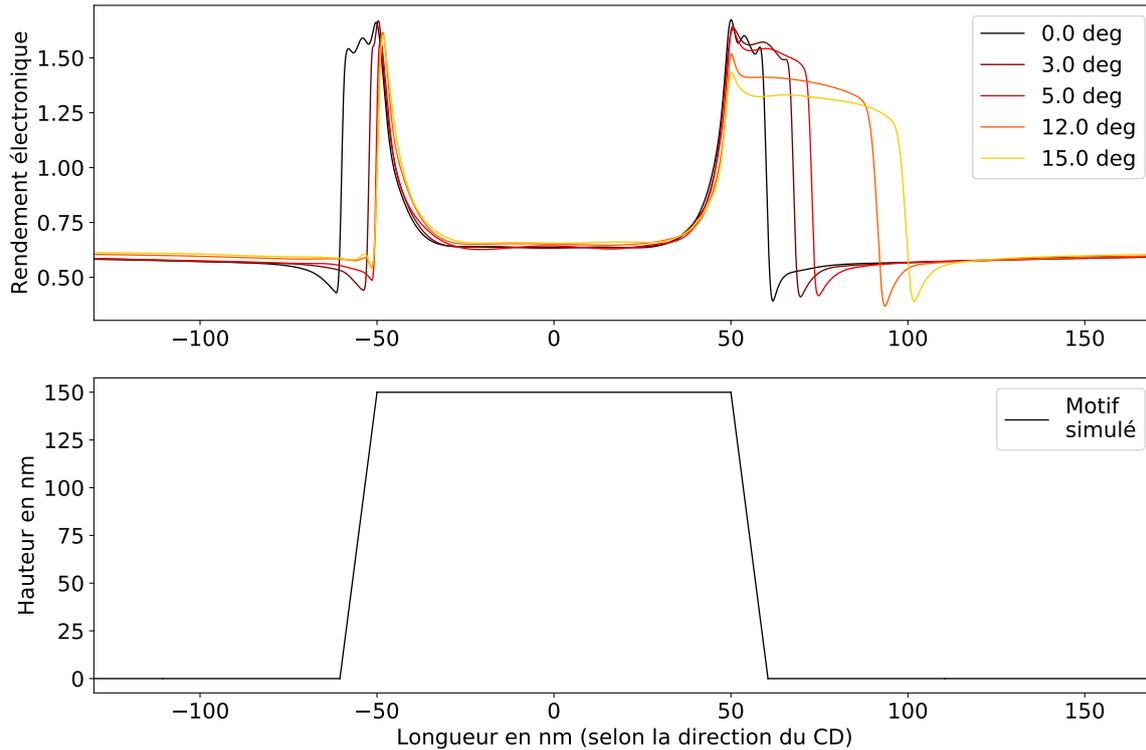


FIGURE 4.13 – Résultat JMONSEL d’un motif unique imagé à 5 tilts différents. Le faisceau arrive de la droite.

désormais familière. Lorsque l’on incline le faisceau de 3° (courbe marron), on peut voir que la taille du flanc de droite (EW_r) est augmentée et comme $T < (90 - SWA) \Leftrightarrow 3 < 90 - 86$, on observe une diminution du flanc gauche, mais son effet n’a pas été réduit au minimum. Comme prévu dans la partie théorique précédente, on peut voir que la distance entre le front de la courbe noir et le front de la courbe marron du coté droit du motif est égale à cette même distance du coté gauche. A partir du tilt 5° on se trouve dans le cas où $T > (90 - SWA)$, le flanc de gauche est donc réduit à son minimum (lié à l’effet de flanc), alors que du coté droit l’EW_r continue d’augmenter à mesure que le tilt augmente.

Tout comme dans la partie 4.3, on peut voir l’effet de diminution du rendement électronique lorsque la pente diminue de type loi de Lambert. Cependant ici ce n’est pas la pente intrinsèque du motif qui baisse mais l’angle relatif entre le faisceau électronique et la surface. Il serait intéressant d’étudier si l’impact du tilt est égal à l’impact du SWA sur le rendement électronique (notamment dans les cas limites proches de 90°), mais cette étude n’a pas été faite car nous voulions utiliser le moins possible l’information des valeurs absolues de niveau de gris.

On va maintenant s’intéresser plus spécifiquement au niveau de gris au pied du motif du coté gauche. Pour les valeur de tilt tel que $T < (90 - SWA)$, c’est à dire $T = 0^\circ$ et $T = 3^\circ$ dans cet exemple, on voit que l’effet de bas de flanc ressemble à l’effet classique (il sera détaillé dans la partie suivante 4.7). Cependant pour les tilts plus élevés on voit que la zone de rendement électronique plus faible disparaît à mesure que le tilt augmente. Ceci est dû à l’effet décrit plus tôt de l’éloignement du point que l’on a nommé ”1” dans la Fig 4.12. En effet le faisceau passe alors directement du rendement en haut du motif au rendement ”loin” du motif. Le faisceau électronique évite donc physiquement la zone où le rendement électronique est plus faible. C’est un effet intéressant, mais en pratique il sera difficilement utilisable pour la reconstruction

topographique du motif car faible d'une part et d'autre part placé à un endroit où l'on rencontre une faiblesse de la spline. En effet comme la pente de la courbe est très forte puis s'arrête net, la spline réagit avec un certain amortissement et subit un dépassement de la valeur de consigne avant de se stabiliser (comme la réponse d'un système du deuxième ordre (circuit RLC) à un échelon de tension), qui est visible dans le Fig 4.12 sur les tilts 12° et 15° ; la donnée sera donc faussée. Il serait possible d'explorer cette problématique, mais encore une fois nous voulons éviter d'utiliser les valeurs absolues de niveau de gris et l'information qu'il est possible d'extraire de cet effet ne semble pas majeure.

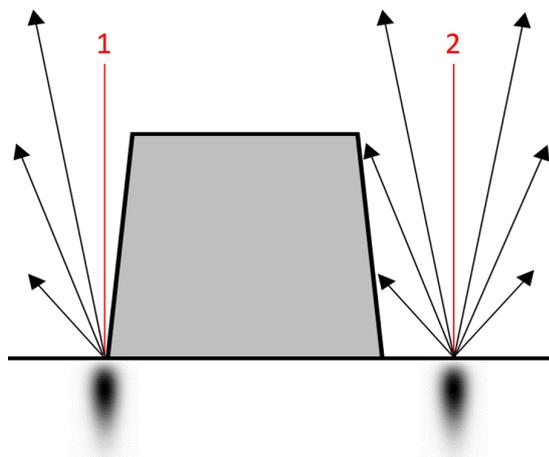
4.7 Analyse de l'effet d'ombrage

Le but de l'analyse de l'effet d'ombrage est d'expliquer son origine puis de voir si cet effet est relié à la géométrie du motif observé. Tout d'abord la définition que l'on a donné à l'effet d'ombrage dans ce travail est la zone de rendement électronique inférieur au rendement électronique d'une surface plane, que l'on observe en bas des flancs des motifs.

Dans cette analyse 48 motifs différents ont été simulés avec les grandeurs suivantes :

Constante : $CD_t = 200 \text{ nm}$, $CR_t = CR_b = 0 \text{ nm}$, $T = 0^\circ$

Variables : $h = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 400, 500] \text{ nm}$
 $SWA = [90, 85, 80]^\circ$



L'origine de l'ombrage est (1) d'une part l'angle solide d'échappement des électrons qui est faible ($180 - SWA$) lorsque le faisceau est proche du bas du flanc motif, puis (2) à plus longue distance par la probabilité de re-captation des électrons secondaires par le motif, en fonction de la topographie de ce motif (Fig 4.14).

FIGURE 4.14 – Représentation schématique de l'effet d'ombrage.

Tant que l'on ne se trouve pas dans un cas de très "petit CD" (tellement petit que les électrons traverseraient le motif de part en part), le CD n'aura pas d'influence sur l'effet d'ombrage. L'effet du CR_t est également négligeable sur l'effet d'ombrage.

On a vu dans la partie 4.4 sur les CR, que le CR_b avait une forte influence sur la zone de bas de motif (Fig 4.6), faisant rapidement disparaître l'effet d'ombrage à courte distance (lié à l'angle solide). En revanche le CR_b devrait faiblement influencer sur l'effet d'ombrage à longue distance. C'est pourquoi on a choisi de faire l'expérience à CR_b nul car il est inutile à longue distance et risque de nous perturber à faible distance (discutable mais il faut faire des choix).

La Fig 4.15 permet d'observer le résultat JMONSEL de 16 motifs de 90° de SWA pour des hauteurs allant de 10 nm à 500 nm. Il est important de noter que la zone qui a été simulée dans cette expérience s'étend beaucoup plus loin du motif que dans les simulations précédentes; en effet on observe ici le rendement électronique jusqu'à une distance de $1.4 \mu\text{m}$. Pour bien observer l'effet d'ombrage, le résultat affiché correspond à la moitié droite du motif : centre du motif en $x = 0$, flanc à $x = 100 \text{ nm}$, puis zone en bas de motif après $x = 100 \text{ nm}$. Le pic lié à l'effet de flanc a été coupé pour ne pas réduire la précision sur la zone de bas de motif.

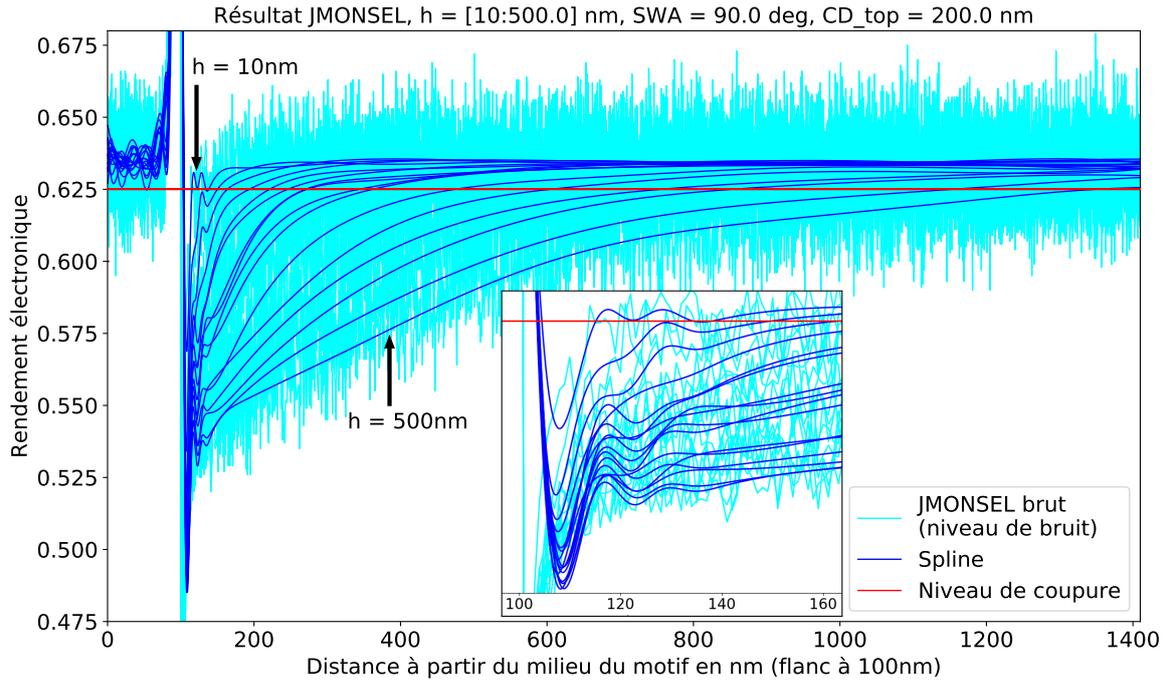


FIGURE 4.15 – Résultat JMONSEL de 16 motifs de hauteurs différentes permettant d’observer l’effet d’ombrage en fonction de la hauteur du motif à mesure que l’on s’éloigne de son flanc droit. Les splines utilisées pour l’analyse sont représentées en bleu, mais le résultat JMONSEL brut a été ajouté (cyan) pour se rendre compte du niveau de bruit pixel à pixel. Le zoom permet de voir l’amortissement due à la spline dans les premiers nanomètres.

On peut voir que les courbes sont ordonnées en fonction de la hauteur du motif : Plus le motif est grand plus la zone de rendement électronique faible s’étend loin du motif. Ce résultat semble logique car plus le motif est grand, plus la probabilité que les électrons soient re-captés par le motif est grande à une distance donnée.

Dans la Fig 4.16, on peut voir que la zone d’influence du motif est même directement proportionnelle à sa hauteur. En effet si l’on choisit un niveau de coupure, on peut voir un rapport de proportionnalité entre la hauteur et la distance à laquelle la spline repasse au-dessus de ce niveau de coupure.

Le choix du niveau de coupure (Droite rouge dans la Fig 4.15) est fait de manière à être légèrement en dessous du niveau moyen de rendement électronique sur une surface plane infinie, mais ”dans le bruit de mesure” d’une surface plane infinie. Bien sûr, la valeur de coupure choisie a un impact énorme sur la valeur du coefficient de proportionnalité. En revanche on retrouve toujours la linéarité du phénomène avec la hauteur, au bruit près (qui peut être important quand on se rapproche du niveau moyen de rendement électronique sur une surface plane infinie). Avec le choix de valeur de coupure fait, on obtient une longueur d’effet d’ombrage égale à 2.6 fois la hauteur du motif environ.

Ce critère pourrait donc être utilisé pour considérer un motif comme étant ”isolé”. En effet si deux motifs sont espacés d’au moins 5.2 fois la hauteur de ces motifs, alors leur effet d’ombrage ne se superposeront pas (au sens du niveau de coupure), c’est à dire que pour une position donnée de faisceau électronique, les électrons ne seront jamais re-captés par les deux motifs en même temps. Bien que théoriquement la distance d’influence d’un motif soit infinie, une valeur de coupure choisie judicieusement peut permettre une approximation satisfaisante.

D’après la Fig 4.16, on voit également que la distance dépend (en moindre mesure) du SWA

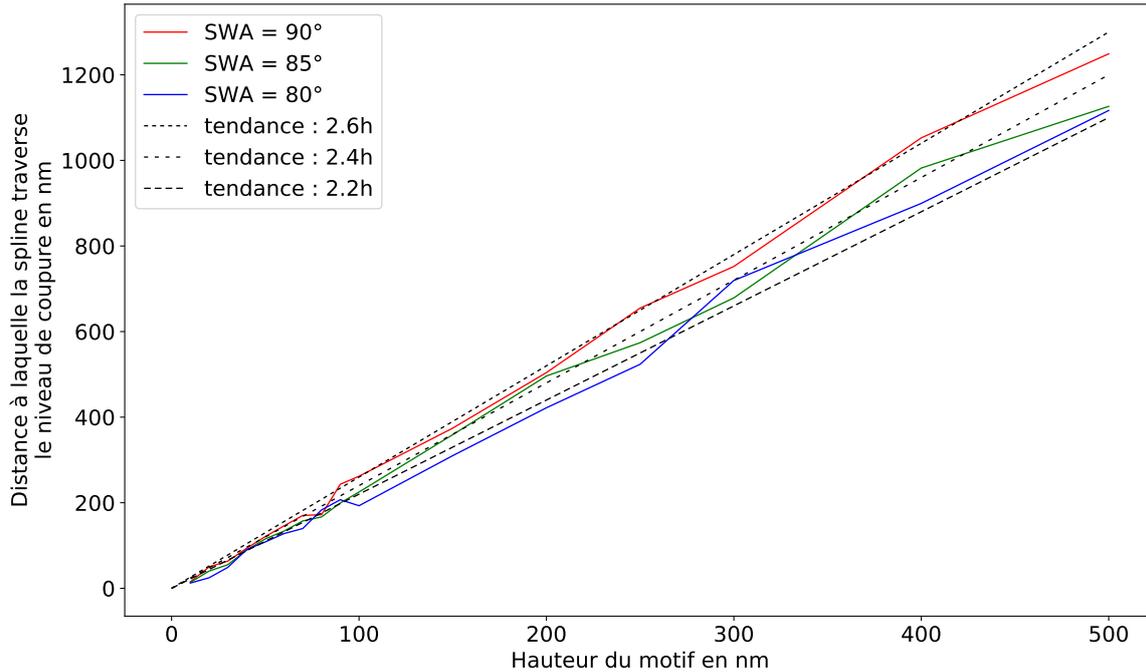


FIGURE 4.16 – Distance à laquelle la spline traverse le niveau de coupure en fonction de la hauteur du motif, pour différents SWA. En pointillés noirs sont représentées les approximations linéaires de l'effet.

car plus le motif s'écarte de 90° plus l'angle de libération s'agrandit. Il faut toutefois noter qu'il est compliqué de comparer les motifs à SWA différents car la "fin du flanc" n'est pas la même. Ici nous avons choisi de recalculer tous les points de bas de flanc au même niveau en x , mais cela implique que pour une position donnée du faisceau électronique, le point du haut du motif est plus éloigné pour les SWA plus faibles. La distance entre deux motifs à considérer "isolés" est donc à prendre en bas de motif.

Encore une fois, cet effet est lié à la valeur absolue de niveau de gris, ce qui n'est pas fiable pour l'estimation quantitative des grandeurs d'intérêts. De plus cet effet ne pourra pas être observé dans le cas de motifs de type "tranchée" ou "dense".

4.8 Conclusion

Grâce aux observations faites à partir des simulations JMONSEL, nous avons pu déterminer des tendances de variation de la position relative des descripteurs entre eux en fonction de la topographie des motifs ainsi que des conditions de prise d'image. A partir de ces observations nous avons pu définir des fonctions de descripteurs permettant de remonter à une estimation des grandeurs d'intérêt dans des sous-espaces limités de variations des topographies d'entrée. Désormais un modèle fonctionnant dans l'entièreté de l'espace des paramètres va être créé grâce à ces observations faites dans des sous-espaces.

Chapitre 5

Le modèle INVSEM3D : Création, calibration et reconstruction pour des lignes isolées en silicium

Dans le chapitre précédent, des lois empiriques ont été établies pour déduire l'impact des paramètres géométriques sur l'évaluation des grandeurs d'intérêt d'un motif de type ligne. Dans ce chapitre nous proposons une approche plus générique où les lois vont être régies par un modèle linéaire de fonctions de descripteurs.

Dans un premier temps la méthode de création d'un tel modèle va être décrite de manière générale, puis sa création pour des motifs de type ligne sera détaillée. Dans un second temps on présentera la calibration expérimentale de ce modèle sur des motifs réels en silicium gravés imaged par SEM et mesurés par AFM comme métrologie de référence. Enfin on présentera l'utilisation de ce modèle pour la reconstruction de motifs de type ligne à partir d'images SEM.

5.1 Création du modèle

5.1.1 Écriture du problème sous forme d'un problème linéaire

Comme on l'a vu, les fonctions de descripteurs que nous appellerons désormais descripteurs secondaires peuvent par exemple être :

- Des distances : $AB = |x_A - x_B|$, $DB = |x_D - x_B|$, $CE = |x_C - x_E|$...
- Une probabilité d'échappement des électrons de la matière : e^{-AB} , e^{-CD} ...
- Une aire sous le profil : $\int_A^B \text{profil}(x) dx$, $\int_D^E \text{profil}(x) dx$...
- Une valeur spécifique ou moyenne de la dérivée sur un intervalle.
- Une compensation due aux conditions de prise d'image (Tilt, accélération des électrons...) : $\tan(T)$.

Ces descripteurs secondaires peuvent être arrangés entre eux de manière non linéaire. En effet, plus les descripteurs secondaires auront un sens physique et un large domaine de validité, plus il sera possible de limiter la complexité du modèle.

Convention de notation : x : scalaire ; \mathbf{x} : vecteur ; X : matrice.
 x_i : i^{eme} valeur de x ; x_I : dernière valeur de x .

Le but de l'utilisation d'un modèle est de trouver les équations permettant d'estimer la valeur des grandeurs d'intérêt notée g à partir d'une combinaison linéaire composée de descripteurs notés d_k et pondérés par un coefficient associé w_k , ainsi qu'un éventuel biais noté w_0 , tel que :

$$g = w_0 + w_1 \times d_1 + w_2 \times d_2 + \dots + w_K \times d_K \quad (5.1)$$

Pour optimiser les coefficients w_k on se sert des valeurs des grandeurs d'intérêt de référence (paramètres d'entrée des simulations JMONSEL), ainsi que de la mesure ou de la connaissance des descripteurs. On va donc pouvoir construire un système de N équations dont on cherche les $(K + 1) w$, tel que :

$$\begin{cases} g_1 = w_0 + w_1 \times d_{1,1} + w_2 \times d_{1,2} + \dots + w_K \times d_{1,K} \\ g_2 = w_0 + w_1 \times d_{2,1} + w_2 \times d_{2,2} + \dots + w_K \times d_{2,K} \\ \dots \\ g_N = w_0 + w_1 \times d_{N,1} + w_2 \times d_{N,2} + \dots + w_K \times d_{N,K} \end{cases} \quad (5.2)$$

On peut écrire ces équations de manière matricielle :

$$\begin{pmatrix} g_1 \\ g_2 \\ \dots \\ g_N \end{pmatrix} = \begin{pmatrix} d_{1,0} = 1 & d_{1,1} & d_{1,2} & \dots & d_{1,K} \\ d_{2,0} = 1 & d_{2,1} & d_{2,2} & \dots & d_{2,K} \\ \dots & \dots & \dots & \dots & \dots \\ d_{N,0} = 1 & d_{N,1} & d_{N,2} & \dots & d_{N,K} \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ \dots \\ w_K \end{pmatrix} \quad (5.3)$$

soit,

$$\mathbf{g} = D \times \mathbf{w} \quad (5.4)$$

En pratique on se placera toujours dans une situation où le nombre de mesures est largement supérieur au nombre de descripteurs utilisés pour remonter à la grandeur d'intérêt ($N \gg K$), on parle de système sur-déterminé, permettant une meilleure estimation des paramètres. On ne va donc pas chercher la solution exacte du système mais une pseudo-solution qui va minimiser la norme au carré du résidu noté \mathbf{e} , avec $\tilde{\mathbf{g}}$ la valeur de référence de la grandeur et \mathbf{g} son estimation :

$$\tilde{\mathbf{g}} = D \times \mathbf{w} + \mathbf{e} \quad (5.5)$$

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} (\|\mathbf{e}\|^2) = \arg \min_{\mathbf{w}} (\|\tilde{\mathbf{g}} - \mathbf{g}\|^2) = \arg \min_{\mathbf{w}} (\|\tilde{\mathbf{g}} - D \times \mathbf{w}\|^2) \quad (5.6)$$

Pour trouver de manière algébrique les valeurs des facteurs w_k qui vont minimiser ce système linéaire, on utilise la matrice pseudo-inverse de Moore-Penrose [54], tel que :

$$\hat{\mathbf{w}} = (D^\top D)^{-1} D^\top \times \tilde{\mathbf{g}} \quad (5.7)$$

5.1.2 La régularisation

En amont de la méthode de résolution proposée ci-dessus, il convient de savoir quels descripteurs vont le mieux décrire la grandeur d'intérêt. En effet pour un lot de mesures donné, plus le nombre de descripteurs d'entrée va être grand, plus l'erreur relative sur la grandeur calculée va être faible. Cependant l'ajout de descripteurs faiblement descriptifs de la grandeur d'intérêt peuvent mener à une instabilité du modèle. Pour éviter ce problème de sur-apprentissage bien connu, nous allons donc à la fois chercher à minimiser la norme au carré du résidu mais en cherchant également à réduire la complexité du modèle par la prise en compte de la somme des valeurs absolues des coefficients (norme L1 de w), on parle de régularisation du modèle. Pour cela on va utiliser l'algorithme dit "LASSO" (Least Absolute Shrinkage and Selection Operator) qui va itérativement essayer de réduire la fonction de coût suivante :

$$Cost(\mathbf{w}) = \|\tilde{\mathbf{g}} - D \times \mathbf{w}\|^2 + \alpha \times \|\mathbf{w}\|_{L1} \quad (5.8)$$

$$Cost(w) = \sum_{n=1}^N \left\{ \tilde{g}_n - \sum_{k=0}^K w_k \times d_{n,k} \right\}^2 + \alpha \times \sum_{k=0}^K |w_k| \quad (5.9)$$

Cette régularisation par pénalisation de la norme L1, résultera en une réduction du nombre de descripteurs dont le coefficient w_k est différent de zéro, réduisant ainsi la complexité du modèle. La résolution de ce problème consiste donc à trouver le sous-ensemble optimal de paramètres parmi ceux proposés, permettant d’approcher au mieux la réponse, c’est à dire dans notre cas la valeur de la grandeur d’intérêt.

Afin de résoudre ce problème il existe différentes méthodes itératives, telles que :

- La sélection directe (forward selection ou forward stepwise regression), qui commence à partir d’un modèle vide (tous les $w_k = 0$) et essaye d’ajouter de manière itérative des paramètres [55][56] (on parle de descente de coordonnée (coordinate descent)). Les paramètres qui sont ajoutés au modèle en premier sont ceux dont la corrélation avec la réponse est la plus grande, ce qui est parfois reproché aux méthodes ”stepwise” [57][56][58] ; en effet la corrélation d’un paramètre avec la réponse peut être très différent en fonction du ”remplissage” du modèle.
- La régression inverse (backward elimination ou backward stepwise regression), qui commence à partir d’un modèle utilisant tous les descripteurs donnés et qui essaye d’enlever de manière itérative des paramètres. Les paramètres qui sont enlevés du modèle en premier sont ceux dont la corrélation avec la réponse est la plus faible.
- La régression bidirectionnelle (bidirectional elimination) qui est une combinaison des deux précédentes.
- La méthode LARS (Least Angle Regression) qui est similaire à la sélection directe (forward), mais qui, à la place d’optimiser les paramètres un par un, peut optimiser plusieurs paramètres en même temps (on parle de descente de gradient (gradient descent)) dans un direction n-dimensionnelle équi-angulaire avec les corrélations avec la réponse [58] [59].

De plus, la valeur du coefficient α doit être choisie de manière à obtenir un modèle qui correspond à nos attentes. En effet si α est nul, on se retrouve à nouveau dans un problème classique de minimisation des moindres carrés ; tous les paramètres seront gardés, résultant en un probable sur-apprentissage. A mesure que α augmente, l’importance donnée à la réduction de complexité du modèle va augmenter et de ce fait le nombre de paramètres gardés dans le modèle va décroître, au prix d’une augmentation du résidu, le but étant de trouver le point optimal.

Il existe des critères permettant un choix automatique de la valeur de α , basés sur la comparaison entre la diminution du résidu et l’augmentation de la complexité lorsque le nombre de paramètres augmente, c’est à dire quand α diminue. Les plus connus et fréquemment utilisés sont le critère d’information d’Akaike (AIC) et le critère d’information Bayésien (BIC).

$$\alpha_{AIC} = -\ln \mathcal{L}(\mathbf{w}) + K \quad (5.10)$$

$$\alpha_{BIC} = -\ln \mathcal{L}(\mathbf{w}) + K \ln(N) \quad (5.11)$$

où \mathcal{L} est la fonction de vraisemblance du résidu \mathbf{e} . Dans ce travail on considère que la répartition de l’erreur suit une loi Gaussienne $\mathcal{N}(0, \sigma^2)$, on a alors :

$$\mathcal{L}(\mathbf{w}) = \frac{1}{(\sigma\sqrt{2\pi})^N} \times \prod_{n=1}^N \left[\exp\left(-\frac{(\tilde{\mathbf{g}}_n - D_n \times \mathbf{w})^2}{2\sigma^2}\right) \right] \quad (5.12)$$

soit la fonction de log-vraisemblance suivante :

$$-\ln \mathcal{L}(\mathbf{w}) = N \ln(\sigma\sqrt{2\pi}) + \frac{1}{2\sigma^2} \sum_{n=1}^N [(\tilde{\mathbf{g}}_n - D_n \times \mathbf{w})^2] \quad (5.13)$$

On sait que le critère AIC tend à sélectionner le modèle avec une performance de prédiction optimale, tandis que le critère BIC tend à identifier les descripteurs principaux permettant

de créer un modèle dit parcimonieux ou creux (sparse model en anglais) [60]. Il existe un grand nombre d'autres critères de sélection (vous pouvez trouver une liste très complète dans [61]). En fonction des circonstances, certains critères sont plus consistants que les autres. Une procédure de sélection de modèle est dite consistante si la probabilité que le modèle trouvé soit égal au meilleur modèle (au sens des moindres carrés) tend vers 1 lorsque le nombre de mesures N tend vers l'infini [61]. Dans la publication de Shao citée, le commentaire de M.Zhang nous a encouragé à utiliser le critère BIC (traduction) :

"D'après les résultats de Shao, les critères de type BIC marcheraient mieux si le vrai modèle a une structure simple (dimension finie) et les critères de type AIC marcheraient mieux si le vrai modèle est compliqué (dimension infinie). Ces résultats sont mathématiquement indiscutables. En pratique, cependant, on peut avoir une interprétation différente. Un argument peut être émis en faveur des critères de type BIC quelque soit le vrai modèle. Tout d'abord, il faut comprendre que les modèles statistiques sont principalement utilisés dans des domaines où l'existence d'un vrai modèle est incertaine. Et même si un vrai modèle existe, il y a toujours de bonnes raisons de choisir un modèle simple plutôt que exact, sachant bien que le modèle sélectionné peut ne pas être le vrai. L'avantage pratique des modèles parcimonieux surpasse généralement le besoin d'exactitude du modèle. [...] Les critères de type BIC sont montrés comme étant optimaux du point de vue de la théorie de l'information."

On peut également citer les méthodes de validation croisée (cross-validation) qui permettent d'estimer si le modèle trouvé est instable par le découpage en plusieurs parties des mesures d'entrée. L'optimisation du modèle est réalisée sur un sous-ensemble des mesures d'entrée, puis le modèle trouvé est testé sur la partie des mesures qui n'a pas été utilisée pour l'apprentissage. Différentes stratégies de découpe et d'itération de la méthode peuvent être utilisées.

Il est important de noter que de légères différences dans les étapes d'analyse avant le modèle peuvent conduire à des résultats significativement différents de la part du modèle. On peut notamment noter :

- Le niveau d'ajustement (fit) de la spline.
- La stratégie de positionnement des descripteurs.
- La normalisation des mesures par rapport à leur nombre d'occurrence dans le lot.
- Le type de méthode de régularisation du modèle et de détermination de α .
- La liste d'entrée des descripteurs secondaires à tester.

5.1.3 Optimisation des paramètres pour un motif de type ligne

Après l'essai de différentes techniques de résolution du LASSO proposées par la librairie python "scikit learn" [62], la méthode LASSO LARS critère BIC a été choisie car elle proposait généralement les modèles les plus satisfaisants en sortie d'optimisation, c'est à dire présentant un résidu acceptable tout en gardant peu de descripteurs secondaires pour l'estimation des grandeurs d'intérêt.

Le modèle qui va être présenté par la suite, est l'un des nombreux modèles qui ont été produits lors de cette thèse, grâce à la méthodologie présentée. Il s'agit d'un modèle spécifique aux motifs de type ligne en silicium gravés. Cependant, bien qu'il s'agisse d'un des meilleurs modèles que nous avons été en capacité de produire, il est tout à fait probable qu'il soit encore grandement améliorable par l'optimisation des étapes d'analyse précédemment citées et notamment par l'ajout de nouveaux descripteurs secondaires plus pertinents. On peut noter que, théoriquement parlant, le meilleur modèle est en fait une formule : un unique descripteur de poids 1 décrivant entièrement la grandeur au bruit près. Ce descripteur contiendrait donc toute la complexité de l'effet à décrire, rendant le modèle non complexe au point d'être une formule.

Les Figs 5.3 à 5.8 permettent de voir le résultat de l'estimation des grandeurs d'intérêt faite à partir des simulations JMONSEL. Pour chacune de ces grandeurs d'intérêt la valeur de référence $\tilde{\mathbf{g}}$ est tracée en orange (c'est à dire la valeur d'entrée de la simulation) et les valeurs estimées \mathbf{g} par le modèle à partir des mesures faites sur le profil simulé via JMONSEL sont représentées en bleu. Pour chacune, la valeur du RMSE (Root Mean Square Error, Eq 5.14) entre l'estimation et la référence est donnée.

On rappelle que l'erreur quadratique notée MSE (Mean Square Error) correspond à la somme au carré de la différence entre l'estimation et la référence, moyennée sur le nombre de mesures (Eq 5.14).

$$\text{MSE} = \frac{1}{N} \times \sum_{n=1}^N (\mathbf{g}_n - \tilde{\mathbf{g}}_n)^2 \text{ et } \text{RMSE} = \sqrt{\text{MSE}} \quad (5.14)$$

Pour chacune des grandeurs d'intérêt, une vingtaine de descripteurs secondaires, dont on a estimé qu'ils pouvaient influencer la grandeur concernée, a été proposée. Seul quelques uns ont été gardés par la régularisation LASSO. On va le voir plus en détail, mais les effets mis en évidence dans les différentes analyses précédentes (Partie 4) ont généralement été conservés par la régularisation, qui a vu en eux une corrélation supérieure à tous les descripteurs "moins physique" que l'on a pu ajouter (comme des dérivés, des maximums de dérivé, des intégrales...).

Il est très important de noter que dans le processus d'estimation des grandeurs d'intérêt, l'ordre a une grande importance. On va commencer par estimer les descripteurs les plus faciles à estimer et/ou ceux dont l'estimation est la moins importante. En effet cela permettra pour les descripteurs suivants de profiter de la valeur des grandeurs d'intérêt précédemment estimées pour leur propre estimation.

Concrètement, les CR sont estimés en premier uniquement à partir des descripteurs secondaires, puis l'EW est estimé à partir des descripteurs secondaires ainsi que de la valeur du CRb et du CRt estimés. Grâce aux EW estimés, on calcule la hauteur et le SWA. Enfin Le CDt est estimé à partir des descripteurs secondaires, du CRt estimé et du SWA (Fig 5.1)

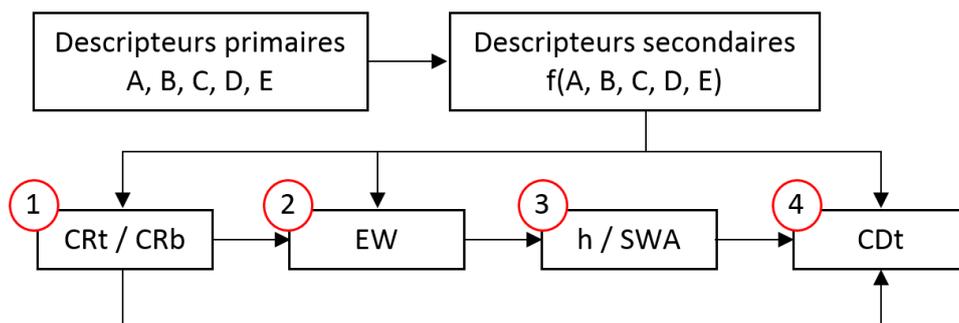


FIGURE 5.1 – Schéma représentant l'ordre d'estimation des grandeurs d'intérêt.

Corrélation des paramètres de simulation

Dans cette partie, nous allons expliquer une faiblesse de notre lot de mesures, c'est à dire des caractéristiques topographiques des motifs de nos simulations, de manière à comprendre les explications qui vont suivre en profondeur. On avait déjà rapidement abordé l'idée que, dans le lot de simulations, les grandeurs d'intérêt devaient varier le plus indépendamment possible. En effet si les grandeurs varient "ensemble", il devient alors difficile voire impossible de déterminer l'origine d'un effet observé.

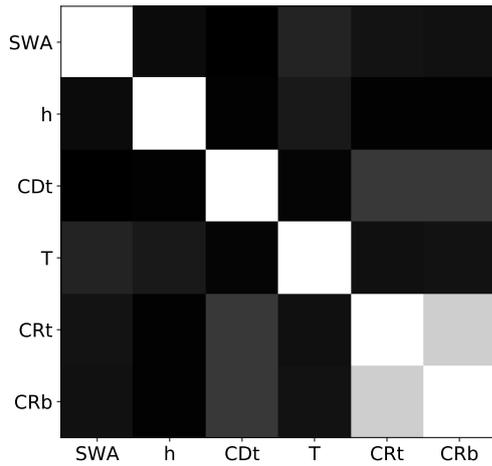


FIGURE 5.2 – Matrice de corrélation des grandeurs d'intérêt $(\tilde{\mathbf{g}}_u, \tilde{\mathbf{g}}_v)$.

La Fig 5.2 ci-contre représente de manière graphique la matrice de corrélation (Eq 5.15) des différentes grandeurs d'intérêts, sur les $N = 1123$ simulations JMONSEL dont on va réaliser l'optimisation. Un pixel blanc correspond à une corrélation parfaite entre deux grandeurs (valeur 1), un pixel noir correspond à l'absence totale de corrélation (linéaire) entre deux grandeurs (valeur 0). La matrice de corrélation idéale dans notre cas, serait une matrice nulle (avec une diagonale de 1 ce qui est toujours le cas dans une matrice de corrélation car une grandeur est toujours corrélée à 100% avec elle même). Avec une telle matrice les différentes grandeurs varieraient indépendamment les unes des autres, il ne serait donc pas possible de confondre l'effet de deux grandeurs.

$$\text{CORR}(\tilde{\mathbf{g}}_u, \tilde{\mathbf{g}}_v) = \frac{\text{COV}(\tilde{\mathbf{g}}_u, \tilde{\mathbf{g}}_v)}{\sigma_{\tilde{\mathbf{g}}_u} \sigma_{\tilde{\mathbf{g}}_v}}, \text{ avec } \text{COV}(\tilde{\mathbf{g}}_u, \tilde{\mathbf{g}}_v) = \frac{1}{N} \sum_{n=1}^N (g_{u,n} - \bar{\mathbf{g}}_u)(g_{v,n} - \bar{\mathbf{g}}_v) \quad (5.15)$$

où $\mathbf{g}_u = [\text{SWA}, h, \text{CDt}, T, \text{CRt}, \text{CRb}]$ et indépendamment $\mathbf{g}_v = [\text{SWA}, h, \text{CDt}, T, \text{CRt}, \text{CRb}]$

On voit que la matrice de corrélation de nos simulations présente une faiblesse : Le CRb et le CRt sont fortement corrélés (0.8). Cela veut dire qu'il est possible de confondre les effets du CRb et du CRt. Heureusement, on a déjà observé que les zones d'effet du CRt et du CRb sur l'image SEM ne se superposaient pas, dans un sous-ensemble de simulations dans lequel le CRt et le CRb étaient faiblement corrélés (partie 4.4). On a donc utilisé cette connaissance pour empêcher le modèle de faire une erreur : Si l'on avait donné tous les descripteurs possibles pour l'estimation du CRt par exemple, il est possible que la régularisation ait gardé un descripteur secondaire lié au CRb, à cause de la corrélation entre ces deux grandeurs. Pourtant ce descripteur n'aurait pas eu de sens physique et aurait conduit à une erreur d'estimation sur un lot différent de simulation ou de mesure SEM réelles. C'est pourquoi, connaissant ce problème, nous n'avons pas autorisé le modèle à utiliser des descripteurs non représentatifs d'une grandeur, pour les CR notamment.

Cette corrélation pourrait tout de même nous gêner pour répondre aux questions de type : "Cette grandeur est-elle plus influencée par le CRt ou le CRb ?"

Les 1123 simulations sur lesquelles on s'appuie pour créer le modèle ont été faites petit à petit de manière à repérer les effets (partie 4). De ce fait les CR avaient souvent la même valeur (et souvent zéro pour ne pas être perturbé par les CR) pour éviter de devoir simuler un nombre trop important de cas. En effet quand on teste un certain domaine, qu'on a déjà essayé de discrétiser de manière à réduire le nombre de simulations, ajouter une valeur supplémentaire dans un des espaces entraîne une multiplication par deux du nombre de simulations à faire. Avec le recul on aurait dû essayer de faire varier les CR de manière aléatoire et surtout de décorrélérer au maximum le CRb du CRt. On pourrait améliorer notre compréhension des effets en ajoutant d'autres simulations avec des valeurs de CR aléatoires pour réduire cette corrélation. Les autres corrélations sont inférieures à 0.25 ce qui est très correct.

Optimisation du CRb

La première grandeur que l'on a décidé d'estimer est le CRb (Fig 5.3) car il est assez facile à estimer et il sera très utile de connaître sa valeur pour la future estimation de l'EW. Le descripteur qui a été utilisé pour son estimation est la distance AB car comme on l'a vu dans la Fig 4.6, le descripteur A permet une très bonne description du CRb.

Le calcul utilisé par le modèle pour estimer le CRb est donc le suivant :

$$\text{CRb}_{\text{JMO}} = 1.52 \times \text{AB} - 1.90 \quad (5.16)$$

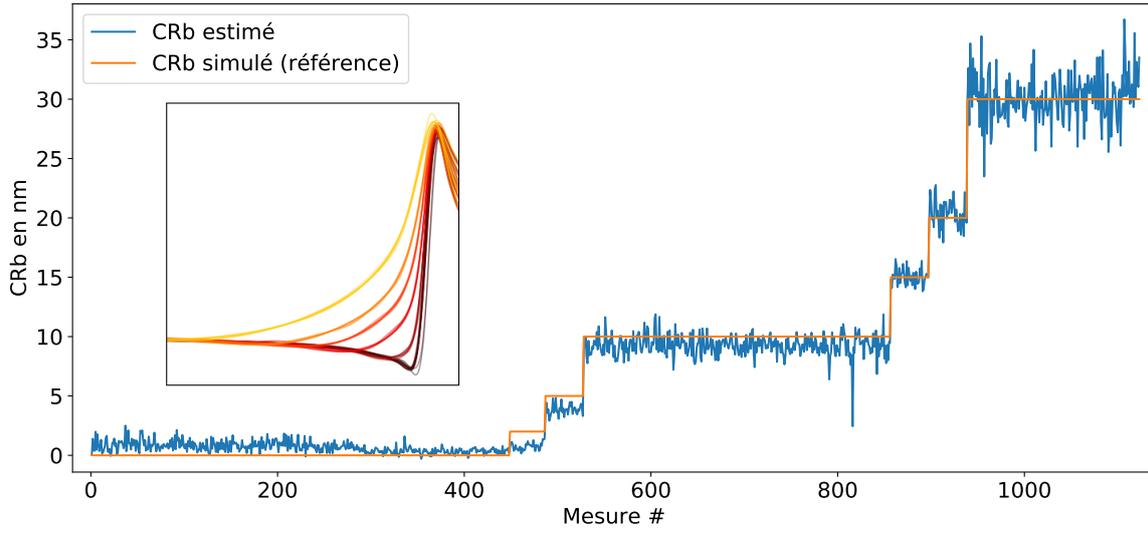


FIGURE 5.3 – Résultat de l'optimisation du CRb. RMSE = 1.23 nm.

On avait noté toutefois que lorsque le CRb diminuait en dessous de 5 nm environ il devenait difficile de différencier les courbes entre elles. On peut ici voir le même résultat, l'estimation des $\text{CRb} = 0 \text{ nm}$ et $\text{CRb} = 2 \text{ nm}$ donnent des résultats équivalents. En utilisant ce modèle, si on trouve un CRb inférieur à 3 nm, alors on pourra seulement conclure que le CRb est compris entre 0 nm et 3 nm.

Ensuite on peut voir que, plus le CRb est grand, plus l'incertitude de l'estimation sera grande. Cela peut être expliqué par l'effet croisé du CRb sur le rendement électronique, du bruit électronique et de l'approximation par une spline. L'encart dans la Fig 5.3 permet de bien comprendre cet effet et provient de la Fig 4.6 précédemment étudiée.

On sait que l'imagerie électronique est caractérisée par un bruit lié à la probabilité des électrons de faire telle ou telle interaction. La simulation physique n'échappe pas à cette règle (on peut même dire qu'elle fait en sorte de ne pas y échapper). Malgré l'envoi d'une grande quantité d'électrons par pixel (10000e-), il subsiste un bruit dont on a déjà pu apprécier l'importance dans la Fig 4.1 avec nos conditions de simulations. Or si ce bruit oscille par hasard à une fréquence inférieure à la fréquence de coupure du bruit que l'on a imposé à la spline par son niveau de fit, alors la spline se met à osciller avec le bruit qui est alors considéré comme du signal. Cet effet a moins de chance d'arriver sur une distance courte et a une importance relative moindre sur des zones pentues (dérivée élevée). Or l'augmentation du CRb augmente la taille de la zone d'influence et diminue la valeur de la dérivée sur cette zone (Fig 4.6 et encart 5.3). On a donc une chance plus élevée de "coincer" un descripteur dans une oscillation non désirée à fort CRb qu'à faible CRb, d'où l'augmentation de l'incertitude avec le CRb.

Enfin on peut noter que l'estimation a été faite seulement sur le CRb de droite, à partir seulement de la distance AB de droite, car cette estimation marche quelque soit l'angle de tilt ; on ne peut donc pas utiliser l'information du coté gauche qui est progressivement perdue avec l'augmentation de l'angle de tilt. Parmi les 1123 simulations présentées, plus de la moitié ont un tilt non nul. De cela découle deux choses : premièrement si cette estimation fonctionne correctement quelque soit le tilt sur des vrais motifs de microélectronique, alors on pourra estimer un CRb par image du motif à plusieurs tilt puis moyenner ces résultats pour diminuer l'erreur d'un rapport $1/\sqrt{N}$, où N est le nombre d'images à différents tilt. De plus, dans un grand nombre de cas on peut faire l'hypothèse que les CRb de droite et de gauche sont égaux, on pourra donc utiliser la distance AB de droite et de gauche en vue du dessus (tilt=0°) réduisant encore les erreurs.

Optimisation du CRt

La seconde grandeur d'intérêt qui est estimée par le modèle est le CRt, car on aura besoin de l'estimation des deux CR pour estimer l'EW. Cependant comme on peut le voir sur la Fig 5.4, l'estimation du CRt est bien plus délicate que celle du CRb. Comme on l'avait déjà souligné dans la partie 4.4, cette difficulté vient principalement du fait que la zone qui est influencée par la taille du CRt est la zone dans laquelle on observe déjà l'effet de flanc.

Le calcul utilisé par le modèle pour estimer le CRt est le suivant :

$$\text{CRt}_{\text{JMO}} = 0.13 \times \text{CE}^2 - 1.36 \times \text{DE} \quad (5.17)$$

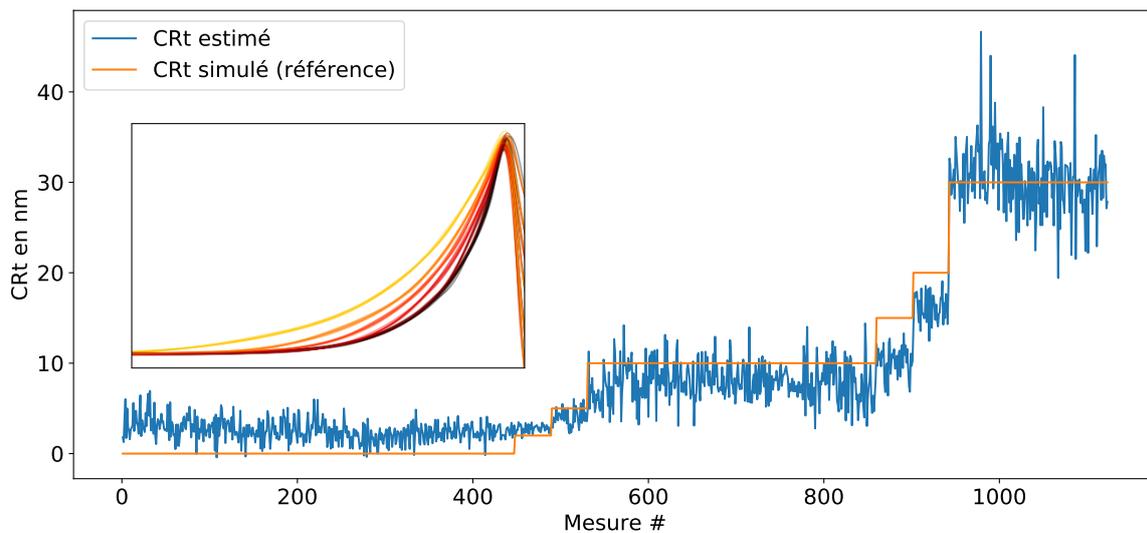


FIGURE 5.4 – Résultat de l'optimisation du CRt. RMSE = 3.13 nm.

Tout comme dans le cas précédent, on observe une augmentation de l'incertitude avec la valeur du CRt mais dans une moindre mesure, avec une incertitude déjà significative à CRt nul. L'effet est le même que celui décrit pour le CRb mais l'effet de flanc rend ce coté du pic moins raide et plus long que du coté du CRb même à CRt nul (Fig 4.6 et encart Fig 5.4).

Le niveau d'erreur de l'estimation du CRt va rendre sa mesure unique pratiquement inutilisable en dessous de 10 nm. Il faut donc prévoir de faire des moyennes d'une manière ou d'une autre pour obtenir un résultat décent (hypothèse d'égalité des CR sur plusieurs motifs, mesure à plusieurs tilt...). On peut toutefois noter que même si la valeur unique du CRt n'est pas utilisable comme valeur du CRt, elle peut tout de même être utile à la mesure d'EW qui vient

après ; Tout d'abord il faut noter que les CR partagent une frontière commune avec l'EW. De ce fait, si le CRT a été fortement surestimé il est probable qu'il ait empiété sur la zone de l'EW. Du coup l'EW risque à son tour d'être sous-estimé, c'est pourquoi on va voir que même mal estimée cette quantité va nous être utile. Bien sûr cet argument est probabiliste et on risque tout de même de voir une dégradation de la reconstruction topographique lorsque les CR sont grands.

Optimisation de l'EW

On va désormais pouvoir estimer l'EW en prenant en compte les valeurs des CR estimées (Fig 5.5). Dans les deux figures précédentes, les mesures étaient classées par ordre croissant de la grandeur d'intérêt simulée, mais ici on a choisi de les ranger d'abord par CRT croissant de manière à voir l'augmentation de l'erreur d'estimation pour les grands CR. Comme on vient de le voir, ranger les mesures par ordre croissant de CRT entraîne également un classement fort du CRb (c'est à dire que dans la partie CRT = 0 nm il y a beaucoup plus de cas où CRb = 0 nm que CRb ≠ 0 nm par exemple et de même quand CRT = 30 nm il y a beaucoup plus de cas où CRb = 30 nm que CRb ≠ 30 nm).

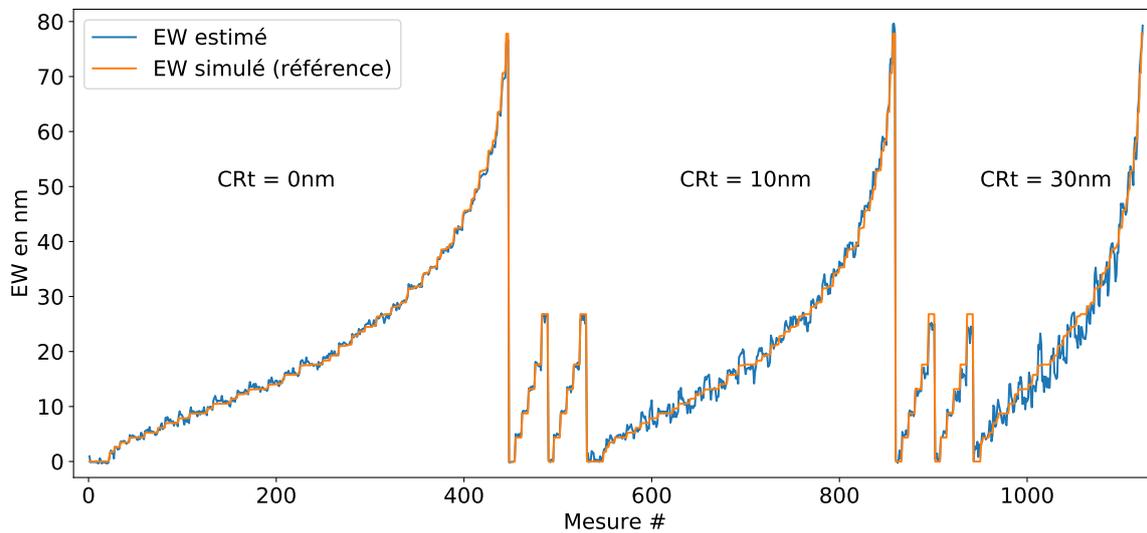


FIGURE 5.5 – Résultat de l'optimisation de l'EW. RMSE = 1.34nm.

Le calcul utilisé par le modèle pour estimer les EW est le suivant :

$$EW_{JMO} = 0.97 \times BD_2 + 0.0203 \times CReff1 + 0.0053 \times CReff2 + 0.0026 \times CReff3 \quad (5.18)$$

avec,

$$BD_2 = BD - 2.8 \exp(-0.16 \times BD) \quad , \quad CReff1 = CRb_{JMO} \times T \quad (5.19)$$

$$CReff2 = BD_2 \times CRb_{JMO} \quad , \quad CReff3 = BD_2 \times CRT_{JMO} \quad (5.20)$$

Comme on l'avait anticipé l'augmentation des CR rend les limites des flancs plus difficiles à repérer avec précision, menant à une augmentation de l'incertitude sur l'estimation des EW. Cette augmentation de l'erreur est résumée dans le tableau 5.1.

CRt simulé (nm)	0	2	5	10	15	20	30	tot
RMSE _{EW} (CR estimés)	0.52	0.46	0.58	1.21	1.38	1.85	2.54	1.34
RMSE _{EW} (CR ref)	0.56	0.65	0.72	1.23	1.29	1.72	2.45	1.32

TABLE 5.1 – Augmentation de l’erreur d’estimation sur l’EW avec le CRt (valeurs en nm). La première ligne correspond aux équations précédentes, la deuxième ligne étant le résultat lorsque l’on remplace les CR_{JMO} par leurs valeurs de référence.

Dans le tableau 5.1, on a choisi de représenter en gris les colonnes dont le nombre de simulations est sensiblement inférieur, ce qui entraîne un risque de bruit plus grand, à cause d’un moyennage sur trop peu de valeurs, que pour les colonnes en noir (facilement visible sur la Fig 5.5).

On voit que l’augmentation de la variance sur l’estimation de l’EW avec le CR est très significative. Comme la valeur de l’EW est estimée à partir de descripteurs utilisant les valeurs des CR estimés précédemment, on peut se demander si l’erreur observée à fort CR est principalement due à la précédente erreur faite sur les CR (on a vu qu’elle augmentait avec le CR). La dernière ligne du tableau correspond donc au même modèle, mais qui utilise la valeur simulée (c’est à dire de référence) des CR plutôt que les valeurs estimées. On peut voir que cela entraîne une légère diminution de l’erreur d’estimation sur l’EW notamment pour les CR élevés, mais pas de façon significative. L’erreur sur l’estimation de l’EW est donc due principalement à l’effet des CR sur la distance BD, qui est le descripteur principal pour l’optimisation de l’EW.

Calcul de la hauteur et du SWA

A partir des valeurs d’EW estimées on peut utiliser les équations de la Fig 2.4 pour calculer la hauteur et le SWA des motifs.

On aurait pu décider de traiter la hauteur et le SWA comme des grandeurs d’intérêts normales, que l’on aurait donc estimées par une combinaison linéaire de descripteurs dont un ou plusieurs auraient contenu des morceaux des équations citées. Cependant au stade de la simulation, comme les données d’entrée sont parfaites, si on fait tendre l’estimation d’EW vers l’EW théorique, alors la hauteur et le SWA tendront vers leur vraie valeur. On verra par la suite lors de l’application à un cas réel que les imperfections du SEM peuvent remettre en cause cette vision et entraîner un traitement différent de la hauteur et du SWA (Partie 5.2).

Premièrement il faut noter que toutes les simulations n’ont pas pu être utilisées dans cette partie, car pour reconstruire la hauteur et le SWA il faut avoir l’information de l’EW sur deux images (ou simulations d’images ici) du même motif à deux angles de tilt différents. Or seuls 204 motifs ont été simulés à quatre angles de tilt différents (soit 816 simulations au total sur les 1123).

Comme ces 204 motifs ont été imagés à quatre angles de tilt différents, il y a donc $C_2^4 = 6$ couples utilisables pour la reconstruction. Le tableau 5.2 présente l’écart-type de la mesure de hauteur faite sur ces 204 motifs à partir des 6 couples de tilts, pour tout CR (204 motifs), seulement pour les motifs à CR nul (96 motifs), à CR=10nm (72 motifs) et à CR=30nm (36 motifs). Enfin la dernière colonne donne la moyenne des écarts-types faite à partir des 3 couples utilisant le tilt 12°.

Parallaxe	12°	9°	7°	5°	3°	2°	
Couple	0°-12°	3°-12°	5°-12°	0°-5°	0°-3°	3°-5°	moy T12°
RMSE _h (tout CR)	5.55	6.39	7.62	10.04	14.84	20.19	5.90
RMSE _h (CR = 0 nm)	4.02	4.92	5.75	6.87	9.24	14.90	4.51
RMSE _h (CR = 10 nm)	5.41	6.20	8.10	10.70	15.37	19.95	5.90
RMSE _h (CR = 30 nm)	8.55	9.49	10.46	14.73	23.42	30.37	8.59

TABLE 5.2 – Erreur sur la hauteur en fonction du couple de tilt utilisé pour différentes valeurs de CR (les valeurs sans unité sont en nm).

Tout d’abord, quel que soit le type de CR considéré, on peut observer l’augmentation de l’erreur de mesure lorsque le parallaxe diminue, conformément à ce que l’on avait repéré dans la littérature et conformément à la tendance de l’équation de la variance théorique de la mesure de hauteur (prouvée dans la partie 7) :

$$\sigma_h^2 = \frac{\sigma_{EW1}^2 + \sigma_{EW2}^2}{(\tan(T2) - \tan(T1))^2} \quad (5.21)$$

Cependant, si l’on calcule σ_h^2 avec cette équation, on va significativement surestimer les erreurs trouvées dans le tableau 5.2, car la présence de biais même faible entraîne $MSE_h > \sigma_h^2$, comme montré plus loin dans les équations 7.2.

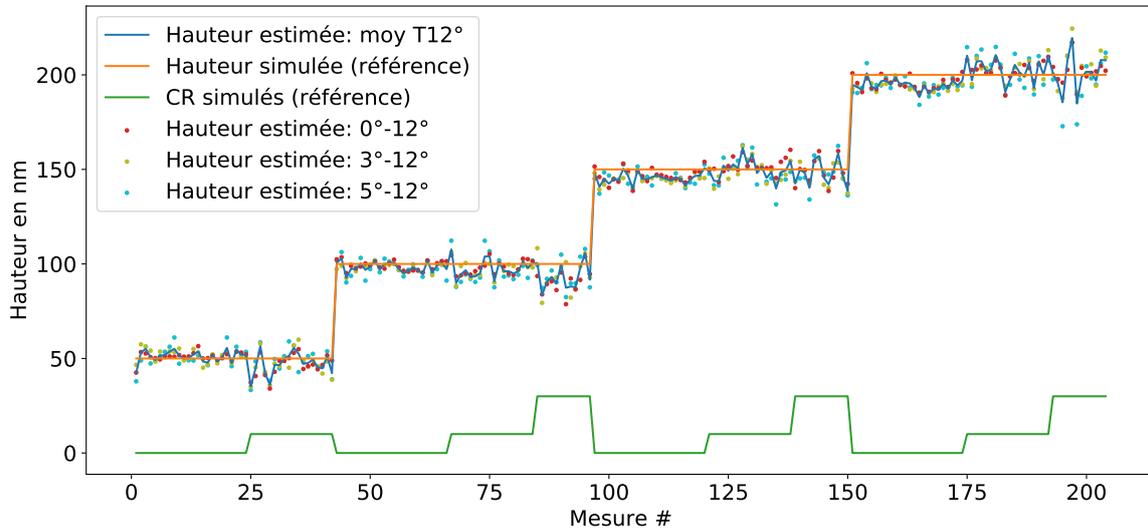


FIGURE 5.6 – Résultat du calcul de la hauteur à partir des EW estimés par le modèle. Les résultats provenant de trois couples d’angle de tilt sont présentés par des points de couleur et leur moyenne par le trait bleu. RMSE = 5.90nm.

Ensuite, on peut voir que l’augmentation de l’erreur sur l’EW avec l’augmentation de CR, s’est propagée à la mesure de hauteur, ce qui est également visible dans la Fig 5.6. Dans cette figure on a tracé en orange la valeur de hauteur simulée et en bleu la moyenne de la hauteur estimée à partir des trois couples d’angles de tilt utilisant le tilt à 12° (cad 0°-12°, 3°-12° et 5°-12°). Le résultat unitaire de chacun de ces couples est représenté par un point de couleur.

Parallaxe	12°	9°	7°	5°	3°	2°	
Couple	0°-12°	3°-12°	5°-12°	0°-5°	0°-3°	3°-5°	moy T12°
RMSE _{SWA} (tout CR)	0.69	0.87	1.13	0.89	1.45	4.49	0.76
RMSE _{SWA} (CR = 0 nm)	0.40	0.58	0.76	0.54	0.66	1.58	0.46
RMSE _{SWA} (CR = 10 nm)	0.73	0.98	1.35	1.09	1.39	7.20	0.84
RMSE _{SWA} (CR = 30 nm)	1.08	1.21	1.41	1.14	2.63	1.95	1.15

TABLE 5.3 – Erreur sur le SWA en fonction du couple de tilt utilisé pour différentes valeur de CR (les valeurs sans unités sont en degrés).

On peut ensuite calculer le SWA grâce à sa formule (Eq dans la Fig 2.4) et tracer les mêmes courbes (Fig 5.7), avec leurs erreurs associées (Tableau 5.3). On retrouve logiquement les mêmes tendances d’augmentation de l’erreur avec la diminution du parallaxe et l’augmentation du CR, avec un peu plus de bruit pour les parallaxes faibles. Pour ne pas alourdir la figure, la valeur du CR n’a pas été tracée sur ce graphique, mais le rangement des données est le même que précédemment, à savoir que sur chaque palier le CR augmente.

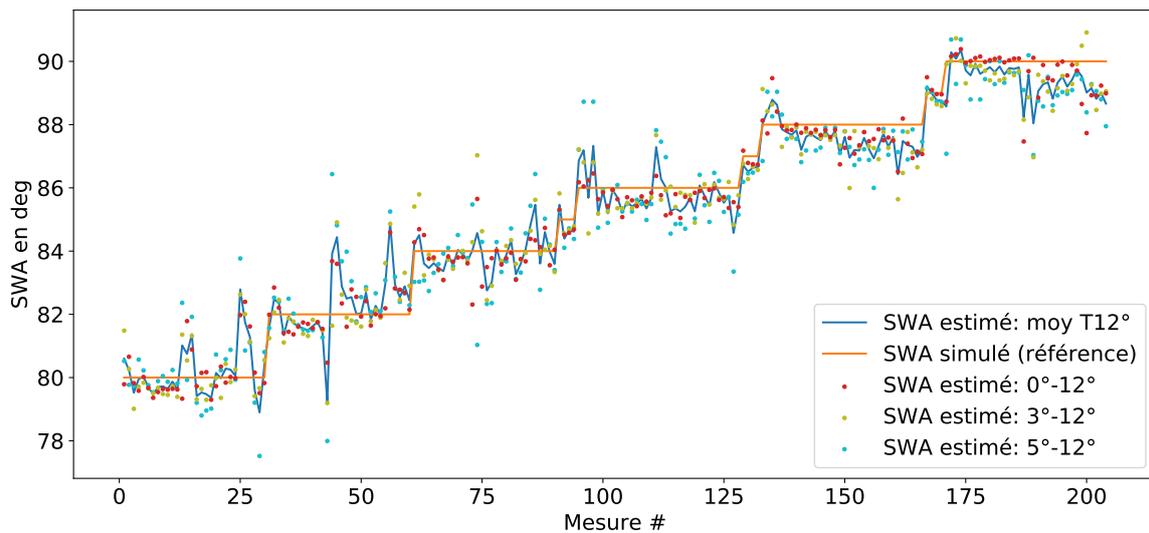


FIGURE 5.7 – Résultat du calcul du SWA à partir des EW estimés par le modèle. Les résultats provenant de trois couples d’angle de tilt sont présentés par des points de couleur et leur moyenne par le trait bleu. RMSE = 0.76°.

Optimisation du CDt

Finalement, la dernière grandeur qui va être optimisée est le CDt. On a choisi de l'estimer en dernier car, d'une part il n'apportait pas d'information nécessaire pour l'estimation des autres grandeurs tant qu'on se trouve dans le régime "grand CD" décrit en partie 4.5 et d'autre part pour qu'il bénéficie au maximum de toutes les connaissances accumulées du motif pour son estimation et notamment les valeurs du CRt et du SWA qui, on l'a vu (Fig 4.2 et Fig 4.9), sont importantes pour bien estimer le CD.

Le calcul utilisé par le modèle pour estimer le CDt est le suivant :

$$CDt_{JMO} = 0.030 \times EE + 0.450 \times CC_{CReff} + 0.523 \times DD_{CReff} \quad (5.22)$$

avec,

$$CC_{CReff} = CC + 1.2 + 1.2(90 - SWA) - 0.28(88.7 - SWA) \times (CRt_{JMO} + 3) \quad (5.23)$$

$$DD_{CReff} = DD + 1.3 + 3.5 \exp(-5 \tan(90 - SWA)) - 0.25(88.7 - SWA) \times (CRt_{JMO} + 3) \quad (5.24)$$

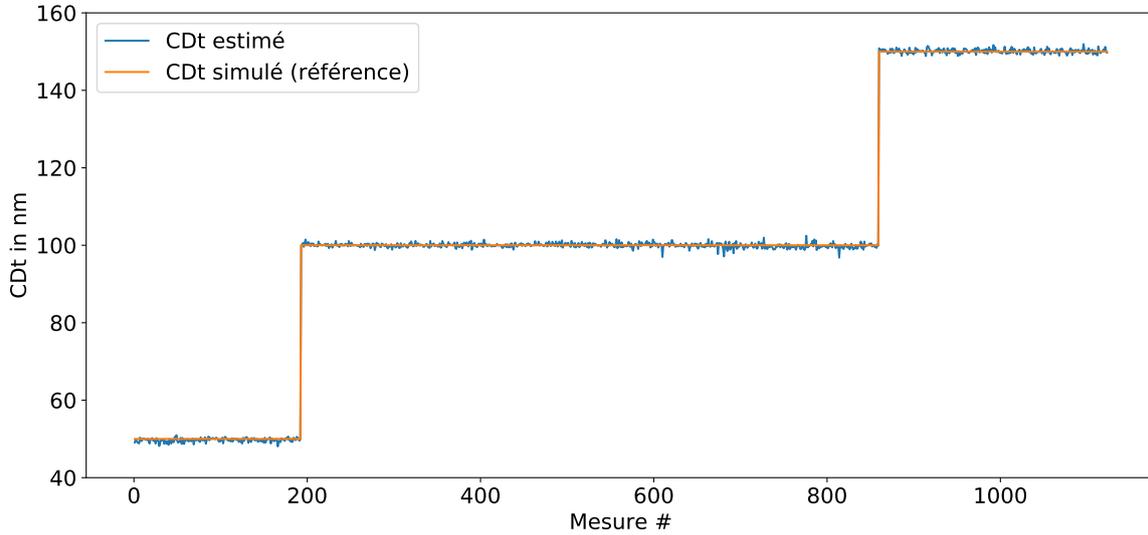


FIGURE 5.8 – Résultat de l'optimisation du CDt. RMSE = 0.58nm.

On rappelle que les CDt estimés de la Fig 5.8 proviennent de motifs de topographies diverses (hauteur, SWA, CRb, CRt) imagés en vue du dessus ou avec un angle de tilt. Le modèle permet donc d'estimer de manière très correcte le CDt dans un large domaine de validité, du moins en simulation. Le tableau 5.4 montre l'erreur attendue en fonction de la valeur du CRt. Ici on peut voir que l'utilisation du CR de référence (dernière ligne) à la place du CR estimé améliore de manière non négligeable l'estimation du CDt. En effet on a vu que l'estimation du CRt présentait une erreur importante, qui s'est propagée sur l'estimation du CDt. Cependant comme on va le voir, même mal estimée, la valeur du CRt améliore grandement l'estimation du CDt (Fig 5.11).

5.1.4 Apport du modèle par rapport aux méthodes classiques

Comme on l'a vu dans l'état de l'art, la plupart des méthodes actuelles de détermination de grandeurs à partir d'images SEM se basent sur la mesure de distance entre deux points correspondant à une valeur de dérivée voire d'une limite de niveau de gris, que ce soit sur plusieurs images pour déterminer la topographie ou sur des images uniques pour déterminer les

CRt simulé (nm)	0	2	5	10	15	20	30	tot
RMSE _{CDt} (CR estimés)	0.48	0.54	0.31	0.52	0.65	0.62	0.86	0.58
RMSE _{CDt} (CR ref)	0.42	0.36	0.31	0.44	0.47	0.57	0.63	0.47

TABLE 5.4 – Augmentation de l’erreur d’estimation sur le CDt avec le CRt (valeurs en nm).

CD notamment. Il est donc légitime de se demander si le modèle présenté améliore les résultats de manière significative par rapport aux méthodes classiques.

Tout d’abord la méthode présentée permet de mesurer les CR du motif, ce qui à notre connaissance n’a jamais été proposé avec autant de précision. Ensuite grâce à la connaissance de ces CR, la mesure d’EW a été fortement améliorée comme on peut le voir sur la Fig 5.9.

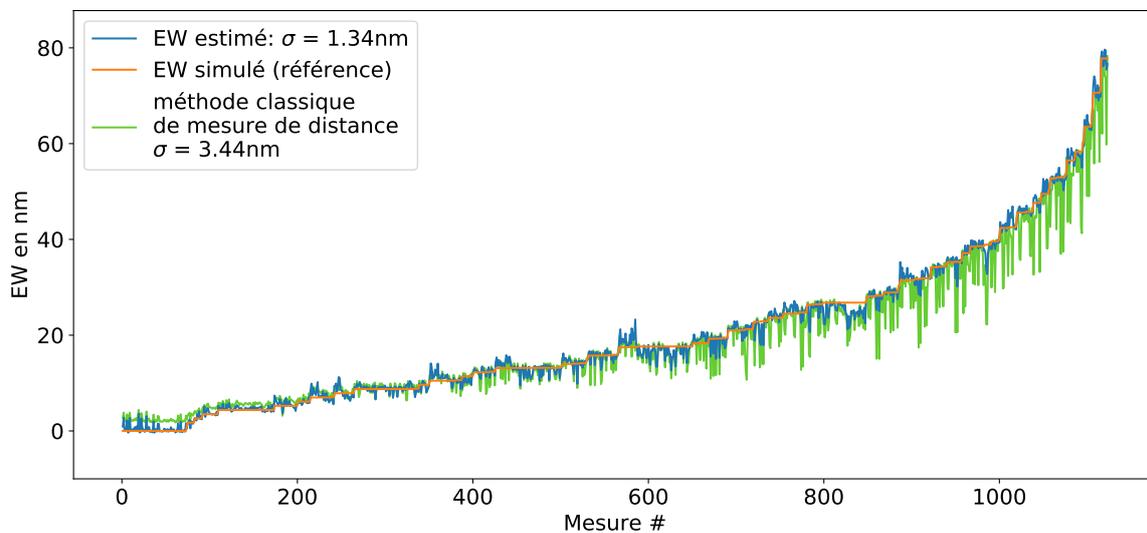


FIGURE 5.9 – Comparaison de la mesure d’EW via le modèle et via une méthode classique.

On peut voir que le modèle corrige la surestimation faite pour des EW faibles. Ensuite, plus les CR sont grands, plus l’estimation est améliorée grâce à leur prise en compte. Il faut bien comprendre que des motifs dont les CR sont nuls n’existent pas dans la réalité. Il faut déjà faire un effort d’élaboration des motifs conséquent pour obtenir des $CR < 5$ nm (couche d’arrêt, gravure sélective, polissage mécano-chimique...). C’est pourquoi l’amélioration de la mesure pour des CR non nuls et inconnus est une amélioration non négligeable pour la mesure d’EW. On rappelle que les angles d’inclinaison du faisceau électronique utilisés pour la reconstruction sont faibles, c’est pourquoi même une amélioration faible de la mesure d’EW peut entraîner une amélioration conséquente de la mesure de hauteur.

En utilisant les mesures de la méthode classique (en vert clair dans la Fig 5.9) pour reconstruire la hauteur du motif on obtient des résultats relativement similaires pour les CR nuls, mais grandement améliorés pour les CR non nuls (Fig 5.10). Pour des hauteurs inférieures à 100 nm, on obtient d’ailleurs des erreurs sur la mesure de hauteur parfois supérieures à 50% comme annoncé dans la littérature. On peut néanmoins être déçu que le niveau de bruit (sans considérer les erreurs de biais) ne soit pas réduit par le modèle. Une méthode de réduction du bruit basée sur l’utilisation de plusieurs splines de niveau de fit différents sera discutée en annexe ”Amélioration”.

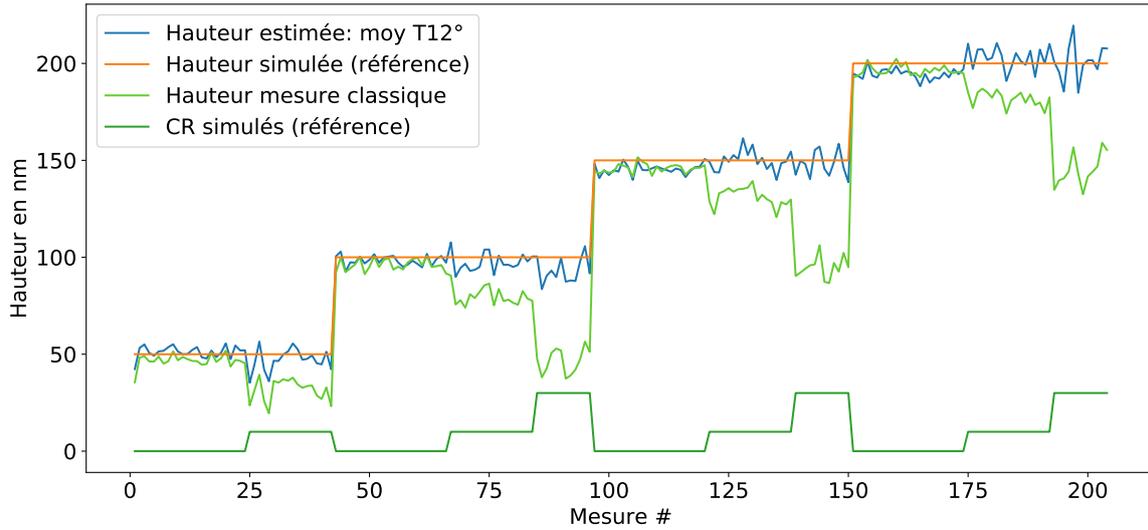


FIGURE 5.10 – Comparaison de la mesure de hauteur via le modèle et via une méthode classique.

Enfin, on peut voir dans la Fig 5.11, l'amélioration de la mesure de CDt. Ici aussi, plus le CRt est important, plus l'amélioration par rapport à une méthode classique est importante. La prise en compte du SWA estimé permet également une amélioration de la mesure du CDt. On peut voir que le RMSE passe de 2.26 nm avec une méthode standard, à 0.58 nm grâce au modèle.

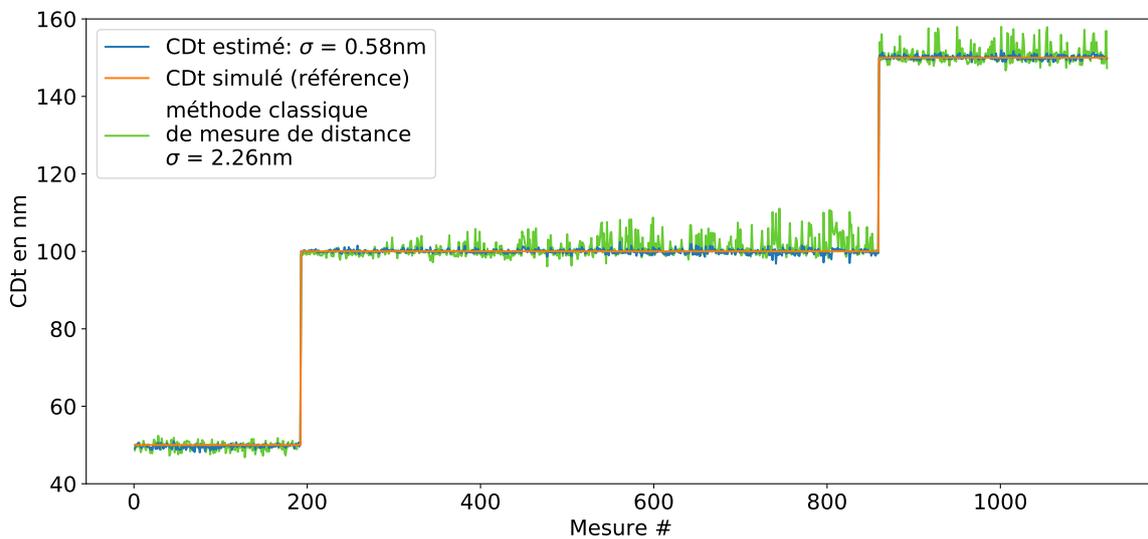


FIGURE 5.11 – Comparaison de la mesure du CDt via le modèle et via une méthode classique.

5.1.5 Conclusion

A ce stade, on dispose donc d'un modèle linéaire de descripteurs secondaires permettant de remonter aux données topographiques des motifs simulés. Ce modèle a pu être construit grâce à l'analyse des images simulées par JMONSEL. Si l'on a décidé de passer par une phase de simulation, c'est parce que l'on souhaitait repérer des tendances de variation de l'image SEM en fonction de la topographie des motifs. Les valeurs exactes des grandeurs d'intérêt estimées par le modèle à ce stade doivent maintenant être ajustées grâce à des mesures SEM réelles. En effet un SEM est une machine très complexe et un logiciel ne peut pas simuler toutes les particularités de chaque SEM (et encore faudrait-il toutes les connaître).

5.2 Calibration expérimentale du modèle INVSEM3D

Une phase de calibration sur des motifs réels est indispensable pour pouvoir utiliser le modèle de manière quantitative. Cette calibration va être réalisée en comparant les résultats obtenus avec le modèle non calibré sur des mesures faites sur des motifs de type ligne en silicium gravés, avec les résultats obtenus à partir de mesures AFM sur ces mêmes motifs.

Grâce à la simulation, on a créé une toile n-dimensionnelle représentant la valeur de l'estimation des grandeurs d'intérêt, que l'on a ondulée en fonction des effets de chaque dimension sur cette grandeur d'intérêt (ces dimensions étant les descripteurs, qui eux même varient en fonction de la topographie du motif observé). Désormais on va tester empiriquement la valeur de certains points de l'espace n-dimensionnel de la toile (bien moins nombreux qu'en simulation) de manière à recaler leurs niveaux sans perdre la forme générale de l'ondulation de la toile. On va autoriser une transformation de translation de la toile (un biais) et d'étirement ou de compression de la toile (coefficient sur la grandeur d'intérêt) soit :

$$\begin{cases} [\text{CRb}_{\text{Calib}}] = a \times [\text{CRb}_{\text{JMO}}] + b \\ [\text{CRt}_{\text{Calib}}] = c \times [\text{CRt}_{\text{JMO}}] + d \\ [\text{EW}_{\text{Calib}}] = e \times [\text{EW}_{\text{JMO}}] + f \\ [\text{CDt}_{\text{Calib}}] = g \times [\text{CDt}_{\text{JMO}}] + h \end{cases} \quad (5.25)$$

$$\Rightarrow [[\text{Modele}_{\text{Calib}}]] = [A] \times [[\text{Modele}_{\text{JMO}}]] + [B] \quad (5.26)$$

5.2.1 Création de motifs en silicium gravé

Pour pouvoir réaliser cette calibration on a donc dû réaliser des plaques présentant différentes topographies. Pour cela quatre plaques ont été réalisées, à des hauteurs de gravure différentes mais présentant les mêmes motifs. Les hauteurs visées étaient 50 nm, 100 nm, 150 nm et 200 nm. Pour cela, nous avons utilisé une méthode de fabrication classique par lithographie puis de gravure. Cette méthode comprend les procédés suivants (Fig 5.12) :

- 1) Ajout d'un masque dur (hard-mask : SOC 130 nm et Si-ARC 30 nm) sur la plaque de silicium vierge (taille 300 mm).
- 2) Dépôt uniforme de résine sur le masque dur.
- 3) Ouverture de la résine par lithographie électronique, c'est à dire par bombardement électronique de la résine sur les zones à enlever.
- 4) Ouverture du masque dur, dans les zones où la résine a été enlevée, par gravure plasma physique (Break-through).
- 5) Enlèvement complet de la résine.
- 6) Gravure du motif dans le silicium par gravure plasma avec passivation des flancs par polymérisation, permettant de protéger les flancs des zones déjà gravées et ainsi mieux contrôler la pente des flancs.
- 7) Lavage HF pour enlever le masque dur restant et les polymères servant à la passivation.

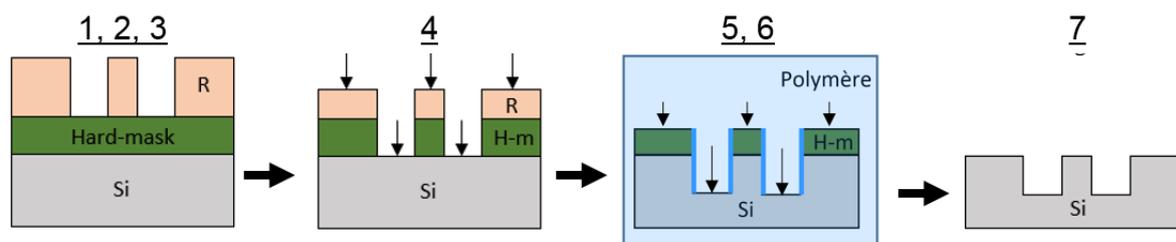


FIGURE 5.12 – Représentation des étapes de fabrication des plaques de silicium.

Sur ces plaques différents motifs ont été gravés :

- Des lignes isolées allant de 30 nm à 300 nm de CD.
- Des tranchées isolées allant de 30 nm à 300 nm de CD.
- Des réseaux denses allant de 30 nm à 300 nm de CD, avec des densités de 1 : 1 à 1 : 5.
- Des champs de contacts et de trous de 30 nm à 300 nm de diamètre, avec des densités de 1 : 1 à 1 : 5.
- Des motifs plus complexes divers (motifs non-Manhattan, motifs type "hotspot", etc...).

Pour cette partie de calibration, seules les lignes isolées seront utilisées car le modèle a été créé pour reconstruire des lignes isolées. On va utiliser les lignes dont le CD est de 50 nm, 100 nm, 160 nm et 200 nm, pour chacune des quatre plaques, on aura donc 16 types de motifs différents. Pour chacun de ces 16 motifs, 50 images SEM vont être réalisées pour chaque angle de tilt disponible (0° , 3° , 5° et 12°), ce qui fait donc 3200 images en tout.

Pour chacune de ces 3200 images, les descripteurs primaires seront mesurés par la méthode d'analyse définie en partie 3.2. A partir de ces descripteurs primaires, les descripteurs secondaires du modèle seront calculés, puis le modèle les utilisera pour retrouver les valeurs de grandeur d'intérêt. Ces grandeurs d'intérêt non calibrées seront ensuite comparées à des mesure de référence AFM de ces mêmes motifs de manière à calibrer expérimentalement le modèle.

5.2.2 Mesures de référence par AFM

La calibration expérimentale s'appuie sur des mesures topographiques de référence réalisées sur les 16 motifs fabriqués, de manière à pouvoir comparer l'estimation des différentes grandeurs d'intérêts à une valeur de référence. Pour cela des mesures AFM-3D ont été réalisées sur les mêmes motifs que ceux observés en SEM. Cependant il est extrêmement difficile de faire les mesures SEM et AFM exactement au même emplacement.

Il faut bien comprendre que dans cette partie on va travailler sur des valeurs moyennes ce qui va avoir des conséquences plus ou moins importantes en fonction des grandeurs d'intérêt observées. Par exemple la hauteur "vraie" pour des motifs de type ligne (non accessible par la mesure) sur une même plaque sur des motifs voisins (moins de $100\ \mu\text{m}$) ne devrait pas beaucoup varier. L'utilisation de sa valeur mesurée moyenne ne sera donc pas une approximation trop grossière. En revanche le CDt "réel" peut fortement varier même sur quelques nanomètres, c'est pourquoi l'utilisation de sa valeur mesurée moyenne pourra se révéler insuffisante, comme on va le voir.

Les profils AFM ont été analysés par la méthode présentée en partie 3.1, permettant un approximation de type "8 paramètres". De cette approximation les différentes grandeurs d'intérêt de référence ont été extraites.

Les hauteurs et les SWA de référence trouvés par cette méthode ont été moyennés par plaque. En effet, leurs écarts-types entre motifs d'une même plaque étant de l'ordre de grandeur de l'erreur de mesure de l'AFM, on peut conclure que le CD des motifs n'a pas influencé leur hauteur et leur SWA.

Les grandeurs trouvées par l'AFM en fonction des plaques sont les suivantes (incertitude 3σ) :

- P1 : $h = 53.3 \pm 0.3\ \text{nm}$, $\text{SWA} = 87.5 \pm 6^\circ$, $\text{CRb} = 23.6 \pm 4.5\ \text{nm}$, $\text{CRt} = 3.2 \pm 2.5\ \text{nm}$
- P2 : $h = 108.0 \pm 0.6\ \text{nm}$, $\text{SWA} = 87.8 \pm 2^\circ$, $\text{CRb} = 26.9 \pm 4.5\ \text{nm}$, $\text{CRt} = 3.2 \pm 2.5\ \text{nm}$
- P3 : $h = 167.7 \pm 1.0\ \text{nm}$, $\text{SWA} = 89.4 \pm 1^\circ$, $\text{CRb} = 29.3 \pm 4.5\ \text{nm}$, $\text{CRt} = 3.2 \pm 2.5\ \text{nm}$
- P4 : $h = 225.8 \pm 1.0\ \text{nm}$, $\text{SWA} = 89.5 \pm 1^\circ$, $\text{CRb} = 33.9 \pm 4.5\ \text{nm}$, $\text{CRt} = 3.2 \pm 2.5\ \text{nm}$

On peut voir que l'incertitude sur le SWA est bien plus grande sur les motifs de faible hauteur que pour les grands motifs. En effet, une même incertitude n'a pas du tout la même influence sur un petit motif que sur un motif haut, car cette erreur d'angle se propage sur une distance plus

longue lorsque le motif est haut, c'est pourquoi qualitativement une erreur de 6° sur des motifs de 53.3 nm n'est pas plus visible qu'une erreur de 1° sur des motifs de 225.8 nm de hauteur.

5.2.3 Calibration

Pour pouvoir trouver les coefficients de calibration de l'équation 5.25, on va donc comparer le résultat du modèle JMONSEL aux valeurs de référence trouvées par AFM.

Les graphiques qui vont être présentés dans cette partie prendront la forme de la Fig 5.13 (cette figure étant seulement un exemple simple sans intérêt d'analyse, mais permettant de comprendre la visualisation proposée pour la calibration). En bleu est représentée l'estimation de la grandeur faite à partir du modèle JMONSEL trouvé dans la partie 5. En vert est représentée la valeur moyenne de la grandeur trouvée à l'AFM sur le type de motif concerné par paquets de 50 mesures redondantes. Enfin en orange est représentée l'estimation de la grandeur faite à partir du modèle calibré sur la référence.

Si le modèle JMONSEL est correct (et que la mesure de référence AFM est correcte!), on fait l'hypothèse que l'estimation des grandeurs par le modèle JMONSEL devraient suivre les données AFM à une transformation linéaire près (d'où les équations 5.25).

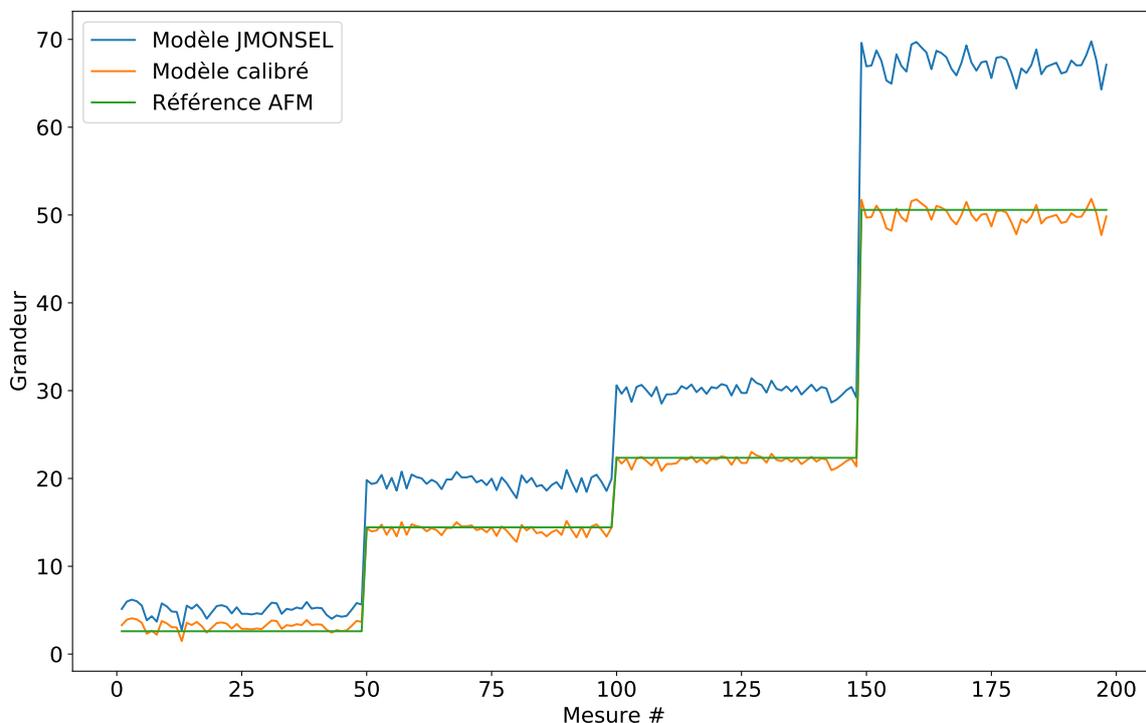


FIGURE 5.13 – Exemple de calibration de l'estimation d'une grandeur quelconque sur les données AFM. $\text{Grandeur}_{\text{calib}} = 0.75 \times \text{Grandeur}_{\text{JMO}} - 0.58$.

Calibration du CRb

La calibration est faite dans le même ordre que la création du modèle, la première grandeur calibrée est donc le CRb (Fig 5.14).

Dans cette figure les données ont été rangées par plaque (donc par hauteur de gravure). Les 200 premiers motifs correspondent donc aux quatre CD de la plaque P1 redondants 50 fois chacun. Les 200 suivants de la plaque P2, puis P3 et P4. On a représenté ici seulement l'estimation du CRb à tilt nul, c'est à dire en vue du dessus, car contrairement à ce que l'on avait observé

via les simulations JMONSEL, le tilt fait varier la valeur de l'estimation du CRb et augmente également fortement l'erreur. C'est pourquoi on va seulement estimer le CRb sur les images en vue du dessus.

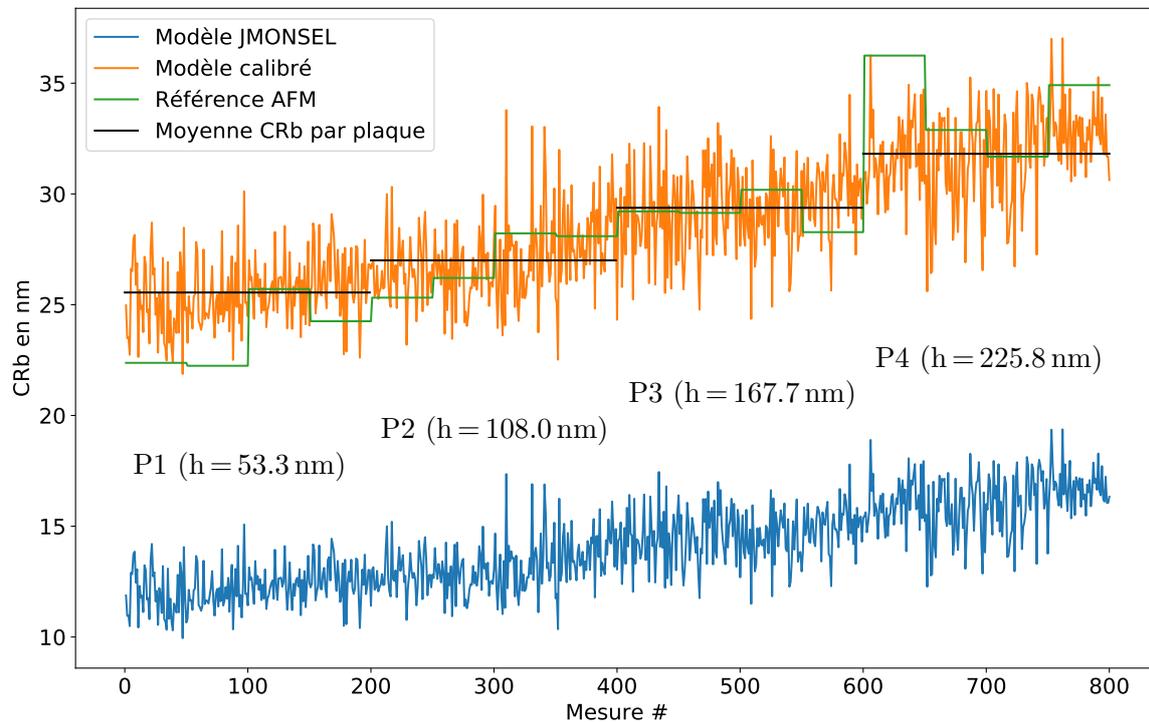


FIGURE 5.14 – Calibration de l'estimation du CRb sur les données AFM.
 $CRb_{calib} = 1.61 \times CRb_{JMO} + 5.88$. $RMSE_{calib} = 2.58$ nm.

Tout d'abord, si l'on regarde la courbe verte qui représente la mesure de référence par AFM du CRb des motifs, on peut voir qu'en moyenne le CRb augmente avec la hauteur de gravure. C'est une très bonne nouvelle car si le CRb avait été constant, il aurait été très difficile de savoir si le modèle était prédictif.

L'estimation du CRb par le modèle JMONSEL (bleu) montre une sous-estimation importante, mais sa variation est cohérente avec les mesures AFM. Après calibration (orange) on obtient donc un modèle qui, malgré un bruit non négligeable, suit en moyenne la tendance d'augmentation du CR avec la hauteur du motif. Pour bien s'en rendre compte on a tracé en noir la moyenne des valeurs du CRb estimé par le modèle calibré sur chacune des quatre plaques. Si on fait l'hypothèse que le CR "réel" des motifs est le même quelque soit le CD du motif, on peut voir que la mesure AFM est également bruitée.

Ce résultat est important, car il prouve la possibilité de mesure du CRb par imagerie SEM (et ici c'est une image simple en vue du dessus sans besoin d'un SEM à faisceau électronique inclinable). De plus la sensibilité de la mesure est remarquable, en effet une différence de 5 nm sur le CR n'est pas si facilement repérable même sur une mesure classique comme l'AFM. Pour se rendre compte de ce que représente une différence de 5 nm sur un CR, la Fig 5.15a compare des motifs dont les CR correspondent à la plaque P1 et à la plaque P4. La mesure SEM présentée, est donc suffisamment sensible pour différencier la topographie de la courbe bleu à celle de la courbe rouge en moyenne.

Enfin la Fig 5.15b permet de se rendre compte de l'impact du bruit sur le résultat topographique final, en l'occurrence sur un motif de type plaque P1. Si on fait l'hypothèse que le modèle ne présente pas de biais systématique (ce qui est une hypothèse forte et pas forcément justifiée vu

la Fig 5.14, mais disons un biais suffisamment faible), alors si la courbe noire correspond au CRb "réel", 68.3% des mesures du CRb par analyse de l'image SEM et utilisation du modèle calibré donneront un résultat compris entre les courbes en pointillés vert, 95.5% entre les courbes en pointillés orange et 99.7% entre les courbes en pointillés rouge.

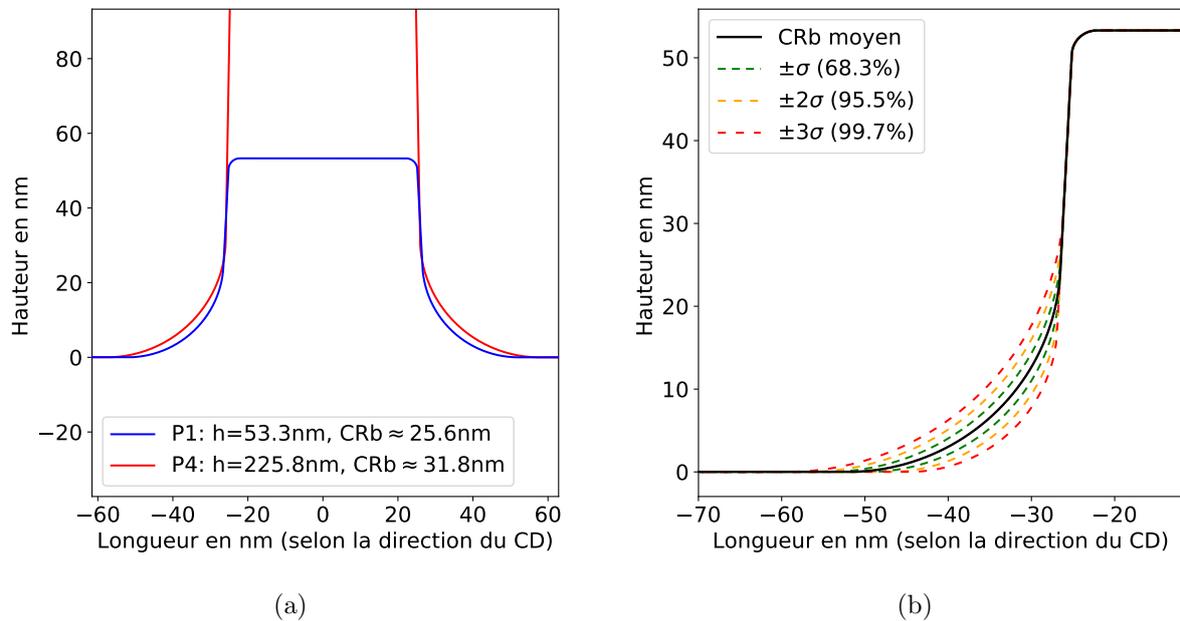


FIGURE 5.15 – (a) : Comparaison du CRb des motifs de la plaque P1 au CRb des motifs de la plaque P4. (b) : Visualisation du bruit de mesure de l'estimation du CRb par le modèle calibré.

Calibration du CRt

Après l'estimation du CRb où toutes les conditions étaient réunies pour créer un modèle prédictif, le cas du CRt va présenter toutes les caractéristiques empêchant une estimation réussie (Fig 5.16).

Premier problème, les mesures de référence AFM montrent que le CRt ne varie pas quelque soit le motif ou la plaque. Deuxième problème, la valeur du CRt est très faible (environ 3.4 nm), or on a vu dans la partie simulation JMONSEL qu'il était très difficile d'estimer les CRt faibles. Troisième problème, le modèle JMONSEL ne varie pas à une transformation linéaire près des données AFM. Il semblerait que l'estimation du CRt soit influencée par le CDt; or cet effet n'était pas du tout présent lors de l'analyse des simulations JMONSEL. Il s'agit donc d'une erreur de simulation du logiciel utilisé.

La combinaison de ces trois problèmes entraîne un effet de moyenne de la calibration, c'est à dire que le coefficient donné au modèle JMONSEL va être très faible, écrasant son impact (et donc son intérêt). Le biais de 3,4 nm va correspondre à la moyenne de la valeur de référence. Concrètement, le modèle ne prédit pas la valeur du CRt et la remplace par une constante, l'écart-type n'a donc pas vraiment de sens hormis d'être environ celui de la mesure de référence.

On pourrait re-travailler pour supprimer l'effet du CDt sur la valeur du CRt directement dans la partie calibration, mais malgré cela le problème de non variation du CRt dans le lot de données de calibration empêchera toute conclusion. Pour pouvoir créer un modèle prédictif du CRt, il faudrait donc également refaire de nouvelles plaques avec des CRt différents et plus grands.

Après avoir comparé les images SEM aux simulations JMONSEL les plus proches des

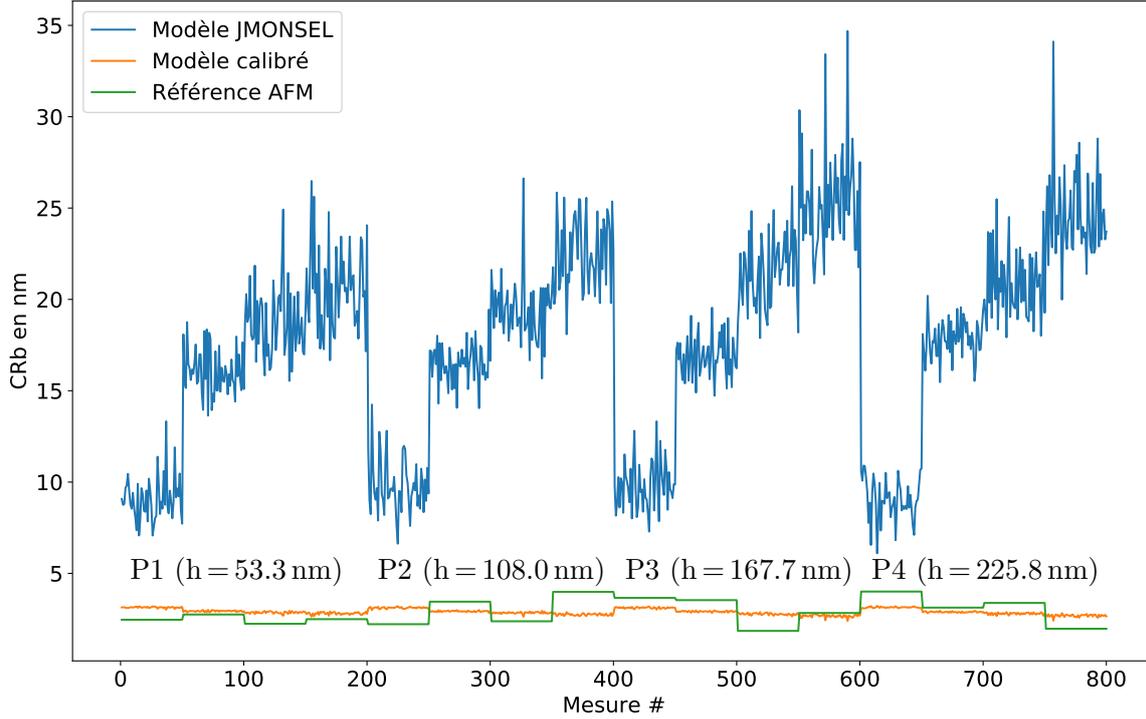


FIGURE 5.16 – Calibration de l’estimation du CRt sur les données AFM.

$$\text{CRt}_{\text{calib}} = -0.03 \times \text{CRt}_{\text{JMO}} + 3.4 \approx 3.4. \quad \text{RMSE}_{\text{calib}} = 0.67 \text{ nm.}$$

motifs de nos plaques, nous avons remarqué que le problème venait du rendement électronique au centre du motif qui, lorsque le CDt diminue, est différent entre la simulation JMONSEL et les images SEM. En effet le rendement électronique du milieu du motif augmente plus vite avec la réduction du CDt sur les images SEM, que via la simulation, ce qui entraîne un changement de position des descripteurs utiles pour l’estimation du CRt.

Calibration de l’EW et utilisation pour le calcul de la hauteur

La grandeur d’intérêt suivante évaluée est l’EW. Pour son évaluation, les descripteurs secondaires utilisant la valeur du CRb et/ou du CRt utiliseront la valeur calibrée, trouvée précédemment (c’est à dire des courbes orange).

On peut voir sur la Fig 5.17 le résultat de la calibration. Les données sont rangées par plaque (4 paquets de 800), puis par CD (16 paquets de 200) et enfin par tilt (64 paquets de 50), ce qui donne ce rangement en escalier (voir encart).

On rappelle que la courbe verte correspond aux valeurs de référence trouvées par AFM. Cependant la mesure AFM ne permet pas de trouver directement la valeur d’EW pour différents tilts ; c’est pourquoi l’EW de référence est en fait calculé à partir de la hauteur et du SWA trouvé par AFM, pour un certain tilt, grâce à la formule :

$$\text{EW}_{\text{ref(AFM)}}(T) = \frac{h_{\text{ref(AFM)}}}{\tan(\text{SWA}_{\text{ref(AFM)}})} + h_{\text{ref(AFM)}} \times \tan(T) \quad (5.27)$$

Le modèle JMONSEL semble relativement satisfaisant, la calibration fonctionne donc correctement. Cependant, si l’on regarde le résultat de la mesure de référence, on peut voir des variations non négligeables pour des mêmes plaques à CD différents, alors que les résultats SEM sont très consistants entre eux. Ce problème vient du fait que l’EW théorique est calculé

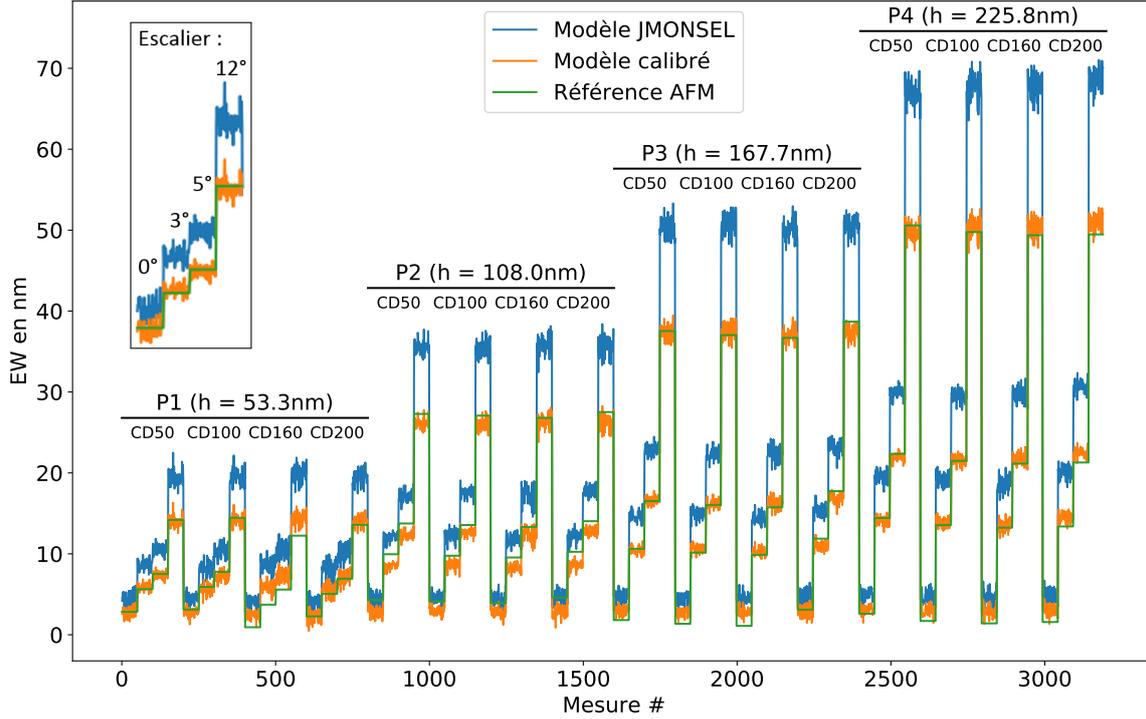


FIGURE 5.17 – Calibration de l'estimation du EW sur les données AFM.

$$EW_{\text{calib}} = 0.75 \times EW_{\text{JMO}} - 0.58. \quad \text{RMSE}_{\text{calib}} = 1.20 \text{ nm.}$$

à partir de la hauteur, mais également du SWA des motifs ; or la hauteur mesurée à l'AFM est généralement très robuste. En revanche la mesure de SWA peut parfois être un peu instable (problème de pointe, topographie locale non-représentative, erreur de l'approximation 8 paramètres...) menant à une instabilité sur la mesure d'EW.

Récapitulons, dans la partie sur les simulations JMONSEL on a minimisé :

$$\|EW_{\text{simu}} - EW_{\text{JMO}}\|^2 \quad (5.28)$$

où EW_{simu} était la valeur exacte d'entrée de la simulation et EW_{JMO} l'estimation de l'EW grâce aux descripteurs secondaire. Dans cette partie on cherche à minimiser :

$$\|EW_{\text{ref(AFM)}} - EW_{\text{calib}}\|^2, \text{ avec } EW_{\text{calib}} = a + b \times EW_{\text{JMO}} \quad (5.29)$$

où $EW_{\text{ref(AFM)}}$ est la valeur d'EW théorique calculé à partir de la hauteur et du SWA trouvé par mesure AFM. Enfin on peut résoudre :

$$h_{\text{calcul}} = \frac{EW_{2\text{calib}} - EW_{1\text{calib}}}{\tan(T_2) - \tan(T_1)} \quad (5.30)$$

avec éventuellement si besoin,

$$\|h_{\text{ref(AFM)}} - h_{\text{calib}}\|^2, \text{ avec } h_{\text{calib}} = a + b \times h_{\text{calcul}} \quad (5.31)$$

Si la calibration des EW est bonne alors cette dernière minimisation devrait être superflue et on devrait donc obtenir un coefficient "b" proche de 1 et un coefficient "a" proche de zéro. Cependant cette étape peut être cruciale si les angles de tilt du SEM utilisé ne sont pas bien calibrés, le coefficient "b" s'éloignerait alors de 1, compensant ainsi le mauvais coefficient

$1/(\tan(T_2) - \tan(T_1))$, ce qui permettrait au modèle de fonctionner malgré cela, mais ce qui le rendrait encore plus spécifique à l'utilisation sur ce SEM en particulier.

Le tableau 5.5 montre les coefficients trouvés avec cette étape de calibration. Comme prévu, la minimisation 5.31 améliore très peu le résultat du calcul brut pour les parallaxes "élevées". Le reste du tableau permet de voir les incertitudes sur la mesure de la hauteur. On peut voir, comme on l'avait déjà vu dans la partie simulation et dans la littérature, que l'incertitude augmente lorsque la parallaxe diminue. On peut également voir que l'erreur augmente avec la hauteur du motif mais de manière bien inférieure à la proportion entre les hauteurs ou même entre leurs EW.

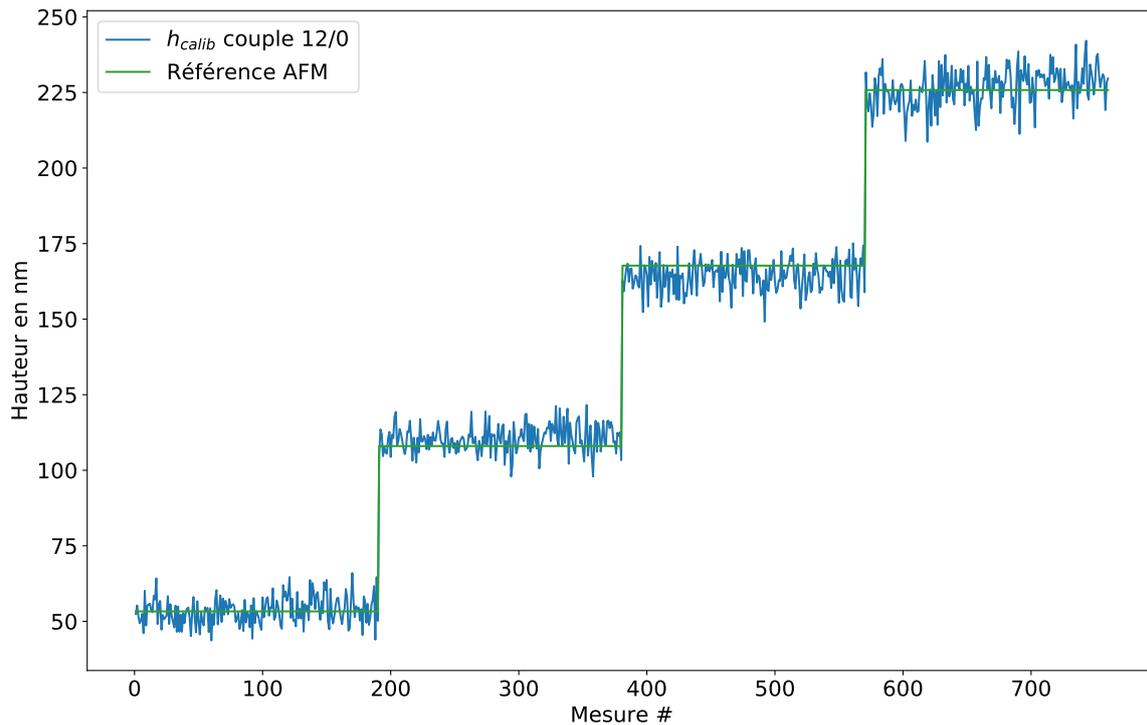


FIGURE 5.18 – Calibration de l'estimation de la hauteur sur les données AFM, en utilisant le couple de tilt $0^\circ/12^\circ$. $RMSE_{calib} = 5.58$ nm.

Parallaxe	12°	9°	7°	5°	3°	2°	
Couple	$12^\circ-0^\circ$	$12^\circ-3^\circ$	$12^\circ-5^\circ$	$5^\circ-0^\circ$	$3^\circ-0^\circ$	$5^\circ-3^\circ$	moy T12°
coeff b	1.03	0.98	0.98	1.06	1.09	0.85	1.00
coeff a	-2.30	1.68	1.48	-2.35	-2.83	20.74	-0.08
$RMSE_h$ (Tout h)	5.58	6.77	8.84	11.31	18.98	22.38	6.32
$RMSE_h$ (h = 53 nm)	4.55	5.84	8.07	10.68	20.08	21.71	5.25
$RMSE_h$ (h = 108 nm)	5.09	6.38	8.01	12.32	16.63	23.38	5.66
$RMSE_h$ (h = 168 nm)	6.19	6.38	8.92	12.32	20.24	19.55	6.51
$RMSE_h$ (h = 226 nm)	6.32	8.24	10.24	10.85	18.76	24.71	7.65

TABLE 5.5 – Incertitude sur la mesure de la hauteur.

La Fig 5.18 représente la hauteur calibrée par rapport à la référence AFM sur le couple d'angle de tilt $12^\circ-0^\circ$. Il s'agit du couple qui permet d'obtenir la meilleure précision sur la mesure

de hauteur et qui correspond à la première colonne du tableau 5.5. Ici on n'a pas représenté la courbe avant optimisation comme nous en avons l'habitude car elles sont si proches que cela nuirait à la visualisation.

On pourrait faire des optimisations de l'EW avec plus de degrés de liberté comme par exemple en minimisant les EW seulement par rapport à la hauteur :

$$\|h_{\text{ref(AFM)}} - (a + b \times (\text{EW2} - \text{EW1}) + c \times (\text{EW3} - \text{EW2}) + d \times (\text{EW3} - \text{EW1}) + \dots)\|^2 \quad (5.32)$$

voire,

$$\|h_{\text{ref(AFM)}} - (a + b \times \text{EW1} + c \times \text{EW2} + d \times \text{EW3} + \dots)\|^2 \quad (5.33)$$

Ce genre d'optimisations (Eq 5.32 ou Eq 5.33 ou autre) vont mener à une réduction forte du MSE; cependant, plus la calibration aura de degrés de liberté, plus on va diminuer le domaine de validité du modèle jusqu'à un domaine inacceptable qui sera strictement celui des données d'entrée de la calibration. En effet ce type de calibration ne permet plus de garder la toile n-dimensionnelle inchangée à une homothétie-translation près, comme prévue dans l'introduction de cette partie.

Calcul du SWA

On va ensuite utiliser la même approche pour le calcul du SWA, dont on rappelle le calcul :

$$\text{SWA}_{\text{calcul}} = \arctan\left(\frac{\text{EW2}_{\text{calib}} - \text{EW1}_{\text{calib}}}{\text{EW1}_{\text{calib}} \times \tan(T_2) - \text{EW2}_{\text{calib}} \times \tan(T_1)}\right) \quad (5.34)$$

avec de la même manière une minimisation possible,

$$\|\text{SWA}_{\text{ref(AFM)}} - \text{SWA}_{\text{calib}}\|^2, \text{ avec } \text{SWA}_{\text{calib}} = a + b \times \text{SWA}_{\text{calcul}} \quad (5.35)$$

Sur la Fig 5.19 est représenté le résultat du calcul direct du SWA en bleu (Eq 5.34), puis de son ajustement en orange (Eq 5.35) par rapport aux données AFM en vert. On remarque tout de suite que l'ajustement semble être mauvais; on va essayer d'expliquer pourquoi.

Tout d'abord on rappelle que ces résultats proviennent des mêmes images que précédemment; or l'estimation de la hauteur est faite sans biais (on a vu que la minimisation 5.31 était même superflue). En revanche, pour ces mêmes motifs, l'estimation du SWA semble difficile. Cela vient du fait que pour le calcul de la hauteur, le cœur de la formule est la différence ΔEW entre EW2 et EW1, c'est à dire de manière imagée si on regarde la Fig 5.17, l'écart relatif entre les marches des escaliers (cet escalier correspond au même motif vu à des angles de vues différentes). En revanche pour le calcul du SWA la hauteur moyenne de l'escalier est également importante.

Or si on regarde la courbe verte sur cette Fig 5.17, on peut voir que les écarts relatifs des escaliers sont très reproductibles alors que la hauteur moyenne de l'escalier peut varier de manière non négligeable (par exemple sur la P1 l'escalier 3 est plus bas que les autres mais conserve la même forme).

Pour les données AFM cela vient du fait que le SWA peut être difficile à estimer de manière très précise, notamment quand le motif a une hauteur faible. On rappelle que les motifs de nos plaques ont des CR d'environ 27 nm en bas et 3 nm en haut; sur les motifs de 53.3 nm cela fait donc $53.3 - 27 - 3 = 23.3$ nm de flanc. De ce fait une différence même de 4° de SWA est en fait très faible sur la topographie du motif. L'encart de la Fig 5.19 permet de voir la différence entre un SWA de 85° en bleu et de 89° en rouge sur un tel motif. On comprend donc plus facilement pourquoi le bruit est fort sur la mesure AFM comme sur la reconstruction SEM pour la plaque P1 dont la hauteur de motif est faible.

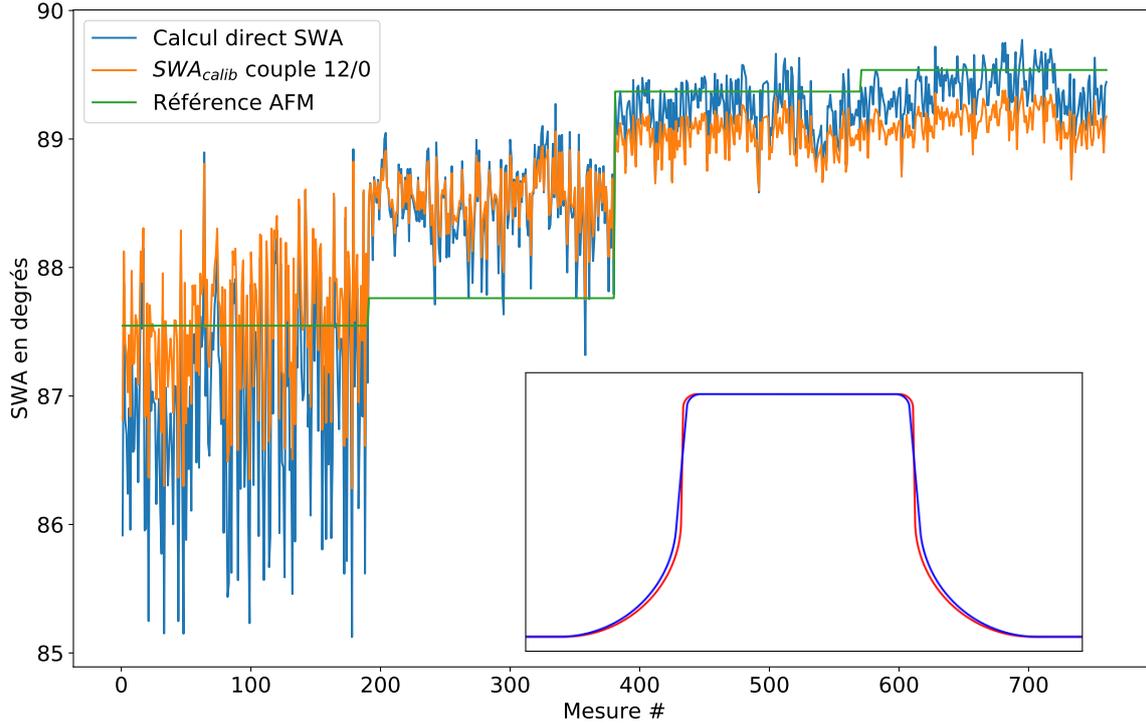


FIGURE 5.19 – Calibration de l’estimation du SWA sur les données AFM. $RMSE_{calib}=0.61^\circ$. L’encart représente la différence entre un SWA de 85° en bleu et de 89° en rouge sur un motif de 53.3 nm.

On peut ensuite voir dans la Fig 5.19 que pour la plaque P2 il y a un biais important entre la référence AFM et l’estimation du SWA. On pouvait déjà le voir sur la Fig 5.17 où les EW étaient sous-estimés par rapport à la référence pour la plaque P2. Et, comme on l’a vu, une différence de translation entre la référence et l’estimation engendre une erreur sur le SWA mais pas sur la hauteur si les écarts relatifs sont respectés (ce qui est le cas).

Si l’optimisation a gardé ce biais (à l’étape de l’EW) c’est que c’était le meilleur compromis au vu de toutes les mesures. Or comme la mesure SEM est extrêmement reproductible (à condition que l’échantillon ne soit pas chargé/endommagé par la mesure, ce qui est une hypothèse acceptable ici), mon sentiment est qu’à ce stade on peut avoir plus confiance dans l’écart relatif entre les SWA trouvés par la reconstruction SEM que par la mesure AFM. En revanche la valeur moyenne et le niveau d’étirement restent guidés par les mesures AFM lorsque les EW sont optimisés. De mon point de vue, la minimisation 5.35 est même inutile, ce qui veut dire qu’il vaut mieux utiliser la courbe bleue de la Fig 5.19 plutôt que la courbe orange (tout comme pour la hauteur, sauf que dans ce cas les courbes avant et après optimisation étaient pratiquement égales).

Il est très difficile de démontrer ce point de vue ; dans le cadre de cette thèse cela restera donc à l’état de croyance (ou de non-croyance ce qui est parfaitement acceptable).

Calibration du CDt

Le CDt est une grandeur assez particulière car on ne peut pas faire, comme pour les autres grandeurs, l'hypothèse qu'elle est invariante (ou même faiblement variante) sur une plaque. En effet le CD réel des lignes, même sur la longueur d'une image SEM, peut varier localement de plus de 10 nm (Fig 5.20a). De plus, d'une image à l'autre, le CD moyen sur l'image peut également varier de plus de 10 nm (Comparaison Fig 5.20a avec Fig 5.20b). De ce fait la référence AFM que l'on a utilisée jusque là ne semble plus adaptée car il faudrait mesurer chacun des motifs mesurés par SEM sur une longueur de 500 nm, ce qui représente un temps de mesure beaucoup trop important et des risques de non-reproductibilité non-négligeables. De plus il faudrait que la mesure SEM et AFM soient faites exactement au même emplacement sur la plaque, ce qui est faisable mais demande un travail très conséquent, non envisageable pour la pérennité de la méthode.

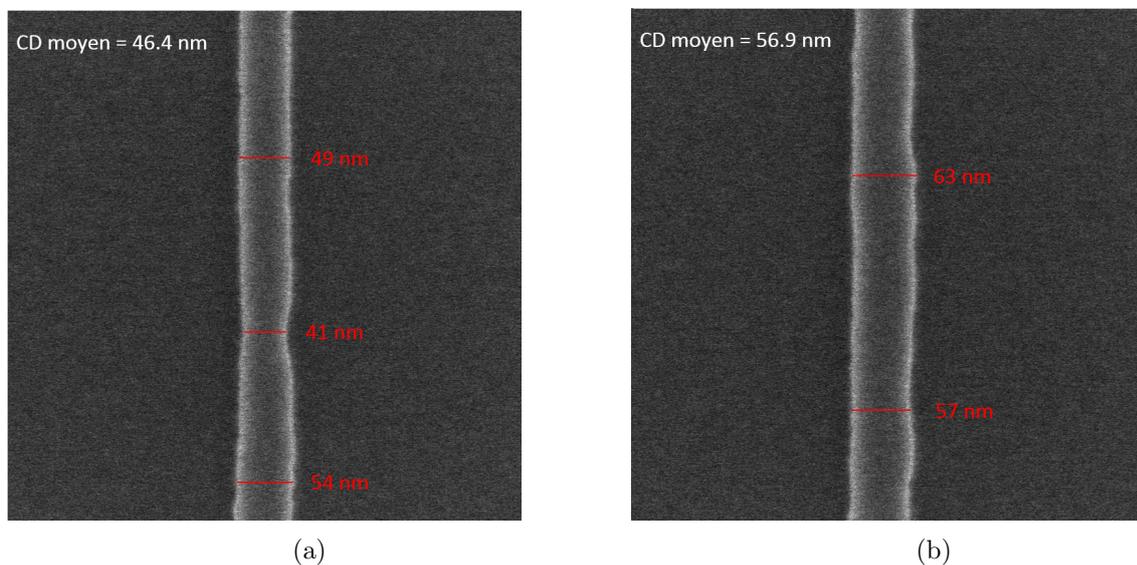


FIGURE 5.20 – Images SEM de lignes isolées planifiées à 50 nm. Comparaison du CD moyen sur l'image (blanc) et de la variation de CD le long des lignes (rouge).

De ce fait la meilleure comparaison que l'on puisse faire est par rapport à une mesure de CD-SEM classique, c'est à dire par l'utilisation d'une valeur de seuil de niveau de gris sur la longueur de la ligne. La Fig 5.21 permet de comparer le résultat de l'estimation du CDt par notre modèle non calibré, c'est à dire directement estimé avec la combinaison linéaire trouvée par JMONSEL, avec une mesure de CD-SEM classique. Cette mesure a été faite en utilisant les mêmes images que celles utilisées pour la calibration, mais seulement en vue du dessus.

Seule la moitié des mesures (P3 et P4) est montrée pour faciliter la lisibilité, mais les résultats de l'autre moitié sont équivalents.

On peut voir que les résultats des deux méthodes sont globalement cohérents ; cependant il est difficile d'aller plus loin dans l'analyse car il n'est pas possible de déterminer quelle méthode "se trompe plus que l'autre". On peut noter que les motifs expérimentaux ont un CRt très faible ; or on a vu que l'amélioration de mesure de CDt grâce au modèle était plus importante lorsque le CRt était grand.

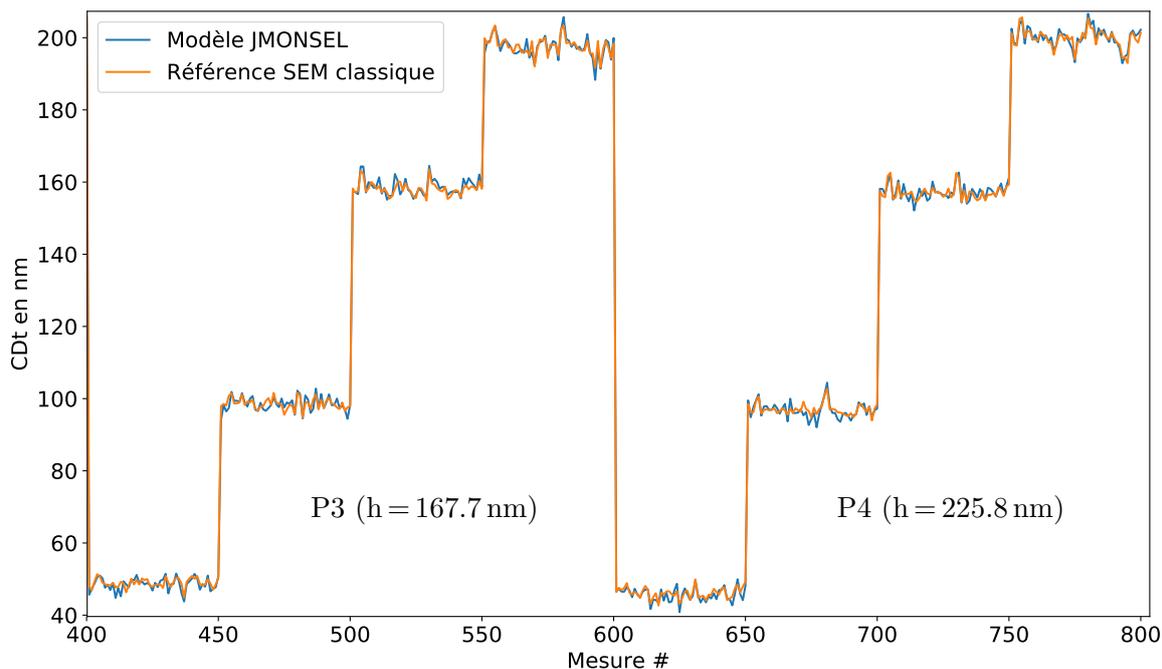


FIGURE 5.21 – Comparaison entre l'estimation du CDt par le modèle et une mesure CD-SEM classique. RMSE=1.4 nm.

Modèle calibré

Grâce au modèle calibré on va donc pouvoir estimer de manière quantitative les différentes grandeurs d'intérêts permettant de remonter à la topographie d'un motif observé par SEM à différents angles de tilt.

On a vu que chacune des grandeurs était estimée à une erreur près. Cependant ces erreurs peuvent parfois se compenser ou au contraire s'exacerber comme montré dans la Fig 5.22 pour un motif de la plaque P1 et dans la Fig 5.23 pour un motif de la plaque P4.

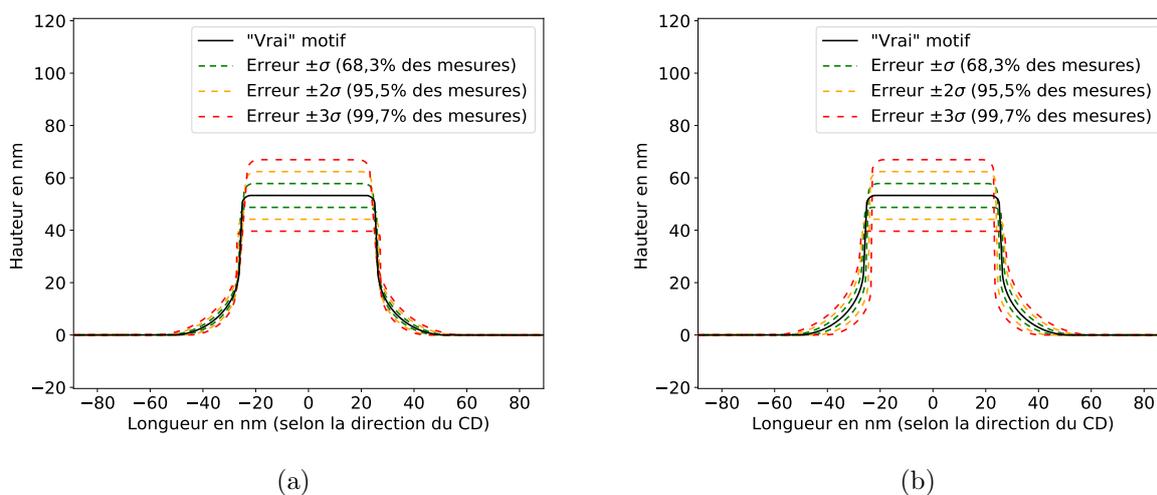


FIGURE 5.22 – Comparaison entre une reconstruction où les erreurs se compensent (a) et une reconstruction où les erreurs s'amplifient (b), sur un motif de type plaque P1.

Les deux cas présentés sont statistiquement très improbables, donc en moyenne on devrait

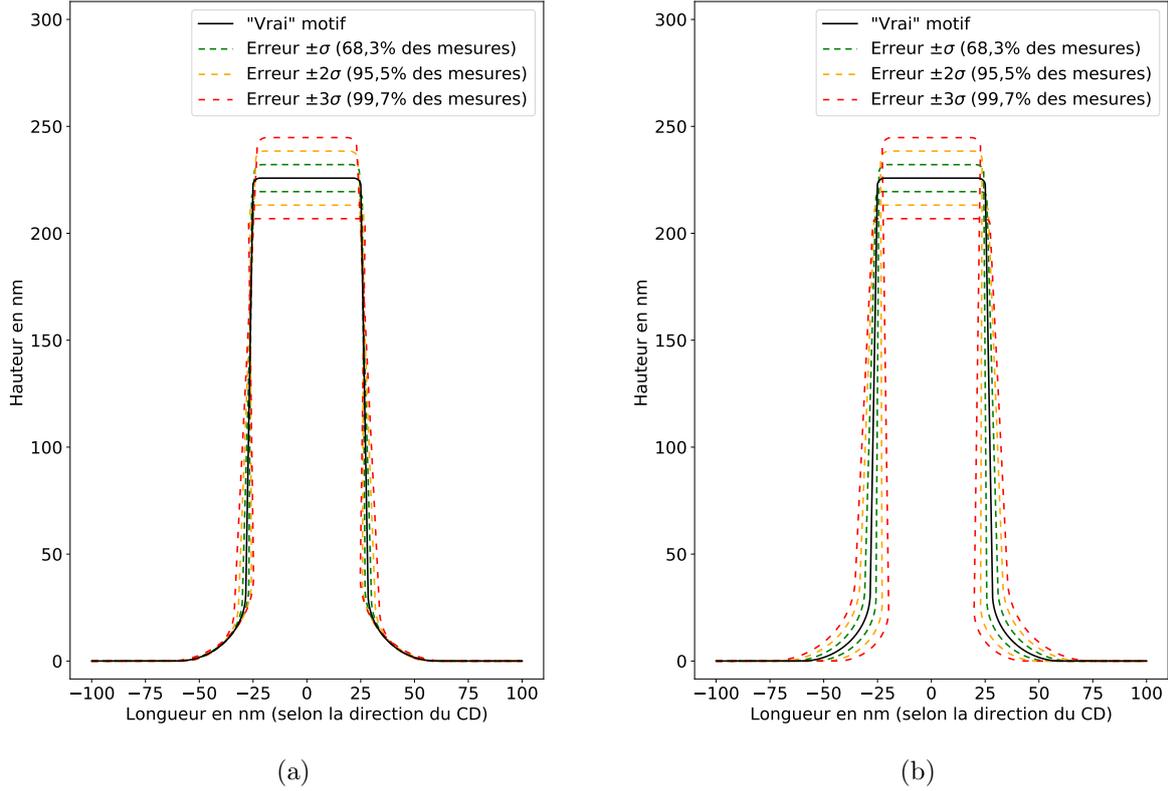


FIGURE 5.23 – Comparaison entre une reconstruction où les erreurs se compensent (a) et une reconstruction où les erreurs s’amplifient (b), sur un motif de type plaque P4.

se trouver quelque part entre ces deux cas extrêmes. Cependant comme le modèle est physique et que les descripteurs partagent souvent des frontières communes, il est physiquement plus probable que l’on se trouve plus proche du cas où les erreurs se compensent.

Finalement, on voit que le point critique de la reconstruction topographique est la valeur de hauteur. On rappelle que la formule de l’incertitude sur la hauteur due à l’incertitude de l’estimation des EW est donnée par (prouvé en partie 7.3) :

$$\sigma_h^2 = \frac{\sigma_{EW1}^2 + \sigma_{EW2}^2}{(\tan(T2) - \tan(T1))^2} \quad (5.36)$$

avec l’hypothèse que $\sigma_{EW1}^2 \approx \sigma_{EW2}^2$ et que l’on travaille avec le couple de tilt $0^\circ/12^\circ$,

$$\sigma_h^2 = \frac{2\sigma_{EW}^2}{\tan(12)^2} \quad , \text{ d'où } \sigma_{EW}^2 = \frac{\tan(12)^2 \sigma_h^2}{2} \quad (5.37)$$

On a vu que l’erreur d’estimation de la hauteur par le modèle par rapport à une référence AFM en utilisant le parallaxe maximum était de : $\text{RMSE}_h = 5.6 \text{ nm} \geq \sigma_h$, ce qui nous donne σ_{EW} majoré par 0.84 nm , ce qui est cohérent avec le $\text{RMSE}_{EW} = 1.2 \text{ nm}$ trouvé en Fig 5.17 (qui présentait des biais non négligeables comme étudié). On voit donc qu’une erreur faible sur la mesure d’EW entraîne une erreur importante d’estimation de la hauteur.

Si on voulait par exemple pouvoir assurer une mesure de hauteur dont l’écart-type serait de 1 nm , alors il faudrait que l’estimation des EW ait un écart-type de 0.15 nm , soit moins de $1/3$ de la taille des pixels des images SEM utilisées.

Une autre solution est d’augmenter le parallaxe, c’est à dire d’utiliser un SEM dont l’angle de tilt maximum est plus important. Par exemple un angle de tilt à 30° permettrait d’obtenir un

$\sigma_h = 2 \text{ nm}$, avec la même incertitude sur l'estimation d'EW que le modèle actuel.
Enfin la dernière technique est de multiplier le nombre de mesures SEM, soit au même endroit, mais dans ce cas il faut être sûr de ne pas altérer le motif pour rester dans le cas de mesures indépendantes, soit à différents endroits si l'utilisateur estime que l'hypothèse de faible variabilité est acceptable sur telle ou telle grandeur.

5.3 Reconstruction de motifs grâce au modèle INVSEM 3D

Dans cette partie on va utiliser la partie du programme informatique (Voir Appendices I.) destinée à l'utilisateur final ayant déjà un modèle créé.

On va donc surtout pouvoir se rendre compte du résultat visuel qu'il est possible d'obtenir grâce à cette technique de reconstruction. Les résultats du modèle de la partie précédente ne devraient pas changer tant que l'on reste dans le domaine de validité du modèle.

En tant qu'utilisateur il est donc possible de donner deux images ou plus d'un même motif imagé à différents angles de tilt, l'angle de tilt devant être renseigné, en général par la dénomination de l'image lors de l'enregistrement par le SEM. L'image en vue du dessus est, comme on l'a vu, indispensable pour pouvoir estimer les CR. Si l'utilisateur décide de travailler avec deux images, plus l'angle de tilt de la seconde image sera grand plus la reconstruction sera précise. Si plus de deux images sont données au modèle alors tous les résultats correspondant aux différentes combinaisons de couples de tilt possibles pourront être calculées.

On rappelle qu'au tout début de l'analyse, la rugosité de ligne est artificiellement supprimée du côté que l'on analyse avec le tilt, grâce à la fonction "ruban". Ceci permet au modèle d'estimer les CR, la hauteur et le SWA sans être perturbé par cette rugosité. En revanche le CDt qui est estimé par le modèle est de ce fait le CD moyen sur la longueur de l'image. En fin d'analyse on peut donc récupérer le profil correspondant à la rugosité de la ligne à droite et à gauche, de manière à adapter le CD moyen en fonction de l'axe de la ligne. Pour cela on va d'abord calculer le centre du motif pour chaque ligne de pixel :

$$x_{\text{centre}_i} = x_{\text{shift}} + \frac{(x_{l,i} - \bar{x}_l) + (x_{r,i} - \bar{x}_r)}{2} = x_{\text{shift}} + \frac{\Delta_i(x_l) + \Delta_i(x_r)}{2} \quad (5.38)$$

avec x_{centre_i} le centre du motif à la $i^{\text{ème}}$ ligne de pixels, \bar{x} la position moyenne du ruban et x_i sa position à la $i^{\text{ème}}$ ligne. Les indices indiquent le côté de gauche (l) ou de droite (r).

Puis on modifiera le CD local autour de la valeur moyenne estimée CD_{calib} :

$$CD_i = CD_{\text{calib}} - (\Delta_i(x_l) - \Delta_i(x_r)) \quad (5.39)$$

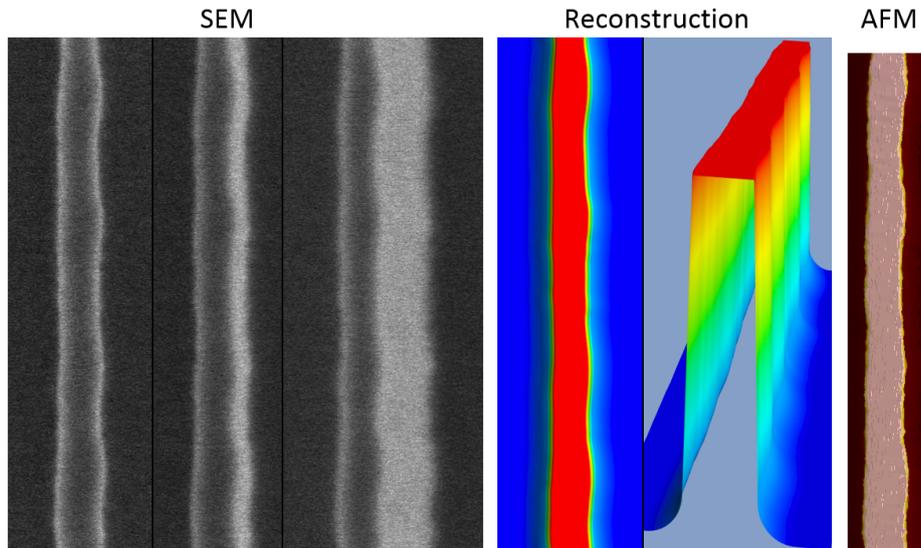


FIGURE 5.24 – Résultat de la reconstruction 3D, à partir des images SEM. On peut voir que la rugosité de ligne a été récupérée grâce au ruban et est conforme à la mesure AFM.

La Fig 5.24 montre un exemple de reconstruction faite à partir de trois images SEM de tilt 0° , 3° et 12° . Ces images sont visibles à gauche de la figure. Au milieu de la figure est représenté

le résultat 3D de la reconstruction, grâce à l'utilisation de la bibliothèque de visualisation 3D "Mayavi" [63], tout d'abord en vue du dessus pour pouvoir apprécier la restauration de la rugosité de ligne puis en vue isométrique. A droite la vue du dessus d'une reconstruction par méthode AFM permet également d'observer la rugosité de ligne par une autre méthode que la mesure SEM.

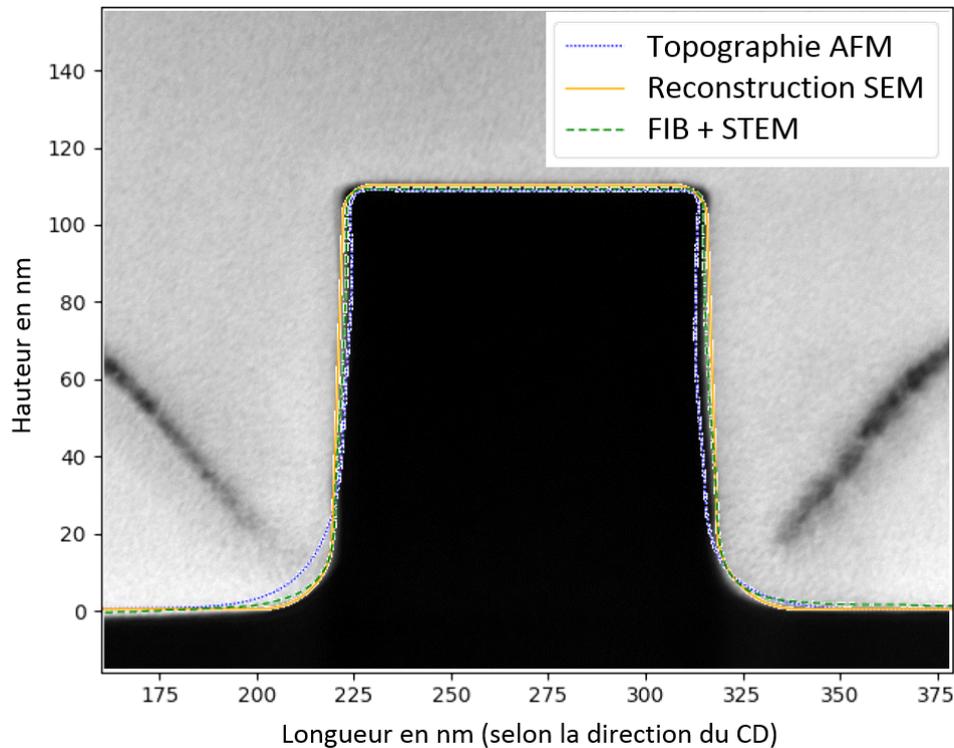


FIGURE 5.25 – Comparaison de la reconstruction SEM par rapport à de l'AFM et du FIB+STEM, avec l'image du STEM en arrière plan.

La Fig 5.25 permet de comparer le résultat d'une reconstruction par le modèle par rapport à une mesure AFM ainsi qu'une mesure issue d'une lamelle FIB imagée en STEM. L'image de la mesure STEM a été ajoutée en arrière-plan du graphique. Comme il est très long de faire une mesure FIB/STEM seules quelques coupes ont été réalisées. La comparaison à une méthode d'imagerie sur une coupe transversale (cross-section) a été intéressante pour s'apercevoir que la méthode de référence par AFM pouvait avoir des limitations par rapport à la mesure des CRb, comme on peut le voir sur la Fig 5.25, sur le CRb de gauche. Ce type d'erreur sur la valeur du CRb trouvé par AFM est typique d'une pointe abimée. On peut noter que l'erreur engendrée sur la valeur du CRb sera forcément positive, ce qui peut être intéressant dans la prise de décision de la meilleure mesure du CRb par mesure AFM (valeur minimale à la place de valeur moyenne par exemple).

Finalement, des reconstructions ont été réalisées sur des motifs n'ayant pas servi à la calibration du modèle, mais provenant des mêmes plaques. La Fig 5.26, permet de voir le résultat de l'estimation de la hauteur pour trois couples d'angles de tilt différents. Il est important de noter que ce résultat n'a pas été fait avec la dernière version du modèle, mais avec un modèle prenant en compte moins d'effets que celui présenté dans les parties précédentes. Ceci explique notamment pourquoi le couple $0^\circ/12^\circ$ présente un biais négatif pour les motifs de hauteur élevée. En effet le descripteur "BD2" (Eq 5.19, dont on peut voir l'effet sur les valeurs proches de

0 sur la Fig 5.9) qui permet de compenser la non nullité des EW, n'était pas encore implémenté et de ce fait les EW à tilt nul étaient surestimés, ce qui mène à une sous-estimation de la hauteur.

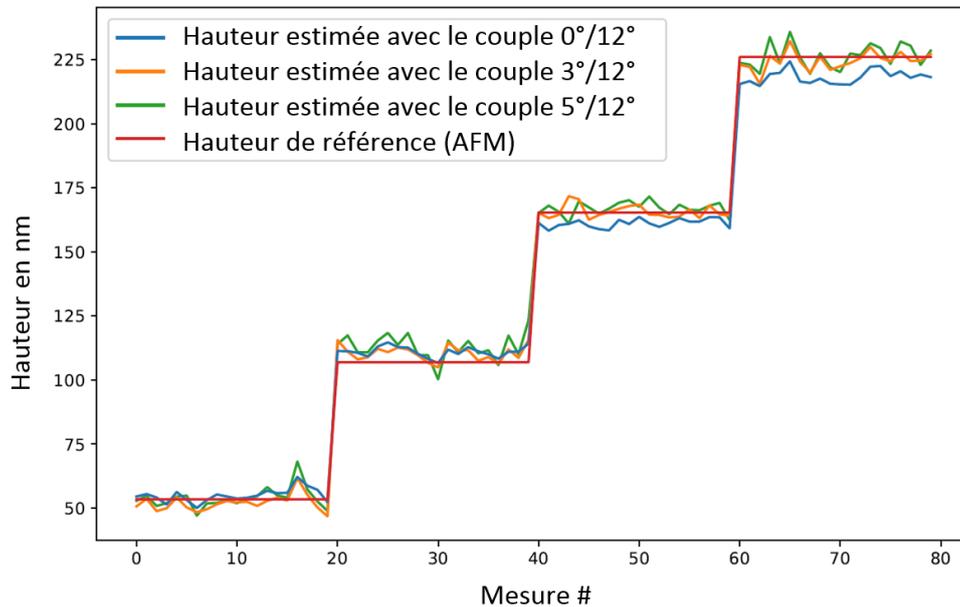


FIGURE 5.26 – Hauteur trouvée par l'utilisation du modèle INVSEM3D pour différents couples d'angle de tilt et comparaison avec la hauteur moyenne par plaque mesurée par AFM.

Il serait intéressant de tester le modèle actuel sur des motifs ne provenant pas des plaques sur lesquelles la calibration a été faite de manière à voir le comportement du modèle, en interpolation (cad sur des valeurs de grandeurs d'intérêt contenues entre les bornes de celles utilisées pour la calibration) et en extrapolation (c'est à dire sur des valeurs de grandeurs d'intérêt plus grandes ou plus petites que les bornes de celles utilisées pour la calibration).

5.4 Conclusion

Dans ce chapitre, nous avons présenté la méthode générale développée, pour la création d'un modèle linéaire permettant de trouver les grandeurs d'intérêt à partir de la mesure de descripteurs provenant d'images SEM. Dans un second temps, nous avons créé un modèle spécifique à la reconstruction de motifs de type "ligne" en silicium gravé. Puis nous avons calibré ce modèle sur des mesures SEM réelles en les comparant avec des mesures de référence AFM. Enfin, grâce à ce modèle créé et calibré, quelques résultats de reconstruction sur des motifs réels ont été présentés.

Dans la partie suivante on va à nouveau utiliser la méthode de création de modèle présentée ici, mais pour créer un modèle qui permettra de reconstruire des motifs différents : les tranchées ainsi que les motifs denses.

Chapitre 6

Extension du modèle aux tranchées isolées et aux denses

Depuis la partie JMONSEL (Partie 4), on s'est concentré sur l'optimisation d'un modèle dédié à la reconstruction de motifs de type lignes isolées. Cependant pour pouvoir répondre à un plus grand nombre d'applications il est important que le modèle puisse également reconstruire des motifs de type tranchées isolées et des motifs denses (succession de lignes suffisamment proches pour qu'on ne puisse pas les considérer comme isolées (voir ombrage partie 4.7)). Si le modèle est capable de reconstruire ces trois types de structures, il permettra alors de répondre à la majorité des problématiques de mesure tri-dimensionnelle de l'industrie de la microélectronique. Un premier étage d'analyse pourrait découper les images SEM en sous-parties correspondant à un de ces trois types de motif (non réalisé dans cette thèse). Ensuite chacune de ces sous-parties serait reconstruite par le modèle correspondant.

6.1 Extension du modèle aux tranchées

6.1.1 Comparaison des descripteurs utilisés pour les lignes et pour les tranchées

La Fig 6.1 permet de voir à quoi vont ressembler les profils d'un motif de type tranchée, en vue de dessus et à un angle de tilt de 12° , ainsi que le positionnement des descripteurs sur ces profils. On va utiliser les mêmes descripteurs pour analyser les motifs de type tranchée que les descripteurs utilisés pour analyser les motifs de type ligne. Cependant ils seront toujours situés de la même manière sur le front, c'est pourquoi ils sembleront inversés par rapport à un motif de type ligne. Cependant cette inversion ne doit pas être vue comme une inversion des descripteurs, mais comme une inversion du front montant et descendant.

On peut voir que le rendement électronique au centre du motif, c'est à dire au fond de la tranchée est faible ; en effet les électrons vont avoir du mal à s'échapper d'un motif de type tranchée car les électrons ressortant du matériau auront une probabilité importante d'être recaptés par les parois de la tranchée dû au faible angle solide de sortie. Cet effet sera d'autant plus grand que le facteur de forme du motif sera grand, c'est à dire que le rapport entre sa profondeur (que l'on notera hauteur pour être homogène avec les motifs de type ligne) et l'ouverture de la tranchée (CDt) sera grand.

Grâce à ces descripteurs, on va pouvoir créer un modèle par l'analyse des simulations JMONSEL de la même manière que pour les lignes. Cependant, contrairement aux motifs de type "ligne" dont la reconstruction tri-dimensionnelle n'a théoriquement pas de limitations liées aux dimensions du motif, la reconstruction de motifs de type "tranchée" ne peut être réalisée que si le faisceau électronique tilté est géométriquement en mesure de sonder le bas du flanc analysé. Il y a donc une limitation liée au facteur de forme du motif, c'est à dire au rapport entre sa

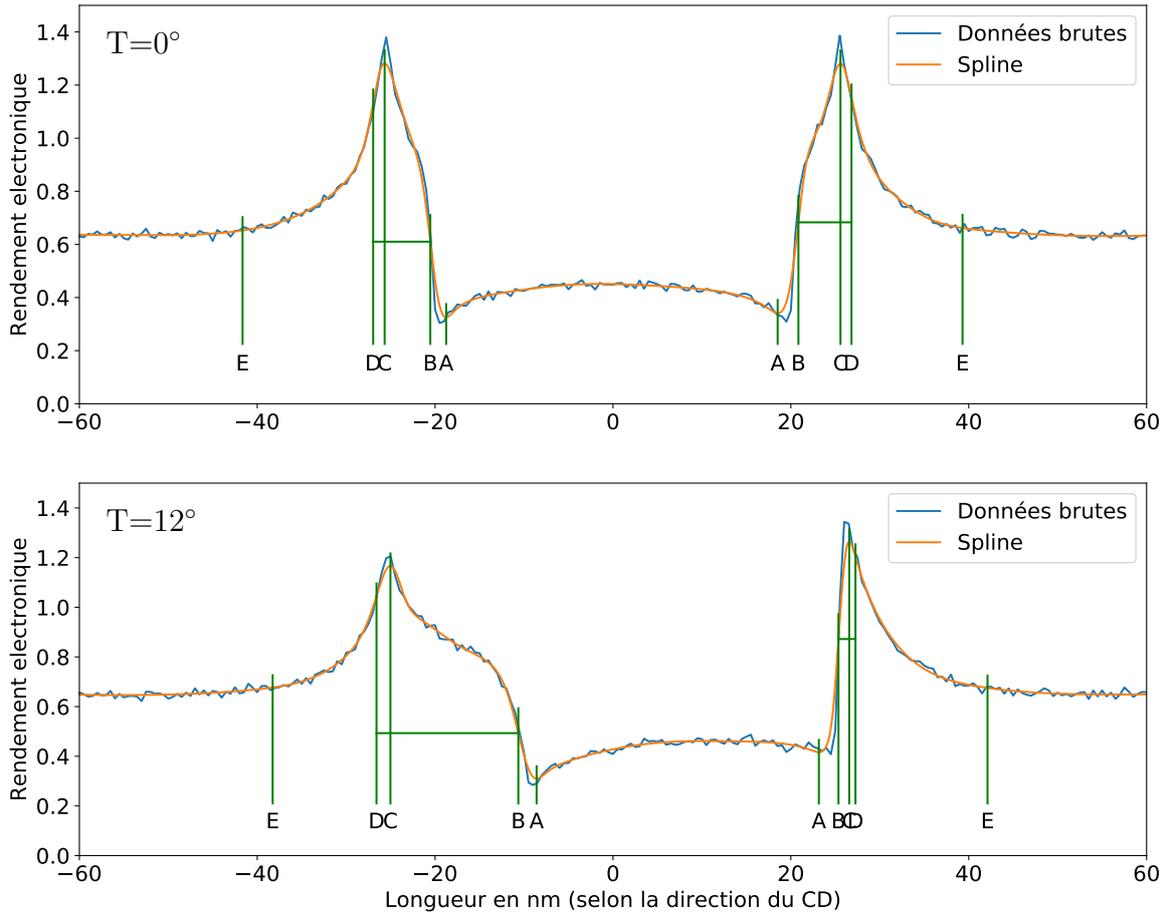


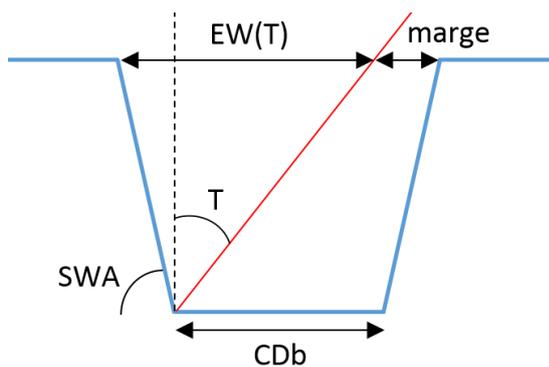
FIGURE 6.1 – Placement des descripteurs sur le profil SEM d’un motif de type tranchée imagé en vue du dessus (haut) et à un angle de tilt de 12° (bas). Motif : $h = 50$ nm, $SWA = 84^\circ$, $CD_b = 40$ nm, $CR = 0$ nm.

profondeur h et l’ouverture de la tranchée (le CD_t) tel que :

$$CD_t > EW(T) \Leftrightarrow CD_t > \frac{h}{\tan(SWA)} + h \tan(T) \quad (6.1)$$

ou

$$\text{Comme } CD_t = CD_b + 2 \frac{h}{\tan(SWA)}, \text{ alors } CD_b > h \tan(T) - \frac{h}{\tan(SWA)} \quad (6.2)$$



Dans la réalité une marge doit être ajoutée à la limite géométrique ci-dessus, pour pouvoir détecter de manière robuste le bas de la pente (Fig 6.2). D’après notre expérience sur des résultats de simulations JMONSEL, nous proposons une valeur de marge minimale de 5 nm sur des motifs à CR nuls.

FIGURE 6.2 – Représentation de la limitation de facteur de forme d’une tranchée.

Comme nous allons le voir dans la partie sur l'analyse des simulations JMONSEL sur des motifs de type tranchée, la reconstruction est fortement compliquée par le changement de comportement de la réponse électronique du SEM, en fonction du facteur de forme des motifs, c'est à dire en fonction du niveau de confinement des électrons dans la tranchée. La réduction du nombre d'électrons collectés avec l'augmentation du facteur de forme est visible dans la Fig 6.3. On peut observer que pour le motif de 50 nm de hauteur et de 200 nm de CDb, le rendement électronique en milieu de tranchée remonte pratiquement au niveau "surface plane infinie", ce qui est cohérent avec les résultats de l'analyse sur l'effet d'ombrage (Partie 4.7).

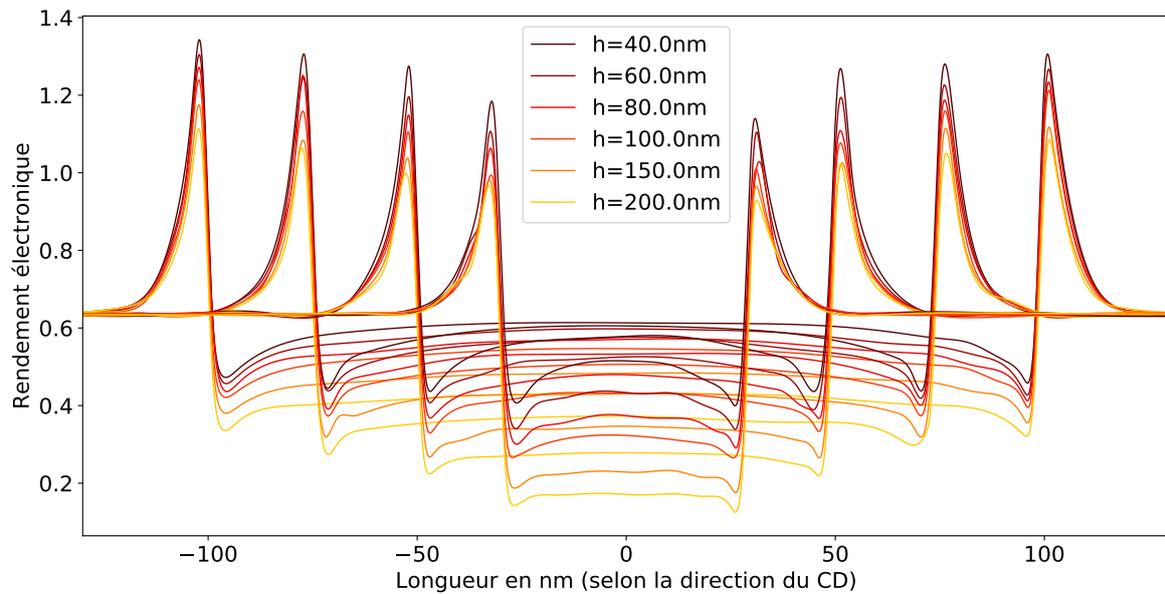
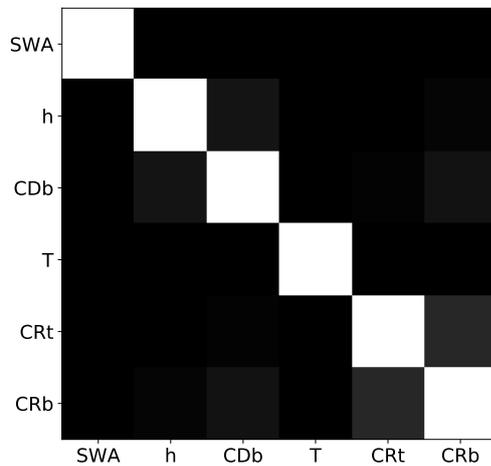


FIGURE 6.3 – Résultats de simulations JMONSEL de tranchées. SWA = 90°, CDb = [60, 100, 150, 200] nm, h = [40, 60, 80, 100, 150, 200] nm.

6.1.2 Analyse JMONSEL pour les tranchées

De la même manière que pour les motifs de type ligne, une analyse basée sur un grand nombre de simulations JMONSEL a été réalisée pour créer le modèle adapté aux motifs de type tranchée. Un total de 2000 motifs différents ont été simulés avec les grandeurs suivantes :



$h = [50, 100, 150, 200]$ nm
 $SWA = [84, 86, 88, 89, 90]^\circ$
 $CDb = [40, 60, 80, 100, 150]$ nm
 $CRb/CRt = [0/0, 5/30, 10/20, 20/5, 30/10]$ nm
 $T = [0, 3, 5, 12]^\circ$

Pour réduire le nombre de simulations nous n'avons pas simulé tous les couples de CRb/CRt possibles mais seulement des couples "au hasard". Certaines des combinaisons formées ci-dessus ne sont pas valides et ont donc été enlevées soit à cause de la limitation de facteur de forme (Eq 6.1) soit car le CRb était trop grand par rapport au CDb (forme invalide si $2CRb > CDb$).

Ce jeu de données est de meilleure qualité que celui utilisé pour les lignes car la corrélation des CR est réduite à 0.16 (Fig 6.4).

FIGURE 6.4 – Matrice de corrélation des grandeurs d'intérêt (\tilde{g}_u, \tilde{g}_v).

La robustesse de positionnement des descripteurs sur des motifs de type tranchée est bien plus difficile que pour des motifs de type ligne. En effet, dans le cas des lignes, la forme générale du signal SEM est relativement régulière quelque soit la valeur des grandeurs d'intérêts. En revanche pour les espaces, l'angle solide de sortie des électrons étant réduit par le facteur de forme du motif, on pourra observer des comportements suffisamment différents du signal SEM, à l'origine d'erreurs dans le positionnement des descripteurs.

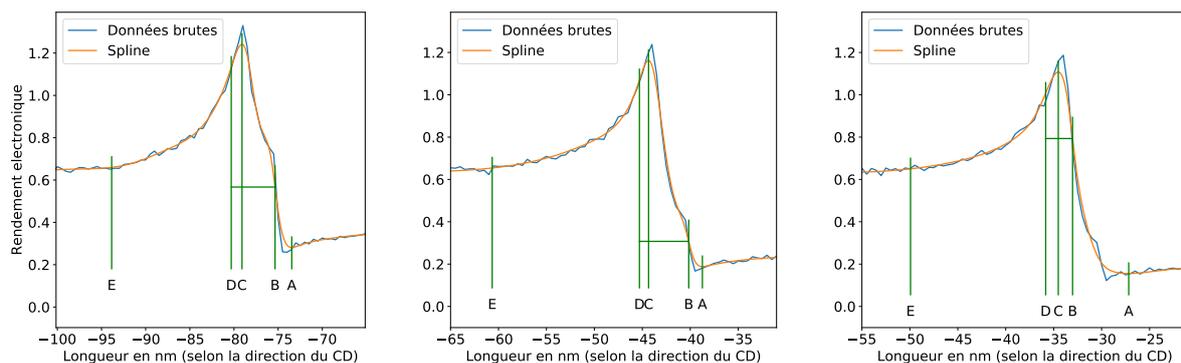


FIGURE 6.5 – Comparaison de la simulation JMONSEL d'un motif tranchée de hauteur 200 nm et 89° de SWA, pour des CDb de 150 nm, 80 nm et 60 nm (de gauche à droite).

Par exemple la Fig 6.5, montre la diminution de l'information du bas de pente lorsque le facteur de forme augmente. Les trois figures correspondent à la réponse électronique d'un SEM sur le flanc gauche d'un espace de hauteur 200 nm et de 89° de SWA. Le motif simulé pour obtenir le profil de gauche a un CDb de 150 nm, c'est pourquoi on obtient une forme qui ressemble à ce que l'on avait sur des flancs de lignes isolées, c'est à dire un maximum d'intensité électronique correspondant au passage de l'arrête du haut de flanc (descripteur C) puis une "épaule" correspondant au balayage du flanc (de C à B), jusqu'à une descente

rapide correspondant au bas du flanc (en B). Cependant dans le cas des espaces, le rendement électronique de "l'épaule" diminue plus ou moins rapidement en fonction du facteur de forme, rendant la pente finale moins importante. On peut donc voir sur le profil central la diminution de l'effet de bas de flanc à un CDb de 80 nm, mais que le descripteur B est encore bien placé. En revanche pour un CDb de 60 nm l'effet de fin de flanc est tellement réduit qu'il est lissé par la spline menant à une erreur de placement du descripteur B.

Il s'agit là d'un exemple parmi une multitude, de la réduction de la visibilité des effets à cause du facteur de forme. C'est pourquoi, dans les analyses JMONSEL qui vont être présentées, on observera bien plus souvent que pour les motifs de type ligne des valeurs aberrantes. De plus lorsque la valeur de CRb est non nulle cette limite de bas de pente, représentée par le descripteur B, va devenir encore plus complexe à déterminer.

De la même manière que pour les motifs de type ligne, les résultats de l'optimisation de l'estimation des grandeurs d'intérêt par méthode LASSO vont être présentés.

Optimisation du CRb

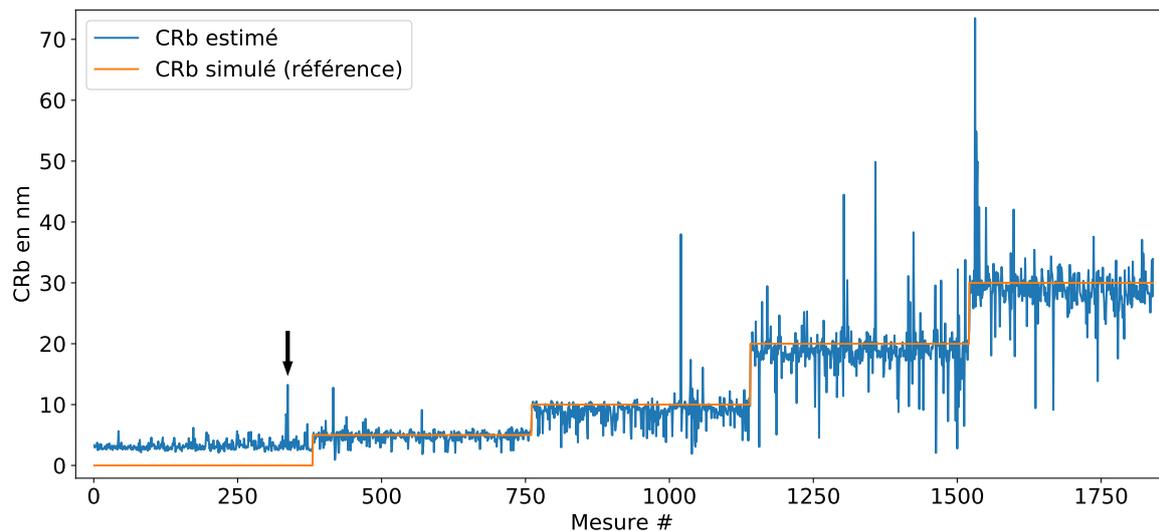


FIGURE 6.6 – Résultat de l'optimisation du CRb. RMSE = 3.47 nm.

Tout comme l'estimation du CRb des motifs de type ligne, c'est également la distance AB qui a été retenue par la régularisation du modèle, comme descripteur permettant l'estimation, telle que :

$$\text{CRb}_{\text{JMO}} = 1.12 \times \text{AB} + 0.92 \quad (6.3)$$

Comme annoncé, un nombre non négligeable de valeurs sont fortement éloignées de la valeur de référence. Ces erreurs très importantes ne sont pas liées à une erreur du modèle mais à un mauvais placement initial des descripteurs A et/ou B. Par exemple la surestimation forte du CRb, indiquée par une flèche sur la Fig 6.6, provient de l'erreur de placement étudiée en Fig 6.5, à droite. De même les autres fortes surestimations proviennent du même genre d'erreur de placement (une dizaine de cas sur les 1800 simulations) et sont plus probables pour les motifs à fort facteur de forme.

Ces valeurs aberrantes sont gênantes car lors de la résolution par pseudo-inverse, pour la détermination du poids des descripteurs, ces valeurs très éloignées vont détériorer le modèle

pour toutes les mesures dont le placement des descripteurs est correct.

Sinon, on observe le même type d'effet que pour les motifs ligne, comme par exemple l'impossibilité d'estimer des CRb en dessous de 4 nm et l'augmentation de l'erreur avec l'augmentation du CRb.

Optimisation du CRt

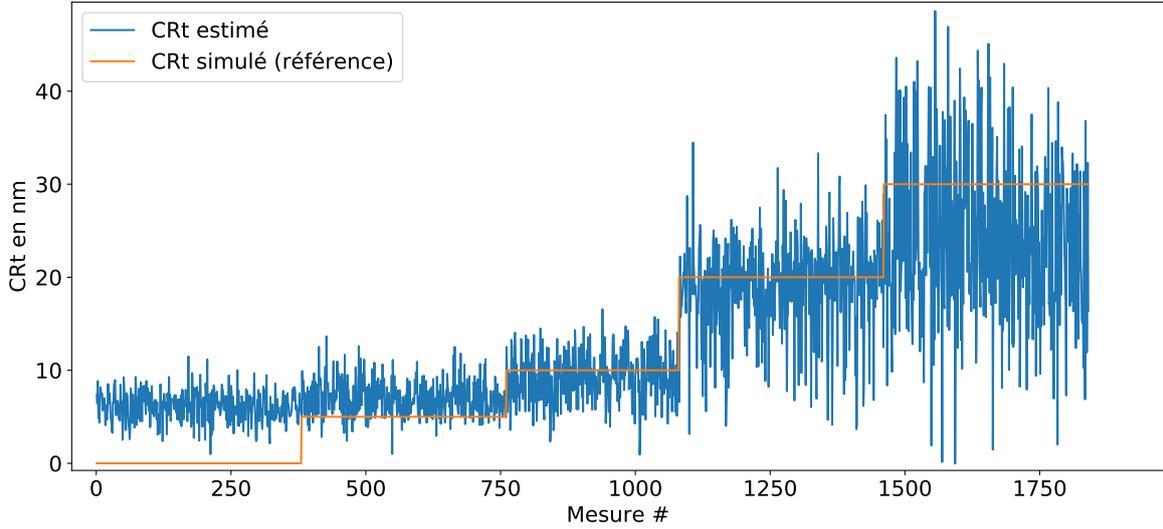


FIGURE 6.7 – Résultat de l'optimisation du CRt. RMSE = 6.39 nm.

Comme on peut le voir dans la Fig 6.7, l'estimation du CRt est, comme pour les lignes, difficile. Lorsque l'on laisse la régularisation choisir les descripteurs, la meilleure estimation trouvée est :

$$\text{CRt}_{\text{JMO}} = 0.07 \times \text{CE}^2 - 0.25 \times \text{CE} - 3.45 \times \text{intDE} \quad (6.4)$$

avec,

$$\text{intDE} = \int_D^E \text{profil}(x) dx - \text{DE} \times \min(\text{profil}(D), \text{profil}(E)) \quad (6.5)$$

On peut noter que la partie " $-\text{DE} \times \min(\text{profil}(D), \text{profil}(E))$ " permet d'éviter qu'une différence de luminosité de l'image, c'est à dire une translation du profil, n'altère le résultat. Pour la simulation cela ne sert à rien car on travaille directement avec le rendement électronique, mais pour le passage aux vrais images cela permet d'éviter des erreurs dues aux éventuels post-traitements du SEM. Cependant ce descripteur est toujours dépendant d'un changement de contraste de l'image.

On remarque que cette formule est sensiblement différente de celle qui a été trouvée pour un motif ligne (Eq 5.17). Cependant si on force le choix des mêmes descripteurs que ceux utilisés pour un motif ligne on obtient :

$$\text{CRt}_{\text{JMO}} = 0.07 \times \text{CE}^2 - 0.76 \times \text{DE} \quad (6.6)$$

Avec cette estimation le résultat est pratiquement équivalent à la Fig 6.7, avec un RMSE de 6.44 nm au lieu de 6.39 nm. Cette estimation pourrait donc être préférée, d'une part pour sa

similarité avec l'estimation du CRt des motifs lignes et par le fait qu'elle n'utilise pas de valeur absolue de niveau de gris.

Optimisation des EW

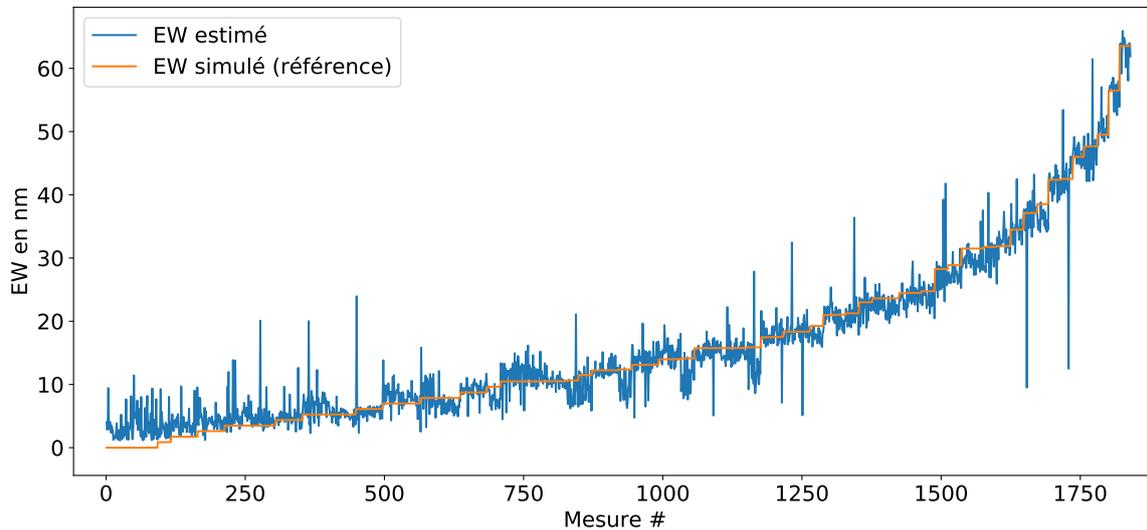


FIGURE 6.8 – Résultat de l'optimisation de l'EW. RMSE = 2.71 nm.

Pour l'estimation de l'EW (Fig 6.8), on observe à nouveau le problème de robustesse de placement des descripteurs, menant à des valeurs aberrantes. L'estimation de l'EW est donnée par :

$$EW_{JMO} = 0.09 \times BD2 + 0.08 \times AB + 0.70 \times BD\text{tilt} + 0.003 \times C\text{Reff}2 + 0.005 \times C\text{Reff}3 \quad (6.7)$$

avec BD2, CReff2 et CReff3, les mêmes fonctions que pour les motifs ligne (Eqs 5.19) et BDtilt tel que :

$$BD\text{tilt} = \frac{BD}{1 + \tan(T)} \quad (6.8)$$

Si on observe seulement les motifs dont les CR sont nuls (Fig 6.9), on peut observer que la présence de CR est le principal facteur d'erreur de placement des descripteurs primaires (par comparaison des Fig 6.9 et 6.8), que ce soient des valeurs aberrantes à cause du mauvais placement des descripteurs primaires ou du bruit de mesure dû à la variabilité du positionnement des descripteurs primaires. En effet, en leur absence, l'estimation est bien meilleure, avec un RMSE de 0.65 nm sur les 380 motifs à CR nuls.

L'estimation est également beaucoup plus simple :

$$EW_{JMO} = 1.02 \times BD - 1.19 \quad (6.9)$$

La clé de l'amélioration d'estimation de la hauteur et donc de la mesure des EW à CR quelconque est donc bien la recherche de descripteurs permettant de compenser les effets des CR.

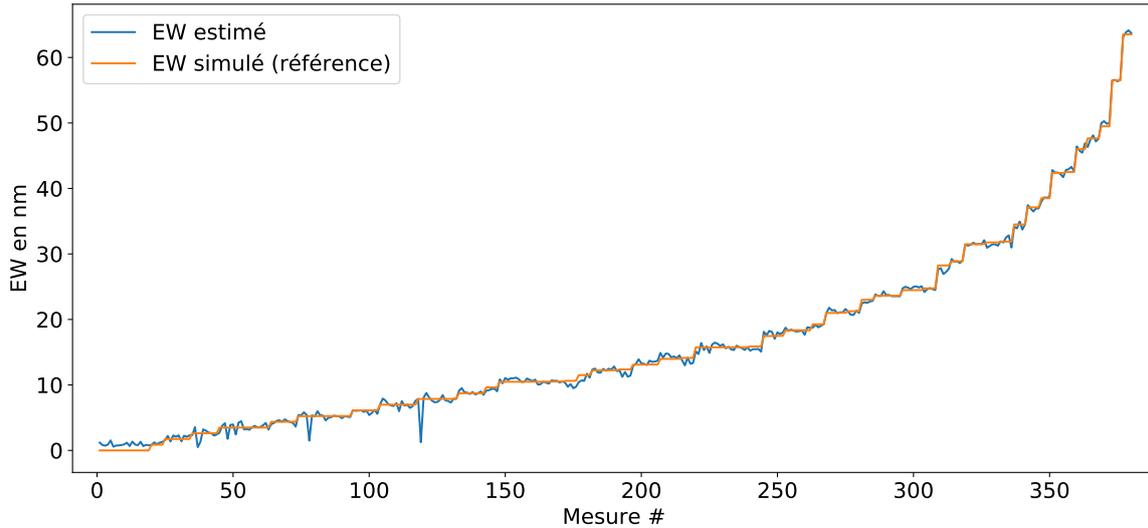


FIGURE 6.9 – Résultat de l’optimisation de l’EW à CR nuls. RMSE = 0.65 nm.

Calcul de la hauteur et du SWA

Finalement, on peut calculer la hauteur et le SWA grâce aux différents couples d’angle de tilt. On peut voir sur la Fig 6.10 que les erreurs sur l’estimation des EW se propagent sur le résultat du calcul de hauteur et de SWA. On peut voir que l’estimation de la hauteur pour les CR nuls est bien plus consistante que pour les CR non nuls.

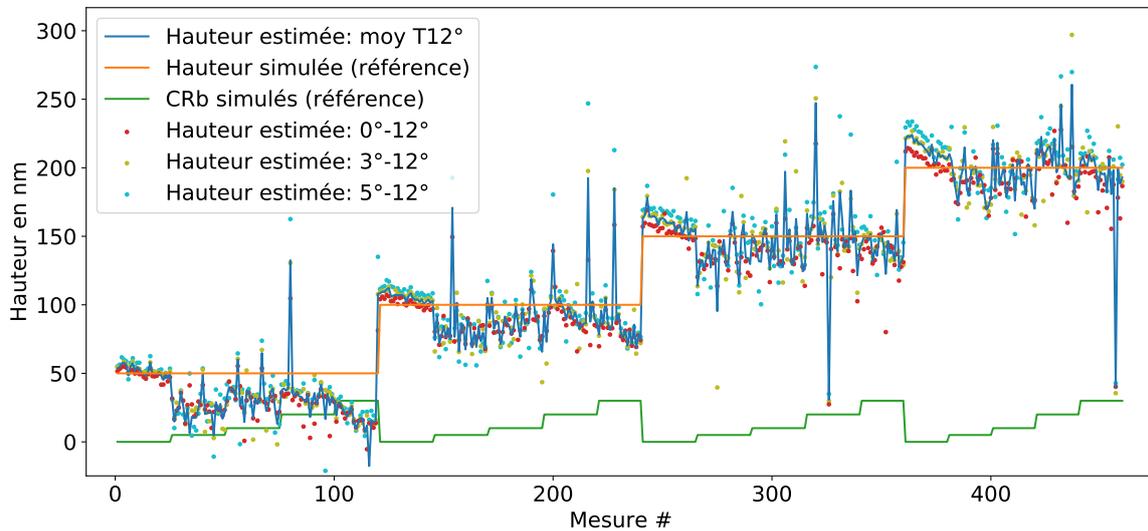


FIGURE 6.10 – Résultat du calcul de la hauteur à partir des EW estimés par le modèle. Les résultats provenant de trois couples d’angle de tilt sont présentés par des points de couleur et leur moyenne par le trait bleu. RMSE = 22.4nm.

Comme on utilise toujours l’image dont l’angle de tilt est le plus grand (12°), la plupart des valeurs aberrantes proviennent d’un mauvais positionnement des descripteurs primaires sur la simulation à 12° ; dans ce cas les trois hauteurs estimées sont toutes mal estimées. Lorsque l’erreur vient d’une simulation à un autre angle de tilt, le moyennage peut permettre

de compenser l'erreur.

Le résultat du calcul du SWA n'est pas proposé ici car il n'apporte pas beaucoup d'information supplémentaire (résultats corrects à CR nuls, résultats médiocres sur CR non nul, avec des valeurs aberrantes), le RMSE étant de $5,3^\circ$.

On peut noter que si l'on utilise le modèle d'estimation de l'EW à CR nul (donc seulement sur les motifs à CR nuls) de l'équation 6.9, on obtient des très bon résultats : RMSE de 4,0 nm sur la hauteur et de 0.4° sur le SWA.

Optimisation du CDt

Enfin l'estimation du CDt est donnée par :

$$\text{CDt}_{\text{JMO}} = 1.11 \times \text{CC} - 0.11 \times \text{DD} \quad (6.10)$$

On peut noter que l'on préfère estimer le CDt que le CDb car l'information est meilleure en haut du motif qu'au fond de la tranchée. La régularisation n'a pas gardé de termes liés au CRt ou au SWA comme pour les motifs de type ligne, mais c'est probablement à cause de leur mauvaise estimation. Si on améliorait les résultats de l'estimation du CRt et du SWA on pourrait probablement améliorer l'estimation du CDt.

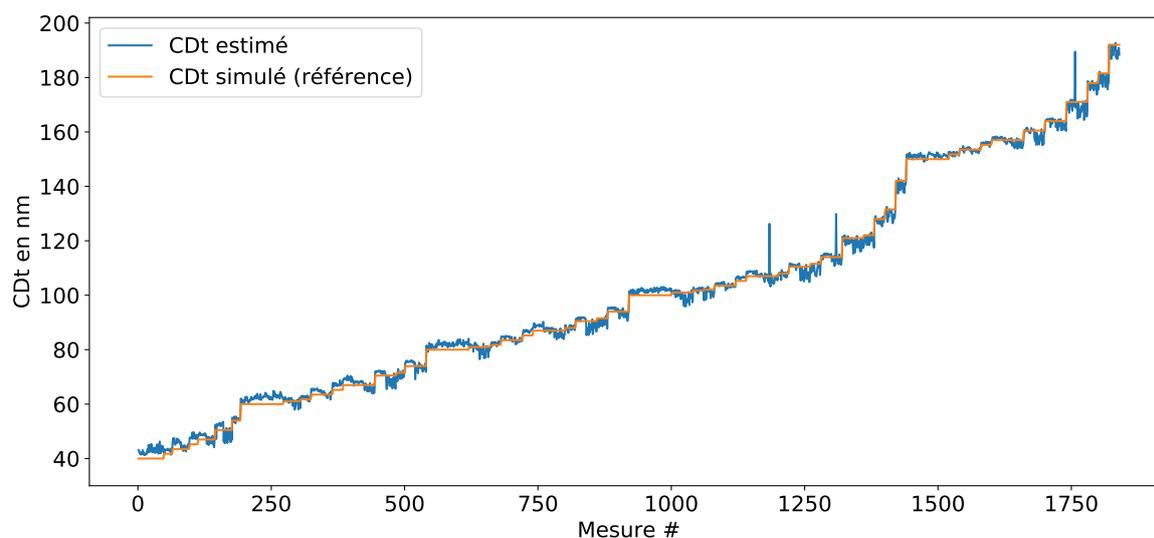


FIGURE 6.11 – Résultat de l'optimisation du CDt. RMSE = 2.05 nm.

On peut ici voir trois valeurs aberrantes qui sont toutes trouvées pour le CRt à 30 nm. Encore une fois, si l'on s'intéresse seulement aux motifs à CR nul, le RMSE descend à 0.77 nm.

Calibration pour les tranchées

On dispose désormais d'un modèle JMONSEL adapté pour les tranchées. Cependant il faudrait ensuite le calibrer sur des mesures de référence. Ce travail n'a pas pu être achevé et ne sera donc pas présenté.

De plus il faudrait d'abord rendre le modèle JMONSEL meilleur avant de le calibrer, la version actuelle étant trop instable, rendant la recherche de descripteurs secondaires fins difficile. Des

propositions d'améliorations de la méthode seront présentées dans la partie sur les perspectives [8.2](#).

D'un point de vue de la métrologie de référence, les tranchées seront également un challenge intéressant, car la mesure AFM par exemple ne peut plus atteindre le fond des tranchées lorsque le facteur de forme devient trop important. De plus dans le cas des tranchées, la hauteur de gravure à de forte chance d'être dépendante du CDt, augmentant la localité sur les hypothèses d'égalité sur la hauteur voire du SWA et des CR.

6.2 Approximation d'un motif dense par une succession de motifs tranchés

On peut simplifier la problématique en prouvant que les motifs de type dense peuvent être équivalents à une concaténation de motifs de type tranchée, dans un certain domaine de validité. En effet, lors du passage du faisceau électronique dans une tranchée N, si les électrons ne passent pas au travers de la matière qui les séparent des tranchées N+1 et N-1, alors le signal électronique résultant sera exactement le même, que les tranchées N+1 et N-1 existent ou non. Dans ce cas la reconstruction d'un motif de type dense pourra être réalisée grâce au modèle des tranchées isolées, appliqué autant de fois que de zones "tranchée".

Il s'agit désormais de déterminer le domaine de validité de l'approximation précédente ; en effet on sent bien qu'en dessous d'une certaine taille de la zone de matière séparant deux tranchées, on ne pourra plus s'affranchir de créer un modèle spécifique aux motifs de type dense. Pour ce faire nous avons utilisé le logiciel JMONSEL pour visualiser la taille de la poire d'interaction pour différentes topographies de tranchée.

La Fig 6.12 permet de visualiser la taille de la zone d'interaction des électrons sur deux motifs différents pour une énergie d'arrivée des électrons primaires de 800 eV, simulée par JMONSEL. Les électrons primaires sont envoyés sur le flanc gauche. En bleu est représenté une éventuelle tranchée adjacente, à la limite de cette poire d'interaction, correspondant à environ 35 nm.

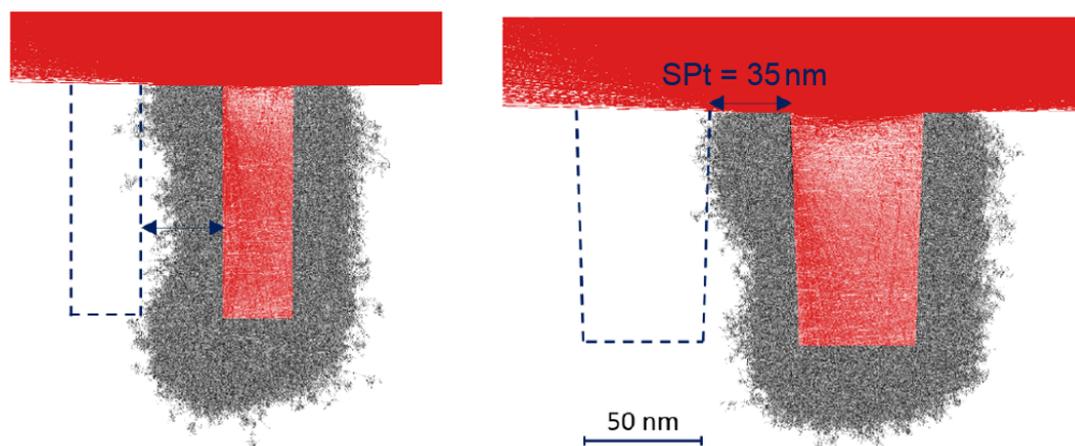


FIGURE 6.12 – Visualisation de la poire d'interaction des électrons sur deux motifs différents pour une accélération de 800 eV. En bleu est représentée une tranchée virtuelle adjacente, à la limite de cette poire d'interaction. La flèche bleue a une largeur de 35 nm.

Cependant il est difficile avec cette représentation de savoir exactement quand le signal électronique sera suffisamment perturbé par les tranchées adjacentes pour que l'approximation ne soit plus valide. C'est pourquoi on a également comparé les signaux SEM provenant d'une tranchée isolée et de la même tranchée entourée par deux autres de même forme, en diminuant progressivement la taille séparant les tranchées. Pour cela 126 simulations représentant 63 motifs imagés à deux tilts ont été réalisées :

Constante : $CD_b = 50 \text{ nm}$, $CR_b = CR_t = 0 \text{ nm}$
Variables : $SP_t = [50, 40, 30, 25, 20, 15, 10] \text{ nm}$
 $h = [50, 100, 200] \text{ nm}$
 $SWA = [90, 88, 85]^\circ$
 $T = [0, 12]^\circ$

Dans le cas des motifs denses, nous considérons que le CD correspond désormais à la taille du CD de la tranchée et le SP correspond à l'espace entre deux tranchées, ce qui est la grandeur importante de cette analyse.

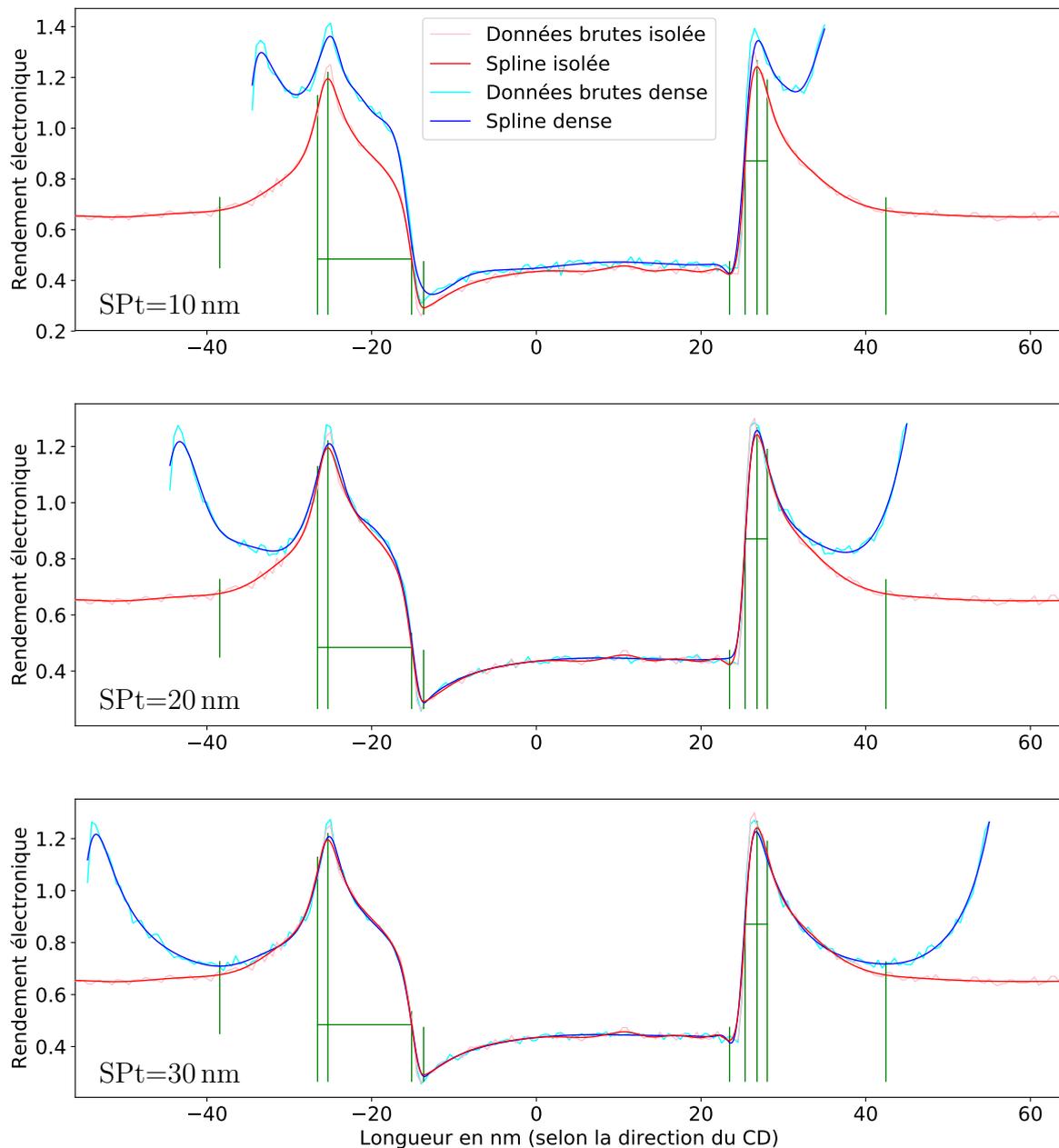


FIGURE 6.13 – Comparaison entre un motif tranchée isolée et un motif dense, pour différents $SPt=[10, 20, 30]$ nm. $h = 50$ nm, $SWA = 90^\circ$, $T = 12^\circ$, $CDb = 50$ nm.

Après avoir observé ces simulations, il apparaît que la hauteur influe très peu sur la limite du domaine de validité de l'approximation "dense = $n \cdot$ tranchées". En revanche plus le SWA diminue moins les électrons traversent la barrière entre deux tranchées, ce qui est logique car la diminution du SWA augmente le CD moyen sur la hauteur du motif (visible sur la Fig 6.12 à droite). Enfin plus l'angle de tilt sera important plus les électrons pourront facilement traverser la barrière, car ils arriveront avec une composante de vitesse horizontale plus importante à fort tilt.

C'est pourquoi la Fig 6.13 permet d'observer le motif "le plus limitant" pour différents SPt. On peut voir que plus la valeur de SPt augmente plus les courbes isolées et denses vont se rapprocher puis se confondre sur une grande distance.

On avait déjà prouvé que le CDt minimum que la méthode était capable de reconstruire était d'environ 20-25 nm (Partie 4.5). Or on peut observer sur la figure que c'est environ à partir de ce SPt que les courbes se confondent aux emplacements des descripteurs A, B, C et D. Le descripteur E quant à lui ne se superpose qu'à partir de 30-35 nm. C'est pourquoi si on utilise le modèle sur des motifs denses dont le SPt est supérieur à 30-35 nm, alors on pourra faire l'approximation qu'un motif dense est en fait une succession de motifs tranchées.

On rappelle que ces résultats sont uniquement valables à $V_{acc}=800$ eV. Si V_{acc} est réduit, ce qui est la tendance pour les études topographiques par SEM pour rendre la mesure la plus locale possible, cette limite sera réduite car la poire d'interaction sera plus petite.

6.3 Conclusion

Dans ce chapitre, nous avons utilisé la même méthode de création de modèle que celle présentée au chapitre 5, mais sur des données provenant de simulations de motifs de type "tranchée". On a vu que pour ce type de motif il était plus difficile d'obtenir des informations précises à cause de la réduction de l'information électronique dans les zones confinées, c'est à dire au fond des tranchées. On a vu que ce problème augmentait avec le facteur de forme du motif ainsi qu'à cause de la présence de CR.

Ce modèle n'a pas été calibré sur des mesures réelles car dans l'état actuel il était trop instable pour donner un résultat suffisant.

Enfin on a montré qu'il était possible d'utiliser le modèle de reconstruction adapté aux tranchées pour reconstruire des motifs denses tant que l'espace entre deux tranchées consécutives n'était pas trop faible.

Chapitre 7

Analyses statistique

Dans cette partie, on va présenter quelques analyses statistiques permettant de déterminer l'impact de certains paramètres d'entrée que l'on a choisi d'utiliser pour les mesures/simulations, sur la variabilité du positionnement des descripteurs et l'estimation des grandeurs d'intérêt. Si jusque là on avait jamais simulé deux fois la même géométrie avec les mêmes conditions de prise d'image, ici au contraire on va simuler un grand nombre de fois le même motif, de manière à pouvoir étudier la variance au sens mathématique du terme. On rappelle que la variance et l'écart-type sont définis comme :

$$\text{variance} = \text{écart-type}^2 = \sigma_{\mathbf{g}}^2 = \frac{1}{N} \times \sum_{n=1}^N (\mathbf{g}_n - \bar{\mathbf{g}})^2, \text{ avec } \bar{\mathbf{g}} = \frac{1}{N} \times \sum_{n=1}^N \mathbf{g}_n \quad (7.1)$$

La variance n'est donc pas la comparaison d'un résultat en fonction de la valeur qu'il devrait avoir, mais la comparaison d'un résultat avec sa valeur moyenne. Si on veut tenir compte de l'éventuel écart (appelé biais) entre cette moyenne et la valeur considérée comme "vraie", on doit alors utiliser l'erreur quadratique moyenne, qui dans le cas où on fait N fois la même expérience, peut être reliée à la variance par :

$$\left[\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\text{Biais}^2 + \sigma_{\mathbf{g}}^2} = \sqrt{(\bar{\mathbf{g}} - \tilde{\mathbf{g}})^2 + \sigma_{\mathbf{g}}^2} \right] \geq \left[\text{écart-type} = \sigma_{\mathbf{g}} = \sqrt{\sigma_{\mathbf{g}}^2} \right] \quad (7.2)$$

7.1 Impact du paramètre d'ajustement de la spline, sur le placement des descripteurs

Dans tout le travail qui a été présenté, la première étape de traitement des profils issus des images SEM ou issus de la simulation, est la transformation en spline. Cette transformation permet à la fois de transformer les données brutes en un objet mathématique utilisable (et notamment dérivable) et à la fois comme filtre passe-bas, pour filtrer ce que l'on considère comme le bruit dû aux effets probabilistes des interactions électrons-matière.

Cependant au fur et à mesure de l'avancée du travail, nous nous sommes rendu compte de l'importance du paramètre d'ajustement de la spline (niveau de fit ou de smooth noté "S") sur tout le reste de la chaîne de traitement.

Dans ce travail, on a utilisé des B-splines (combinaison linéaire de fonctions polynomiales) de la bibliothèque python "scipy.interpolate" [64], qui se base sur les algorithmes de Dierckx [65]. Dans ce cadre, le paramètre d'ajustement "S" permet de contraindre la distance moyenne (au moindre carrés) entre la spline et les données d'entrée (Fig 7.1). Un S nul impliquera donc que la spline passera en chaque point des données d'entrée, la spline servira alors simplement de régularisation, c'est à dire à approximer les données par une suite de fonctions régulières afin de lisser les discontinuités. Ensuite plus on augmentera S, plus la spline pourra passer loin des

données, menant à un effet de lissage de plus en plus important. Dans notre cas, on pourra approximer cette augmentation de S , à une diminution progressive de la fréquence de coupure d'un filtre passe-bas sur notre signal.

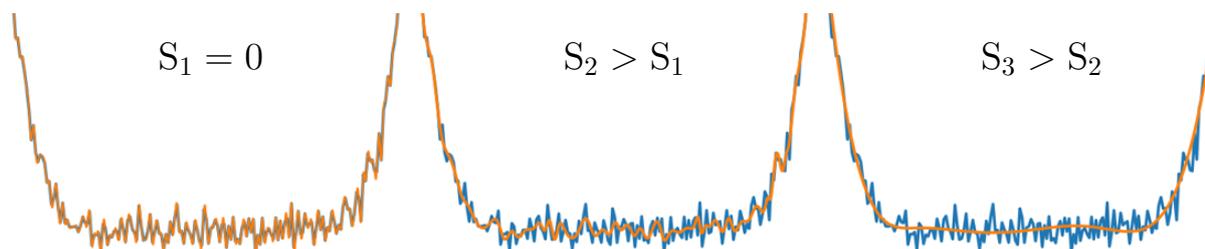
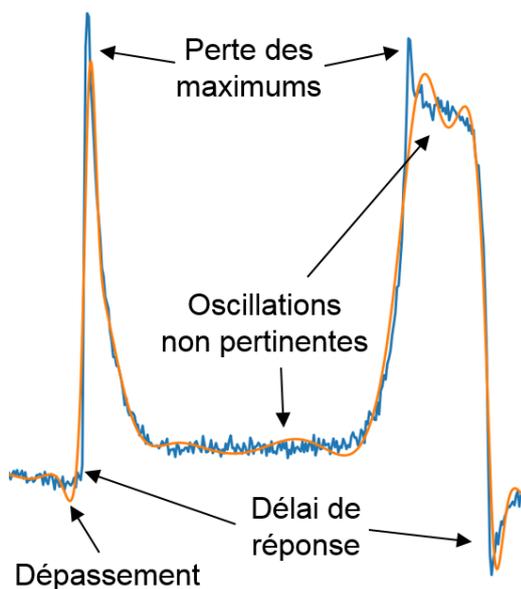


FIGURE 7.1 – Impact du paramètre d'ajustement "S" sur le lissage des données (en bleu) par la spline (en orange).

La difficulté de ce paramètre d'ajustement, est que l'on souhaite supprimer ce que l'on considère comme le bruit, tout en collant le plus possible aux données brutes, c'est à dire en altérant le moins possible ce que l'on considère comme l'information du signal. Le niveau d'ajustement optimal sera donc à la limite de disparition du bruit. Cependant il est dangereux de se placer à ce niveau optimal car cette limite va être différente pour chaque profil et même pour chaque sous-partie d'un même profil, en fonction des particularités du bruit (fréquence et amplitude par exemple) et du signal (différence de gradient notamment).



C'est pourquoi en pratique on va devoir s'éloigner suffisamment de ce point optimal pour s'assurer de supprimer le bruit. En effet si on ne supprime pas le bruit on va se retrouver avec des erreurs de positionnement des descripteurs voire même un échec de la stratégie de positionnement des descripteurs. Au contraire si on choisi un S "trop grand", l'altération du signal va dans un premier temps seulement détériorer la qualité du placement des descripteurs. On peut toutefois noter les effets dérangeants d'un S trop grand sur l'approximation du signal, comme décrit dans la figure 7.2.

FIGURE 7.2 – Visualisation des problèmes liés à un paramètre d'ajustement de la spline S trop grand.

On va maintenant quantifier l'impact sur le positionnement des descripteurs du niveau d'ajustement S de la spline, en allant d'un sur-ajustement, soit un S trop faible, c'est à dire une spline trop proche des données (couleur noire sur les graphiques), à un sous-ajustement soit un S trop grand, c'est à dire trop loin des données (couleur jaune sur les graphiques).

Sur la Fig 7.3, on peut voir le résultat qualitatif de l'augmentation du paramètre d'ajustement, sur un profil SEM unique. Pour les S les plus faibles (couleurs foncées) on peut voir que la spline est très bruitée, car le bruit n'est pas lissé par la spline. Pour les S les plus grand (couleurs claires) on peut voir l'apparition d'oscillations non pertinentes et donc une altération du signal utile. Entre ces deux extrêmes, on peut voir un intervalle de valeur de S sur lequel la spline approxime correctement les données (de orange à rouge foncé). Enfin il est très

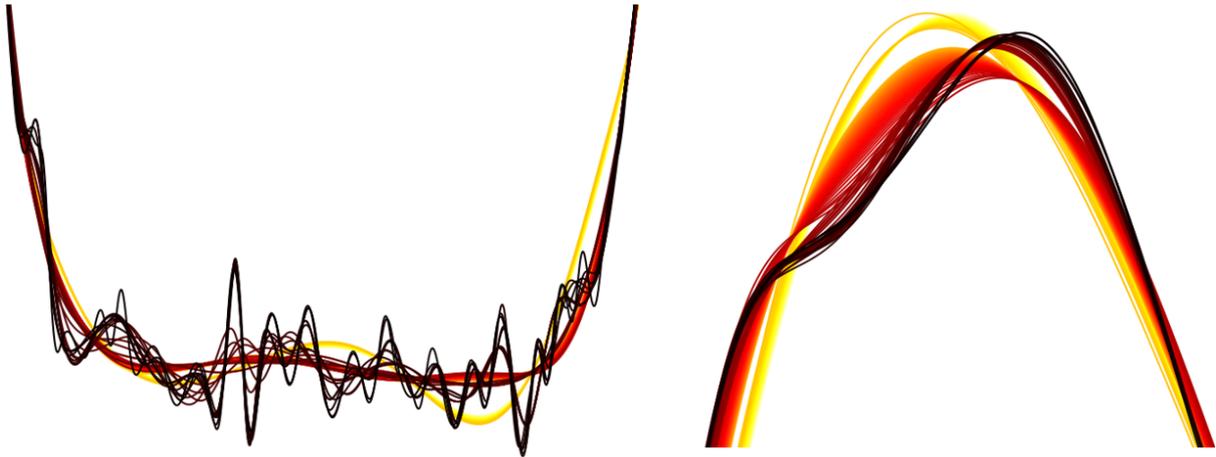


FIGURE 7.3 – Comparaison entre des splines dont le paramètre d’ajustement varie pour les mêmes données d’entrée.

important de noter que les transitions entre les régimes de splines sont très abruptes. En effet sur un certain intervalle on va avoir un déplacement plus ou moins régulier de la spline, puis au passage à une valeur de S à peine supérieur, la spline peut fortement changer de comportement. Cet effet de non linéarité peut amener à des discontinuités dans le placement des descripteurs.

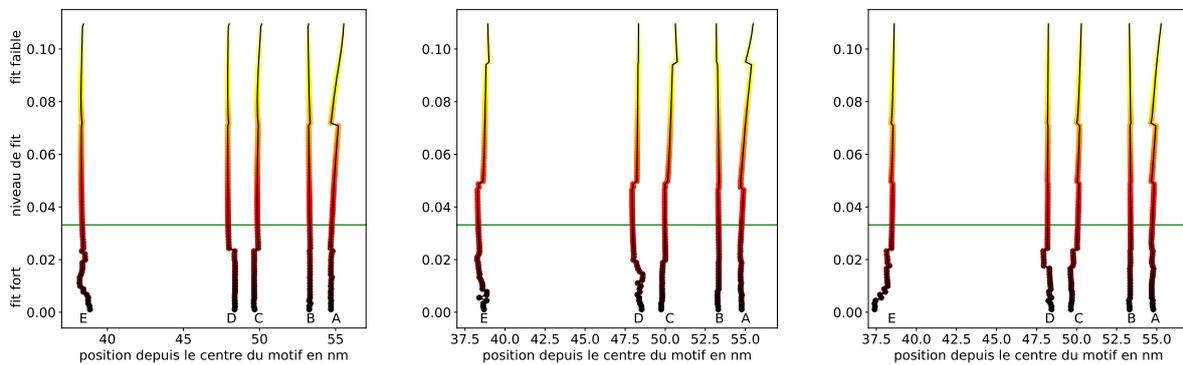


FIGURE 7.4 – Positionnement des descripteurs primaires en fonction du paramètre d’ajustement, sur trois simulations différentes du même Motif1 à $T=0^\circ$. La droite verte correspond à la valeur de S choisi dans la thèse pour ce type de motif.

La fig 7.4 montre le positionnement des descripteurs primaires en fonction de S . Tout d’abord on peut voir que lorsque S est faible (fit fort), on observe une instabilité du positionnement des descripteurs car le bruit n’est plus lissé et engendre donc des erreurs ”aléatoires”. On peut facilement observer que pour certaines valeurs de S , toute la spline change de positionnement et engendre donc un déplacement généralisé des descripteurs, qui peut être abrupt. A partir d’une valeur élevée de S , on observe des glissements progressifs des descripteurs, notamment pour le descripteurs A. Ceci est dû à un aplatissement régulier de la spline dans la zone du descripteur.

Les positionnements des descripteurs primaires dû au paramètre d’ajustement vont donc être fortement corrélés.

On peut noter que ce comportement n’est pas dû au fait que l’on observe des profils issus de simulations JMONSEL, en effet lorsque l’on fait la même expérience sur des profils

issus d'images SEM réelles, on obtient les mêmes cassures, dû à un repositionnement général de la spline pour certaines valeur de S difficilement anticipables. Cependant ces cassures correspondent au changement du nombre de points de contrôle de la spline. Pour mieux maîtriser ces problématiques il faudrait donc ne pas laisser la spline calculer elle-même ces points de contrôle, mais les forcer, cependant cela n'a pas été réalisé dans la thèse.

Il est donc très difficile de trouver une valeur optimale pour S pour le domaine de validité entier du modèle. Cela n'a pas été déployé dans cette thèse, mais on va voir dans les perspectives, une idée d'utilisation de plusieurs splines à différents paramètres d'ajustement, pour essayer de réduire les erreurs dues à l'approximation par une spline.

7.2 Impact du nombre d'électrons envoyés par pixel dans JMONSEL, sur le placement des descripteurs

Dans cette partie on va s'intéresser à l'influence du nombre d'électrons envoyés par pixel lors de la simulation JMONSEL sur la variabilité du positionnement des descripteurs primaires. En effet plus on envoie d'électrons par pixel plus cette variabilité diminue, mais au prix d'un temps de calcul plus long. De plus tous les descripteurs primaires ne vont pas avoir la même variabilité, ce qui va ensuite influencer sur la variabilité des descripteurs secondaires et donc de la variabilité de l'estimation des grandeurs d'intérêt.

On peut noter que le nombre d'électrons par pixel ($nbep$) des simulations peut être comparé avec le nombre de lignes sur lesquelles on somme pour obtenir le profil de niveau de gris sur des images SEM réelles. La valeur de $10000\text{ nbep} = 10\text{ knbep}$ qui a été utilisée pour les simulations engendre un bruit similaire à des profils moyennés sur 500 lignes d'images SEM, avec les conditions de prise d'image choisies (ici la taille de pixel et la vitesse de balayage du SEM sont importantes).

Jusqu'ici seule une simulation était réalisée pour chaque type de motif, en revanche pour étudier cette variabilité nous allons lancer un grand nombre de fois la même simulation de manière à voir comment les effets probabilistes de l'approche Monte-Carlo vont influencer sur le positionnement des descripteurs primaires. On pourra donc parler de variance et d'écart-type au sens mathématique strict.

Pour cela nous avons choisi de tester deux motifs différents à deux angles de tilt différents :

- Motif1 : $h = 100\text{ nm}$, $SWA = 88^\circ$, $CDt = 100\text{ nm}$, $CR = 0\text{ nm}$, $T = 0^\circ$ et $T = 12^\circ$
- Motif2 : $h = 100\text{ nm}$, $SWA = 88^\circ$, $CDt = 100\text{ nm}$, $CR = 20\text{ nm}$, $T = 0^\circ$ et $T = 12^\circ$

Pour chacune de ces quatre conditions de simulation, 1000 simulations ont été réalisées avec un nombre d'électrons par pixel ($nbep$) de 1000. Comme on l'a déjà expliqué, il est équivalent de simuler une ligne à $2X\text{ nbep}$ ou deux lignes à $X\text{ nbep}$, compte tenu de la nature indépendante du bruit, d'une simulation à l'autre. De ce fait, à partir de ces 1000 simulations à 1000 nbep on a pu, par moyennage, obtenir 500 simulations à 2000 nbep , 100 simulations à 10000 nbep etc...

Sur la Fig 7.5, on peut voir la variabilité du positionnement des splines (pour un niveau d'ajustement défini), sur le même motif simulé 100 fois à 1 knbep (a) et à 10 knbep (b). On observe facilement la réduction de la variabilité lorsque l'on considère un plus grand nombre d'interactions électron-matière par unité de longueur, qui va faire tendre la réponse électronique vers l'effet moyen des interactions possibles.

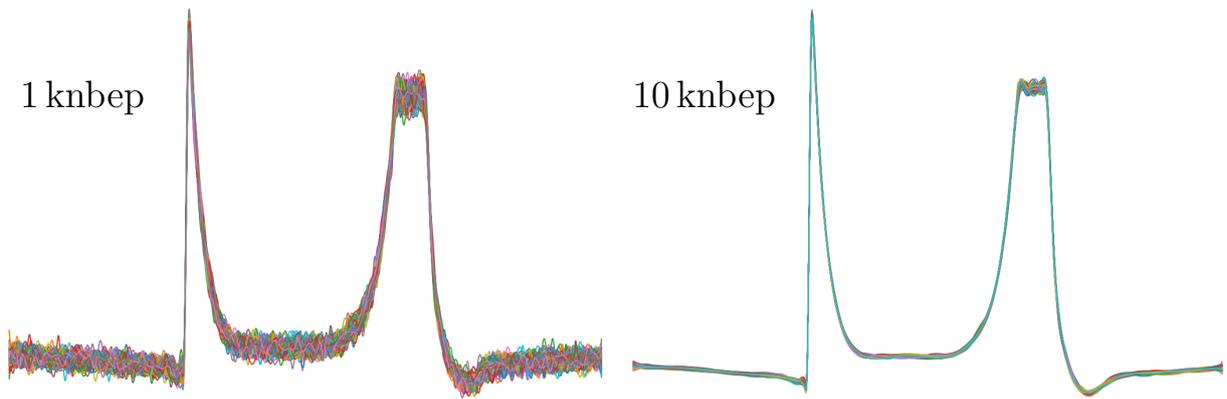


FIGURE 7.5 – Comparaison du bruit entre les splines d’un même motif simulé 100 fois à 1 knbep et à 10 knbep par JMONSEL (Motif2 à 12°).

Toutes ces simulations ont été analysées avec la méthode présentée, c’est à dire par approximation par une spline puis placement de points caractéristiques liés à la dérivée de la spline. La fig 7.6 montre la répartition du positionnement de chaque descripteur, pour les deux nbep simulés. On peut voir d’une part la forte différence de variabilité entre les différents descripteurs pour un même nbep et d’autre part, en comparant les deux graphiques, la réduction de la dispersion du placement des descripteurs, lorsque la simulation moyenne sur un plus grand nombre d’interactions électrons-matière.

On peut expliquer, en première approximation, la différence de variabilité des descripteurs par le gradient local (ou moins local pour C (effet PID)) de la spline à l’emplacement du descripteur.

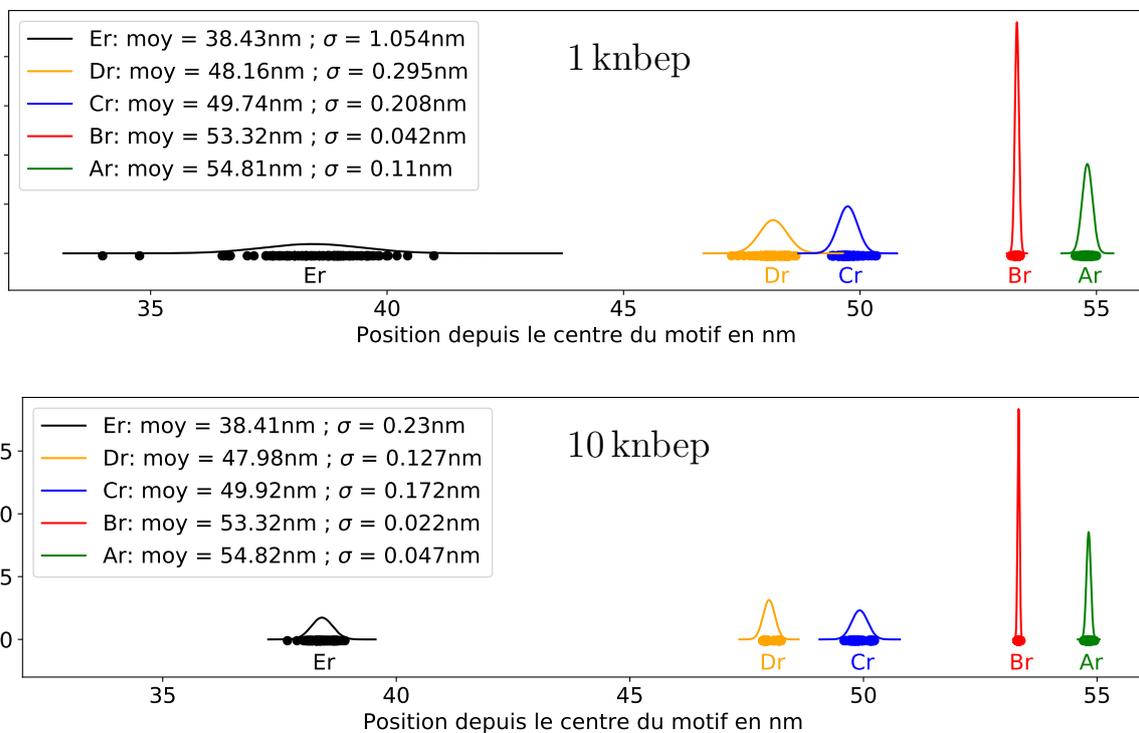


FIGURE 7.6 – Comparaison de la répartition du positionnement des descripteurs pour 1 knbep et 10 knbep sur le Motif1 à $T=0^\circ$.

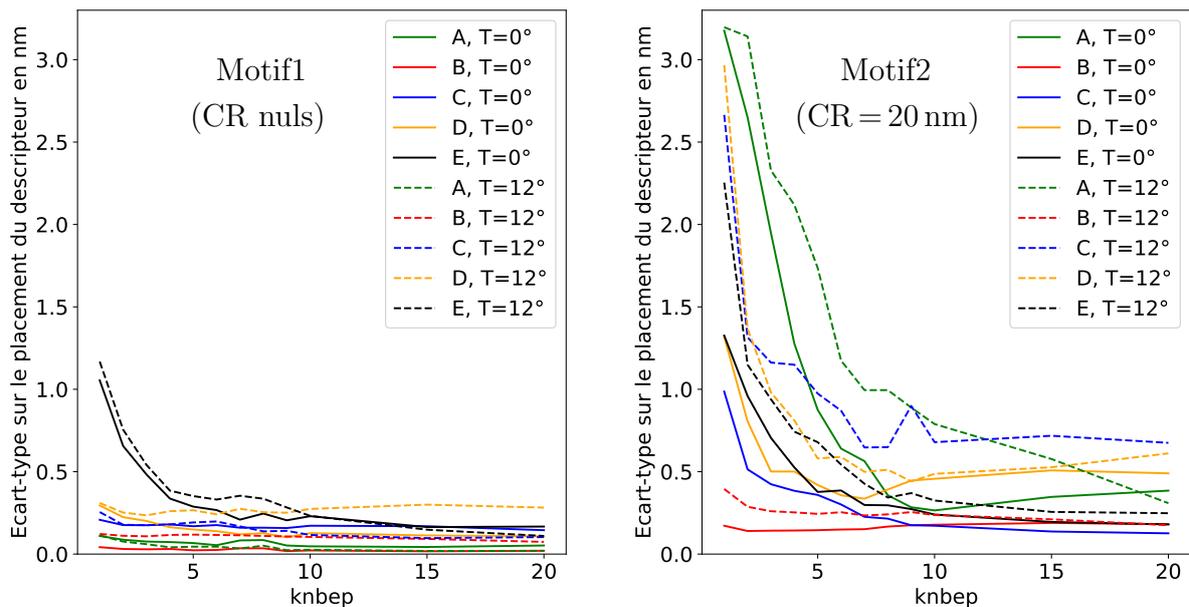


FIGURE 7.7 – Comparaison de la variabilité de positionnement des descripteurs en fonction du nbep, pour le Motif1 et le Motif2.

La Fig 7.7, permet de comparer la variabilité de positionnement des descripteurs en fonction du nombre d'électrons envoyés par pixel (nbep) de 1 knbep à 20 knbep, pour le Motif1 (CR nuls) et le Motif2 (CR = 20 nm).

On peut voir que plus le nbep est grand, plus l'erreur sur le positionnement des descripteurs va être faible. On rappelle que toutes les simulations réalisées pour les parties de création du modèle ont été faites à 10 knbep, ce qui semble être un choix judicieux, car au delà la réduction de la dispersion des descripteurs réduit fortement.

La présence de CR augmente très fortement l'erreur de positionnement des descripteurs même à fort nbep. Cette observation est cohérente, avec l'augmentation de la complexité du modèle, lorsque l'on considère des CR non nuls. De plus on peut voir que les descripteurs ne sont pas tous affectés de la même manière à la fois par le nbep et par le type de motif. On peut en première approximation retrouver l'effet expliqué dans la partie 4.4, de dépendance de la variabilité des descripteurs au gradient local du profil SEM à l'endroit de ces descripteurs. Par exemple on voit que le positionnement du descripteur B (en rouge), qui se trouve à un emplacement où le gradient est fort, est plus précis que les autres descripteurs, ou encore on peut voir que la variabilité du descripteur A (en vert) explose lorsque le CR passe de 0 nm à 20 nm (réduction du gradient en présence de CR).

7.3 Étude de la propagation de l'incertitude de placement des descripteurs sur l'estimation des grandeurs d'intérêts

7.3.1 Calcul empirique des erreurs, sur des résultats de simulation

Tout d'abord, nous allons calculer la variabilité des grandeurs d'intérêt, estimées à partir de la répétition de 100 simulations JMONSEL du même motif, à 10knbep. Nous avons à nouveau travaillé sur les deux motifs présentés précédemment ($h = 100$ nm, $SWA = 88^\circ$ et $CDt = 100$ nm), le Motif1 n'ayant pas de CR et le Motif2 des $CR = 20$ nm, à $T = 0^\circ$ et $T = 12^\circ$.

Grandeur	M1 CR0 T0	M1 CR0 T12	M2 CR20 T0	M2 CR20 T12
CRb (moy)	0.38 nm	0.66 nm	19.30 nm	21.01 nm
CRb (σ)	0.09 nm	0.17 nm	0.41 nm	1.32 nm
CRt (moy)	4.22 nm	4.04 nm	19.22 nm	24.14 nm
CRt (σ)	0.44 nm	0.73 nm	1.69 nm	3.96 nm
EW (moy)	4.08 nm	25.20 nm	5.22 nm	25.64 nm
EW (σ)	0.13 nm	0.31 nm	0.69 nm	0.87 nm
Hauteur (moy)	99.38 nm		96.05 nm	
Hauteur (σ)	1.54 nm		5.32 nm	
SWA (moy)	87.65°		86.86°	
SWA (σ)	0.09°		0.56°	
CDt (moy)	100.08 nm	99.48 nm	97.64 nm	97.15 nm
CDt (σ)	0.24 nm	0.26 nm	0.56 nm	0.68 nm

TABLE 7.1 – Incertitude sur l’estimation des grandeurs d’intérêt, d’un même motif simulés 100 fois. Motif1 à gauche et Motif2 à droite.

On peut observer dans le Tableau 7.1 la propagation des erreurs des descripteurs primaires (Fig 7.7 à 10knbep) vers les grandeurs d’intérêt. On remarque encore une fois que la présence de CR augmente la variabilité sur l’estimation des grandeurs d’intérêt.

Même si ce tableau seul ne prouve pas le résultat suivant, on peut noter que l’estimation du CDt est très sensible aux erreurs sur l’estimation du SWA ; En effet, on voit que l’erreur de sous estimation du SWA faite sur le Motif2 (86.86° au lieu de 88°) entraîne une sous-estimation non-négligeable du CDt, qui se retrouve éloigné de la valeur vrai de plus de 3σ .

Incertaince sur l’estimation des EW

La variabilité sur la mesure d’EW peut avoir plusieurs origines :

- Origine physique : Les mesures sont prises sur une même ligne à différents endroits sur sa longueur (60 μm) pour avoir une moyenne de la topographie de la ligne, de manière à ne pas fausser le résultat par un effet local non représentatif. La variabilité peut donc refléter une variabilité réelle de la topographie de la ligne.
- L’appareil de mesure : Dans un microscope électronique le faisceau électronique n’est pas ponctuel. Les électrons ont donc une probabilité d’arriver sur l’échantillon que l’on peut approximer par une gaussienne 2D (admis dans la littérature). La taille de la surface circulaire de probabilité de présence est appelée le spot size et son ordre de grandeur est de 1 à 2 nm.
- La technique de traitement du signal pour débruiter et lisser le signal, pouvant entraîner des variabilités sur environ 2 pixels (soit 1nm dans notre cas).
- La variabilité de l’angle de tilt de la machine entraîne directement une variabilité sur l’EW mesuré, via la formule suivante :

$$EW = \frac{h}{\tan(\text{SWA})} + h \tan(T) \quad (7.3)$$

En posant σ_T^2 la variance sur le tilt donné par le constructeur, la proportion de variabilité due à la variabilité du tilt (notée $\sigma_{EW_T}^2$) serait donc donnée par :

$$\begin{aligned} \sigma_{EW_T}^2 &= \left(\frac{\partial EW}{\partial T} \right)^2 \sigma_T^2 \\ \sigma_{EW_T}^2 &= \frac{h^2 \sigma_T^2}{\cos^4(T)} \end{aligned} \quad (7.4)$$

Si on pose $\sigma_{EW_{\text{mesure} + \text{traitement}}}^2$ la variabilité due à "tout sauf le tilt" de manière à ce que $\sigma_{EW_{\text{mesure} + \text{traitement}}}^2$ et σ_T^2 soient décollées, alors par construction la variabilité totale est la somme quadratique de ces incertitudes. On pourra donc en déduire si la variabilité de l'EW est principalement due à la variabilité du tilt ou à la variabilité entraînée par la mesure, le bruit, le traitement etc, en posant :

$$\sigma_{EW}^2 = \sigma_{EW_{\text{mesure} + \text{traitement}}}^2 + \sigma_{EW_T}^2 \quad (7.5)$$

Dans [66], il est proposé une valeur de σ_T de 0.075° , ce qui semble être une valeur pertinente pour un SEM calibré. Avec cette valeur on obtient alors une erreur sur l'EW mesuré qui dépend principalement de la hauteur :

$$\sigma_{EW} \approx 1.37 \times 10^{-3} \times h \quad (7.6)$$

L'erreur due à l'incertitude sur la valeur de l'angle de tilt pourrait donc être non-négligeable pour la reconstruction de motif haut. Cette augmentation de l'erreur en fonction de la hauteur du motif pourrait expliquer l'augmentation de l'erreur sur les motifs plus grands, observée dans le Tableau 5.2. Pour un motif de 100 nm on obtient un écart-type de 0.14 nm, sur l'erreur moyenne de 0.84 nm observé soit 15% de l'erreur due à la variabilité du tilt.

7.3.2 Calcul théorique des erreurs

On aurait pu calculer de manière théorique la propagation des incertitudes depuis l'incertitude des descripteurs primaires jusqu'aux grandeurs d'intérêt, cependant nous nous sommes rendu compte que l'hypothèse d'indépendance des descripteurs était rarement acceptable, c'est pourquoi, pour éviter une section de calculs très lourde, nous avons préféré présenter la variabilité expérimentale.

En revanche pour le calcul de la hauteur et du SWA, on utilise la valeur de l'EW provenant de deux images différentes, c'est pourquoi l'hypothèse d'indépendance est acceptable ; Les calculs théoriques de l'incertitude de la hauteur et du SWA en fonction de l'erreur sur l'estimation de EW, sont donc présentés.

Incertitude sur l'estimation de la hauteur

On rappelle que la formule de la hauteur est donnée par :

$$h = \frac{EW2 - EW1}{\tan(T2) - \tan(T1)} \quad (7.7)$$

On veut alors connaître la valeur de l'incertitude sur h en fonction des incertitudes sur les mesures d'EW.

On a alors :

$$\sigma_h^2 = \left(\frac{\partial h}{\partial EW1} \right)^2 \sigma_{EW1}^2 + \left(\frac{\partial h}{\partial EW2} \right)^2 \sigma_{EW2}^2 \quad (7.8)$$

Or,

$$\left(\frac{\partial h}{\partial EW1} \right)^2 = \left(\frac{\partial h}{\partial EW2} \right)^2 = \frac{1}{(\tan(T2) - \tan(T1))^2}$$

d'où,

$$\sigma_h^2 = \frac{\sigma_{EW1}^2 + \sigma_{EW2}^2}{(\tan(T2) - \tan(T1))^2} \quad (7.9)$$

On observe donc que l'incertitude tend vers l'infini lorsque T1 tend vers T2. Pour obtenir la meilleure précision on essayera donc de maximiser l'écart entre les deux angles de tilt. Dans notre cas le tilt maximum est de 12°, on devrait donc l'utiliser en couple avec un tilt de 0° pour minimiser l'incertitude, à condition que la variance en 0° de l'EW, $\sigma_{EW(T=0)}^2$ ne soit pas trop supérieure aux autres.

Incertainude sur l'estimation du SWA

On rappelle que la formule du SWA est donnée par :

$$SWA = \arctan \left(\frac{EW2 - EW1}{EW1 \times \tan(T2) - EW2 \times \tan(T1)} \right) \quad (7.10)$$

On veut maintenant calculer la valeur de l'incertitude sur SWA en fonction des incertitudes sur les mesures d'EW.

On a alors :

$$\sigma_{SWA}^2 = \left(\frac{\partial SWA}{\partial EW1} \right)^2 \sigma_{EW1}^2 + \left(\frac{\partial SWA}{\partial EW2} \right)^2 \sigma_{EW2}^2 \quad (7.11)$$

Pour résoudre l'équation on pose :

$$SWA = g \circ f \quad \text{avec} \quad g(x) = \arctan(x) \quad \text{et} \quad f(EW1, EW2) = \frac{EW2 - EW1}{EW1 \times \tan(T2) - EW2 \times \tan(T1)}$$

$$\text{On pose, } f = \frac{u}{v}, \text{ avec, } \frac{\partial u}{\partial EW1} = -1 \text{ et, } \frac{\partial v}{\partial EW1} = \tan(T2)$$

$$\text{d'où, } \frac{\partial f}{\partial EW1} = \frac{-(EW1 \times \tan(T2) - EW2 \times \tan(T1)) - (EW2 - EW1) \times \tan(T2)}{(EW1 \times \tan(T2) - EW2 \times \tan(T1))^2}$$

$$\text{On rappelle que, } \frac{\partial \arctan(f)}{\partial f} = \frac{1}{1 + f^2}$$

On obtient donc,

$$\frac{\partial SWA}{\partial EW1} = \frac{EW2 \times (\tan(T1) - \tan(T2))}{(EW1 \times \tan(T2) - EW2 \times \tan(T1))^2} \times \frac{1}{1 + \left(\frac{EW2 - EW1}{EW1 \times \tan(T2) - EW2 \times \tan(T1)} \right)^2}$$

d'où,

$$\frac{\partial SWA}{\partial EW1} = \frac{EW2 \times (\tan(T1) - \tan(T2))}{(EW1 \times \tan(T2) - EW2 \times \tan(T1))^2 + (EW2 - EW1)^2}$$

De manière analogue on obtient,

$$\frac{\partial SWA}{\partial EW2} = \frac{EW1 \times (\tan(T2) - \tan(T1))}{(EW1 \times \tan(T2) - EW2 \times \tan(T1))^2 + (EW2 - EW1)^2}$$

Finalement,

$$\sigma_{SWA}^2 = \frac{(\tan(T2) - \tan(T1))^2 \times (EW1^2 \times \sigma_{EW2}^2 + EW2^2 \times \sigma_{EW1}^2)}{((EW1 \times \tan(T2) - EW2 \times \tan(T1))^2 + (EW2 - EW1)^2)^2} \quad (7.12)$$

7.3.3 Vérification des calculs théoriques sur les résultats empiriques

On peut maintenant vérifier si ces équations théorique de propagation de l'erreur de mesure des EW vers la hauteur et le SWA sont vérifiées dans le cas des mesures empiriques que l'on a réalisé précédemment et résumé dans le Tableau 7.1.

Pour la propagation de l'erreur des EW sur la hauteur du Motif1 on a :

$$\sigma_h = \sqrt{\frac{0.13^2 + 0.31^2}{(\tan(12))^2}} = 1.58 \text{ nm (empiriquement : 1.54 nm)} \quad (7.13)$$

Pour la propagation de l'erreur des EW sur la hauteur du Motif2 on a :

$$\sigma_h = \sqrt{\frac{0.69^2 + 0.87^2}{(\tan(12))^2}} = 5.22 \text{ nm (empiriquement : 5.32 nm)} \quad (7.14)$$

Pour la propagation de l'erreur des EW sur le SWA du Motif1 on a :

$$\sigma_{\text{SWA}} = \frac{180}{\pi} \sqrt{\frac{(\tan(12))^2 \times (4.08^2 \times 0.31^2 + 25.2^2 \times 0.13^2)}{((4.08 \times \tan(12))^2 + (25.2 - 4.08)^2)^2}} = 0.096^\circ \text{ (empiriquement : 0.09}^\circ\text{)} \quad (7.15)$$

Pour la propagation de l'erreur des EW sur le SWA du Motif2 on a :

$$\sigma_{\text{SWA}} = \frac{180}{\pi} \sqrt{\frac{(\tan(12))^2 \times (5.22^2 \times 0.87^2 + 25.64^2 \times 0.69^2)}{((5.22 \times \tan(12))^2 + (25.64 - 5.22)^2)^2}} = 0.53^\circ \text{ (empiriquement : 0.56}^\circ\text{)} \quad (7.16)$$

On obtient donc des résultats cohérents avec les valeurs d'erreur trouvées empiriquement sur la hauteur et le SWA. On est donc rassuré sur l'hypothèse d'indépendance faite entre les valeurs d'EW provenant de deux images différentes.

Chapitre 8

Conclusions et perspectives

8.1 Conclusions

Le travail de thèse qui a été présenté dans ce manuscrit a permis de créer une méthodologie pour la mesure topographique de motifs typiques de la microélectronique, à partir d'images de microscopie électronique à balayage à faisceau inclinable. Pour cela une paramétrisation d'un motif typique de la microélectronique (ligne ou tranchée) a été proposée : l'approximation "8 paramètres". Cette approximation permet une définition des différentes grandeurs d'intérêts, permettant par la suite d'avoir une base commune sur laquelle comparer quantitativement les résultats provenant des différentes techniques de reconstruction utilisées, par AFM, FIB/STEM et SEM tilté.

La reconstruction 3D s'appuie sur une première étape de transformation des images SEM en profils de niveau de gris, correspondant au rendement électronique. Grâce à la fonction "ruban" qui permet de récupérer la forme de la rugosité des motifs observés, les profils ne sont pas altérés par la rugosité de ligne et de ce fait la reconstruction est grandement améliorée. Ces profils sont ensuite transformés en spline de manière à obtenir un objet mathématique dérivable et également à réduire le bruit de mesure, caractéristique de la mesure SEM. Enfin des descripteurs primaires, correspondant à des valeurs particulières de la dérivée de la spline sont positionnés de manière à pouvoir analyser ces profils.

A partir de cette méthodologie, un modèle a été créé et calibré pour la reconstruction de motifs de type "ligne isolée" en silicium gravé à partir d'images SEM dont l'énergie des électrons primaires est de 800 eV. Ce modèle, qui lie les descripteurs aux grandeurs d'intérêts recherchées, a été utilisé sur des images SEM provenant de motifs réels et a permis d'obtenir une estimation de la hauteur avec un écart-type de 5.6 nm environ et une estimation du SWA avec un écart-type de 0.6° environ. La méthode présentée permet notamment une forte amélioration de la précision de la reconstruction de motifs présentant des CR non-nuls par rapport à l'état de l'art. Ces résultats sont encourageants, mais ne sont pas forcément suffisants pour tout type d'application. Cependant un grand nombre d'améliorations à ce travail sont possibles comme nous allons le voir dans les perspectives.

A partir de cette même méthodologie, un deuxième modèle linéaire a été créé pour la reconstruction de motifs de type "tranchée isolée" en silicium gravé à partir d'images SEM dont l'énergie des électrons primaires est de 800 eV. On a prouvé que dans un certain domaine de validité, ce modèle pouvait également permettre la reconstruction de motifs de type "dense". Ce modèle cependant n'a pas été calibré car il présentait trop d'instabilité. On a démontré que ces instabilités étaient en partie dues à la réduction de l'information électronique lorsque le facteur de forme des motifs était grand. Des améliorations doivent être apportées à la méthode

pour rendre la reconstruction sur les motifs de type "tranchée" stable et précise.

Enfin des analyses statistiques ont permis de comprendre l'impact de certains paramètres d'entrée que l'on a fixé dans la thèse sur la variabilité du positionnement des descripteurs primaires. On a notamment vu que le paramètre d'ajustement de la spline avait des conséquences non négligeables sur le positionnement des descripteurs. Les caractéristiques topographiques des motifs observés et notamment les CR, ont également une influence importante sur la variabilité du positionnement des descripteurs et donc sur le résultat topographique.

8.2 Perspectives

8.2.1 Propositions d'amélioration

Généralisation à la reconstruction de motifs courbes et 2D

Dans ce travail de thèse, nous avons, par souci de simplicité, présenté la reconstruction de motifs dit "Manhattan", c'est à dire horizontaux ou verticaux. Cependant la méthode présentée pourrait être étendue aux motifs "non-Manhattan".

Pour prendre en compte les formes courbes, il serait possible d'ajouter un facteur correctif sur la mesure de l'EW tel que :

$$EW = \frac{h}{\tan(SWA)} + h \times \tan T \times \sin(\theta) \quad (8.1)$$

où θ représente l'angle entre la direction de tilt du faisceau électronique et la tangente au flanc du motif (Fig 8.1). Pour les valeurs de θ supérieures à 80° (c'est à dire une inclinaison inférieure à 10° de la verticale), $\sin(\theta) \approx 1$, donc la méthode peut reconstruire des motifs légèrement incurvés.

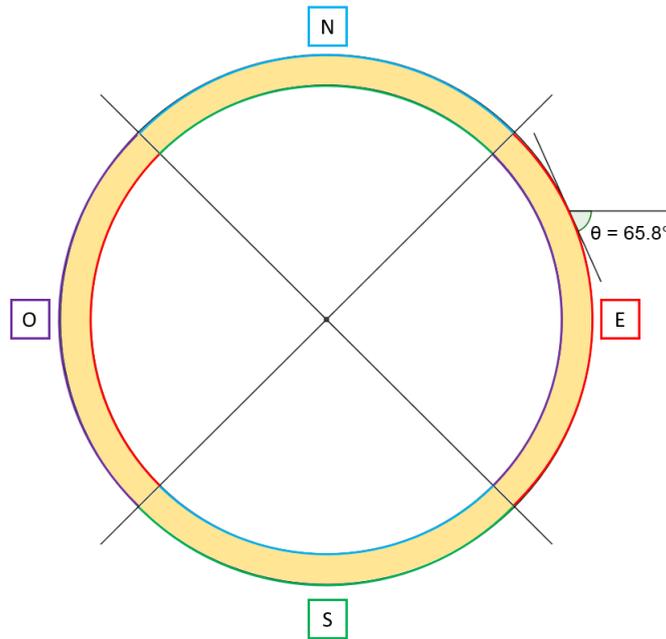


FIGURE 8.1 – Schéma d'un motif circulaire (typique pour la photonique) pour expliquer la reconstruction azimutale ici par quatre azimuts. Les flancs du motifs sont représentés par une couleur qui correspond à la couleur de l'azimut (N, E, S, O) qui en effectuera la reconstruction.

Cependant cette solution ne peut pas permettre de résoudre toutes les géométries 2D ; en effet, plus θ va diminuer, plus l'information apportée par le tilt va réduire, jusqu'à être nulle en $\theta = 0$. Géométriquement parlant, quand la direction du tilt est tangente au flanc observé, alors l'augmentation du tilt ne change pas la taille apparente (EW) du flanc.

Pour pouvoir prendre en compte toutes les géométries 2D possibles, il faudrait donc utiliser différents azimuts comme proposé dans la Fig 8.1, chacun permettant la reconstruction des flancs proches de la normalité par rapport à leur direction. Un minimum de quatre azimuts semble un choix judicieux, car chaque azimut pourrait alors reconstruire les flancs à $\pm 45^\circ$ de la normalité. L'information apportée par le tilt serait alors de $\sin(45) \approx 0.7$ soit 70% de l'infor-

mation optimale à 90° .

Dans le cas de l'utilisation de huit azimuts (N, NE, E, SE, S, SO, O, NO) chaque azimut reconstruirait les flanc orthogonaux ainsi que les flancs jusqu'à $\pm 22.5^\circ$ de cette direction. On pourrait également faire des recouvrements jusqu'à $\pm 45^\circ$ pour doubler les résultats de reconstruction et faire une moyenne.

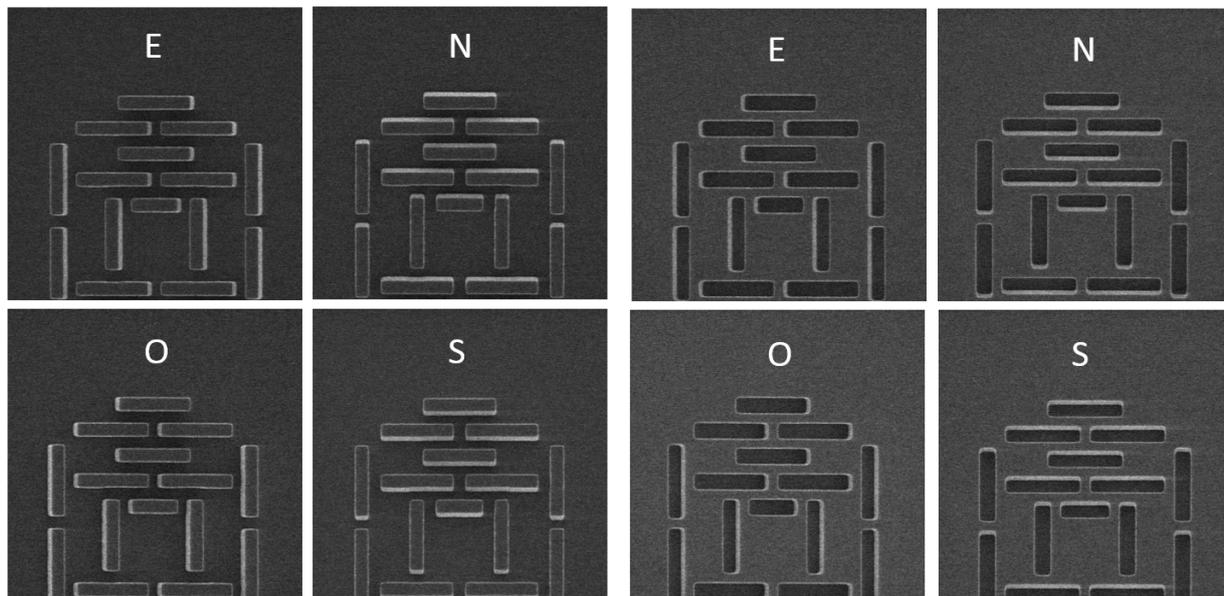


FIGURE 8.2 – Images azimutales sur des motifs de hauteurs positives (à gauche) et négative (à droite).

Enfin, dans ce travail on a utilisé la fonction "ruban" pour supprimer la rugosité de ligne, mais une version plus avancée de cette fonction pourrait être mise en œuvre pour redresser une forme quelconque, afin de la rendre équivalente à une ligne. Ceci permettrait par exemple de reconstruire tout type de motif incurvé ou même circulaire, car il serait alors possible de sommer le niveau de gris sur la longueur du périmètre du motif.

Amélioration du traitement par les splines

On a vu dans la partie 6.1 sur l'extension aux profils tranchées, qu'il était difficile de trouver le paramètre d'ajustement de la spline permettant une analyse optimale d'un grand nombre de profils SEM différents. Pour améliorer les résultats et pour repérer les valeurs aberrantes, on pourrait donc ajuster un grand nombre de splines sur chacun des profils avec des paramètres d'ajustement différents, comme ce que l'on a fait dans la partie 7.1.

On obtiendrait donc une fonction du positionnement de chaque descripteur primaire en fonction du paramètre d'ajustement de la spline (Fig 7.4). On pourrait alors réfléchir à un traitement de cette information pour maximiser l'information provenant de chaque descripteur. On pourrait, par exemple, imaginer qu'un ajustement "au plus proche des données d'entrée" serait meilleur pour déterminer la distance BD (car gradient fort donc peu affecté par le bruit), alors qu'un ajustement plus "mou" serait meilleur pour déterminer les points A et E (car gradient plus faible donc affecté par le bruit). Enfin cela permettrait de faire de la statistique "artificielle" à partir d'une seule image SEM.

Quand on approxime un profil par une spline, on réalise de ce fait un lissage qui va réduire le bruit de mesure. Cette approximation fait toujours disparaître une portion de ce que l'on

considère comme du bruit mais également une portion que l'on considère comme du signal. Cette technique pourrait donc permettre de s'approcher du niveau maximum d'information de l'image, c'est à dire à réduire l'erreur due au traitement des images et de tendre vers l'erreur minimale possible considérant la précision de la représentation des motifs dans les images.

Au delà ces idées, qui sont un peu du bricolage, un travail approfondi sur la compréhension du comportement des splines pourrait être bénéfique à la méthode de reconstruction présentée, notamment par la gestion fine des points de contrôle de la spline.

Création de modèles divers

Grâce à cette méthodologie on pourrait créer des modèles permettant la reconstruction de structures variées. On a vu que, dans l'état actuel, le modèle n'était pas capable de reconstruire des motifs dont le CDT est inférieur à 20-25 nm. Cependant on pourrait utiliser de nouveaux descripteurs primaires, spécifiquement conçu pour l'analyse de motifs dont le CDT est suffisamment petit pour que les effets de deux flancs se superposent (pour une certaine accélération des électrons).

On pourrait également vouloir faire de la reconstruction sur des motifs qui ne sont pas en silicium gravé. Pour les étapes de lithographie, on aimerait notamment faire des reconstructions sur des motifs en résine. Cependant l'utilisation du SEM pour imager des motifs en résine entraîne deux difficultés : les phénomènes d'accumulation de charges (charging) et la réduction du volume (shrinkage) de la résine due au bombardement électronique.

Pour réduire au maximum le problème de réduction de volume, il faut utiliser une accélération des électrons la plus faible possible. En général les énergies 500 eV et 300 eV sont utilisées pour les mesures SEM sur résine. Cependant même avec des accélérations inférieures, l'effet est seulement réduit. Pour supprimer l'effet de réduction de volume due à la prise d'image sur la topographie du motif de résine il faudrait donc étudier cet effet (dépendant de V_{acc} et du type de résine) pour utiliser le résultat "après réduction" obtenu avec notre technique de manière à retrouver la topographie "avant réduction".

Pour la partie sur les phénomènes d'accumulation de charges, si cet effet est suffisamment régulier, il pourrait alors être intégré par le modèle (cela peut même être une source d'information supplémentaire).

Obtenir les informations tri-dimensionnelles par une autre méthode

Pour pouvoir faire la reconstruction tri-dimensionnelle d'un motif sans connaître ni sa hauteur ni son SWA, il est obligatoire d'avoir deux informations. Dans ce travail, c'est grâce à la différence géométrique d'un même objet observé depuis deux positions différentes que l'on a pu tirer la double information. Cependant on pourrait également étudier la différence de rendement électronique entre deux images prises à des accélérations électroniques différentes. Cette technique a d'ailleurs été récemment étudiée dans la publication [67]. Ce n'est pas encore une technique de précision et la méthode doit être testée dans des cas plus variés que ceux présentés, mais les premiers résultats sont encourageants.

8.2.2 Applications spécifiques

Dans ce manuscrit on a présenté la méthode d'un point de vue purement métrologique. Cependant de multiples applications spécifiques de ce travail sont envisageables dans le milieu des micro/nanotechnologies.

Étude de la topographie pour la réalisation de micro-lentilles

Tout d'abord cette méthode pourrait permettre de créer des cartes de hauteur/SWA de motif, permettant de repérer un effet de variabilité au niveau d'un wafer entier. Cela pourrait permettre d'améliorer le contrôle des pentes pour les procédés de gravure par exemple. Pour les applications micro-lentilles, la forme du plot de résine avant l'étape de fluage et notamment la valeur des SWA est très importante pour obtenir la forme voulue. Une méthode rapide et massive de mesure des hauteurs et SWA serait donc intéressante pour assurer l'uniformité sur un capteur ou un wafer, ou encore assurer la symétrie entre le bord gauche et le bord droit des plots. Cette application est envisageable à condition d'avoir un modèle adapté à la résine et une mesure SEM qui n'endommage pas les plots de résine.

Modèles de correction des masques de photolithographie

En lithographie optique, les techniques d'amélioration de résolution (Resolution Enhancement Techniques) sont utilisées pour optimiser les motifs présents sur les masques afin de corriger les effets de proximité optique (Optical Proximity Correction). La mise en œuvre de ces corrections se base notamment sur des modèles optiques permettant de prédire les contours de résine sur le wafer en fonction des motifs présents sur le masque. La calibration de ces modèles nécessite une grande quantité de mesures SEM sur des motifs variés. La qualité des informations de CD extraites de ces mesures est critique pour obtenir un modèle OPC suffisamment précis. Pour autant qu'elle puisse être appliquée sur des motifs de résine, la méthodologie développée au cours de ce travail de thèse pourrait permettre de rendre plus robustes des données de calibration des modèles en assurant une mesure des motifs à hauteur constante quelle que soit la forme du motif, son environnement, sa pente...

De plus, il existe maintenant des modèles OPC avancés prenant en compte la topographie 3D des motifs et pas seulement un contour "topview" comme c'était historiquement le cas. Ce type de modèle nécessite pour sa calibration des informations de CD à différentes hauteurs. Les problématiques classiques de la métrologie 3D développées dans le premier chapitre (lente, destructive, ...) sont un véritable frein à la mise en œuvre de ces modèles étant donné le grand nombre de mesures nécessaires pour les calibrer. Notre méthode peut répondre de façon assez directe à ces difficultés en permettant l'extraction des CD à toutes les hauteurs des motifs à partir d'images SEM prises à différents tilts.

Enfin, l'étape d'OPC est également utilisée pour compenser en avance de phase les distorsions qui auront lieu pendant l'étape de gravure permettant de transférer les motifs de résine dans le matériau sous-jacent. Pour cela, il est nécessaire de quantifier ces distorsions. La façon classique de le faire consiste à mesurer par SEM les motifs d'intérêt en résine (après l'étape de lithographie), puis les motifs dans le matériau après l'étape de gravure. La différence de CD nous permet d'évaluer le biais induit par la gravure et cela sur un grand nombre de motifs différents afin de couvrir l'espace de design de la technologie considérée. L'avantage de la méthode développée est qu'elle peut permettre de s'assurer que l'extraction de CD est faite à hauteur constante sur tous les motifs, ce qui n'est pas évident dans le cas de mesures SEM classiques en vue du dessus avec des pentes variables suivant les motifs.

8.3 Remerciements

Je remercie toutes les personnes qui, avec bienveillance, m'ont expliqué et convaincu que j'avais tort.

Appendices

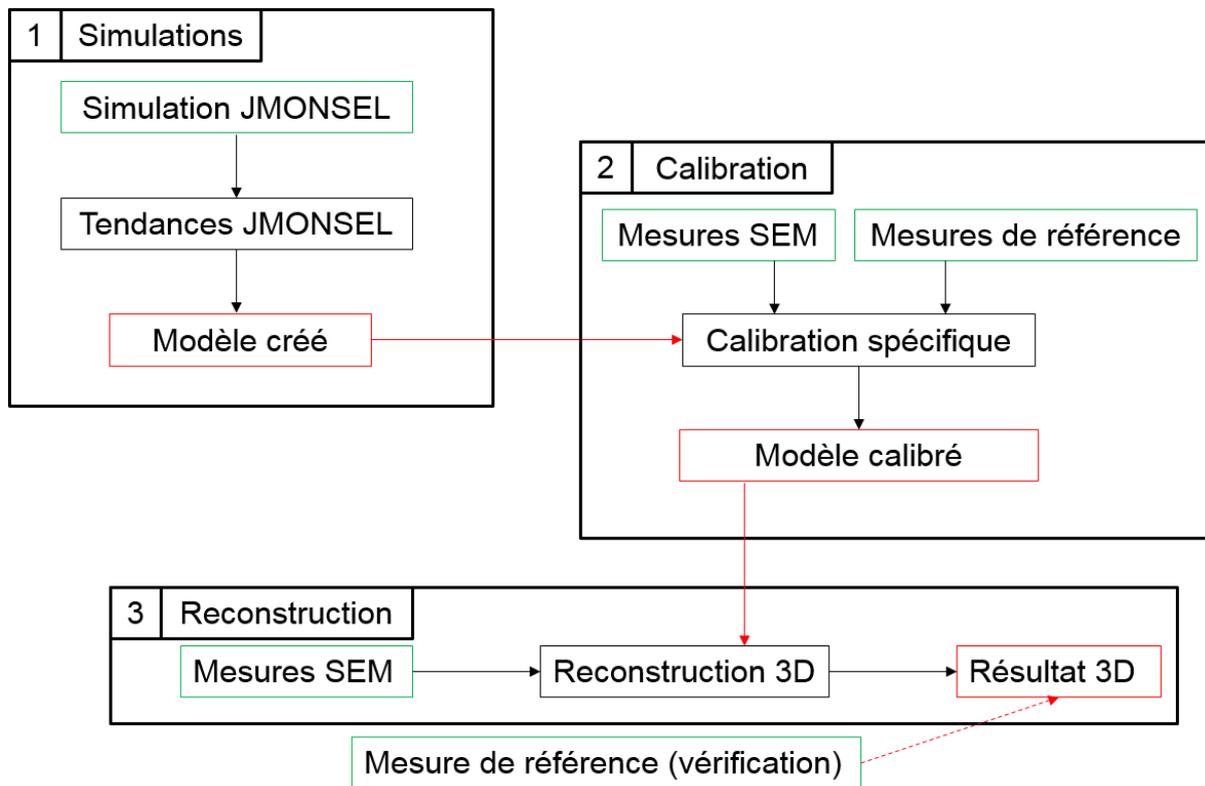
I. Organisation du code

Le code est organisé en trois parties distinctes, en langage python.

La première partie prend en entrée les simulations JMONSEL, et crée le modèle linéaire qui sera utilisé pour l'estimation des grandeurs d'intérêts.

La seconde partie prend en entrée le modèle linéaire issu de la simulation, ainsi que des mesures SEM et des mesures de référence (AFM par exemple) de ces mêmes motifs. Ce code permet de calibrer le modèle issu de la simulation pour coller avec les résultats expérimentaux.

La troisième partie, est la partie qui sera utilisée par un éventuel utilisateur final pour la reconstruction. Elle prend en entrée le modèle calibré et les images SEM des motifs que l'on souhaite reconstruire. En sortie on obtient l'estimation des différentes grandeurs d'intérêt. On peut éventuellement diversifier ce code pour faire de la statistique, des visualisations 3D etc...



Ces trois parties sont organisées de la même manière. Un fichier de configuration (sous forme d'un document texte) permet de choisir les paramètres que l'on souhaite. Ensuite le lancement du programme va exécuter les différentes tâches les unes à la suite des autres en enregistrant les données de manière régulière. Grâce à cela, si on souhaite seulement refaire tourner une partie du traitement, on peut toujours repartir des dernière données enregistrées et donc gagner beaucoup de temps.

Pour l'organisation des données on a choisi d'utiliser la bibliothèque "pandas", qui fonctionne à peu près comme un fichier Excel. C'est grâce à cette organisation des données qu'il a été possible de rapidement produire des graphiques en classant les données en fonction d'une grandeur ou d'une autre.

II. Stratégie de positionnement des descripteurs

La principale difficulté pour le positionnement des descripteurs est de réussir à trouver une stratégie permettant un positionnement robuste des descripteurs, quelque soit les caractéristiques dimensionnelles du motif (dans la limite du domaine de validité étudié) mais également quelque soit l'angle de tilt d'observation du motif.

Lignes isolées

La première étape du placement des descripteurs est de trouver le centre du motif. Pour cela on va chercher les deux pics de rendement électronique centraux. Grâce à la position du centre du motif, on peut travailler d'une part sur le flanc montant du motif (coté gauche) et d'autre part sur le flanc descendant du motif (coté droit).

Les descripteurs sont positionnés dans l'ordre suivant :

Bl : Position du maximum de dérivée du début au milieu.

Br : Position du minimum de dérivée du milieu à la fin.

Dl : Position du minimum de dérivée du début au milieu.

Dr : Position du maximum de dérivée du milieu à la fin.

Al : Position du dernier point de la dérivée inférieur à un niveau de coupure (constante arbitraire) du début au point Bl.

Ar : Position du premier point de la dérivée inférieur à un niveau de coupure (constante arbitraire) du point Br à la fin.

Cl : Position du dernier point de la dérivée supérieur à zéro du point Bl au point Dl.

Cr : Position du premier point de la dérivée inférieur à zéro du point Dr au point Br.

El : Position du premier point de la dérivée supérieur à un niveau de coupure (constante arbitraire) du point Dl au milieu.

Er : Position du dernier point de la dérivée inférieur à un niveau de coupure (constante arbitraire) du milieu au point Dr.

Sur la fin de la thèse le descripteur E a été redéfini pour essayer de réduire la variabilité du positionnement de E, et améliorer l'estimation du CRt notamment :

El : Position du premier point de la spline plus petit qu'un niveau de coupure (correspondant à une portion de la distance entre le maximum de la spline et le niveau de la spline au milieu) du point Dl au milieu.

Er : Position du dernier point de la spline plus petit qu'un niveau de coupure (correspondant à une portion de la distance entre le maximum de la spline et le niveau de la spline au milieu) du milieu au point Dr.

Tranchées isolées

Les descripteurs sont positionnés dans l'ordre suivant :

Dl : Position du maximum de dérivée du début au milieu.

Dr : Position du minimum de dérivée du milieu à la fin.

El : Position du dernier point de la dérivée plus petit qu'un niveau de coupure (constante arbitraire) du début au point Dl.

Er : Position du premier point de la dérivée plus grand qu'un niveau de coupure (constante arbitraire) du point Dr à la fin.

Cl : Position du premier point de la dérivée plus petit que zéro du point Dl au milieu.

Cr : Position du dernier point de la dérivée plus grand que zéro du milieu au point Dr.

Comme vu dans la partie sur l'analyse des motifs de type tranchées, les points A et B sont difficiles à placer car ils sont au fond de la tranchée, et les effets sont donc dépendants du facteur de forme. Différentes stratégies ont été testées pour leur positionnement, mais aucune n'a été suffisamment robuste. Contrairement aux autres définitions, celles-ci ne sont pas symétrique, c'est pourquoi elles marcheront uniquement pour un tilt augmentant le flanc de gauche. Si on inverse le sens du tilt il faudrait alors inverser les définitions et l'ordre de définition des flancs de droite et de gauche. Les définitions qui ont données les résultats présentés sont les suivantes :

Ar : Utilisation d'une fonction de détection de pic entre les points Cl et Cr, avec limitation sur la hauteur de recherche.

Br : Position du maximum de dérivée entre les points Ar et Cr.

Al : Position du premier point de la spline égale au minimum locale de la spline entre les points Cl et (Ar-marge), avec une marge arbitraire liée à la limite de facteur de forme vu en Eq 6.1.

Bl : Position du dernier point de la dérivée seconde plus petit qu'un niveau de coupure entre les points Cl et Cr.

III. Publications

Publication dans le journal à comité de lecture JM3 :

C.Valade, J.Hazart, S. Bérard-Bergery, E. Sungauer, M. Besacier, and C. Gourgon, "Tilted beam scanning electron microscopy, 3-D metrology for microelectronics industry", *Journal of Micro/Nanolithography, MEMS and MOEMS*, vol. 18, p. 1, July 2019.
doi : 10.1117/1.JMM.18.3.034001

Publication lors de la conférence SPIE Advanced lithography 2019 :

C.Valade, J.Hazart, S. Bérard-Bergery, E. Sungauer, M. Besacier, and C. Gourgon, "Tilted beam SEM, 3D metrology for industry", in *Metrology, Inspection, and process Control for Microlithography XXXIII*, vol. 10959 of *Proceedings SPIE*, (San Jose, United States) p. 32, SPIE, Feb. 2019.
doi : 10.1117/12.2514977

Bibliographie

- [1] M. S. Bronsgeest, J. E. Barth, L. W. Swanson, and P. Kruit, "Probe current, probe size, and the practical brightness for probe forming systems," *Journal of Vacuum Science & Technology B : Microelectronics and Nanometer Structures*, vol. 26, no. 3, p. 949, 2008.
- [2] J. R. Michael and D. B. Williams, "A consistent definition of probe size and spatial resolution in the analytical electron microscope," *Journal of Microscopy*, vol. 147, pp. 289–303, sep 1987.
- [3] M. M. Waldrop, "The chips are down for moore's law," *Nature*, vol. 530, pp. 144–147, feb 2016.
- [4] M. Neisser and S. Wurm, "ITRS lithography roadmap : 2015 challenges," *Advanced Optical Technologies*, vol. 4, jan 2015.
- [5] J. L. c. Le Coz, *Réduction de la consommation statique des circuits intégrés en technologie SOI 65 nm partiellement désertée*. Theses, Université de Grenoble, Nov. 2011.
- [6] M. Bruel, "Silicon on insulator material technology," *Electronics Letters*, vol. 31, no. 14, p. 1201, 1995.
- [7] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors," in *IEEE International Electron Devices Meeting 2003*, IEEE, 2003.
- [8] A. Hikavy, I. Zyulkov, H. Mertens, L. Witters, R. Loo, and N. Horiguchi, "Use of high order precursors for manufacturing gate all around devices," *Materials Science in Semiconductor Processing*, vol. 70, pp. 24–29, nov 2017.
- [9] B. Bunday, T. A. Germer, V. Vartanian, A. Cordes, A. Cepler, and C. Settens, "Gaps analysis for CD metrology beyond the 22nm node," in *Metrology, Inspection, and Process Control for Microlithography XXVII* (A. Starikov and J. P. Cain, eds.), SPIE, apr 2013.
- [10] R. Attota, B. Bunday, and V. Vartanian, "Critical dimension metrology by through-focus scanning optical microscopy beyond the 22 nm node," *Applied Physics Letters*, vol. 102, p. 222107, jun 2013.
- [11] B. D. Bunday, A. Bello, E. Solecky, and A. Vaid, "7/5nm logic manufacturing capabilities and requirements of metrology," in *Metrology, Inspection, and Process Control for Microlithography XXXII* (O. Adan and V. A. Ukraintsev, eds.), SPIE, mar 2018.
- [12] R. Micheloni, L. Crippa, C. Zambelli, and P. Olivo, "Architectural and integration options for 3d NAND flash memories," *Computers*, vol. 6, p. 27, aug 2017.
- [13] Site internet : <https://fr.wikipedia.org/wiki/Fichier:Interaction.jpg>.
- [14] D. Hussain, K. Ahmad, J. Song, and H. Xie, "Advances in the atomic force microscopy for critical dimension metrology," *Measurement Science and Technology*, vol. 28, p. 012001, nov 2016.
- [15] Entreprise : Orsay Physics. Site internet : <http://www.orsayphysics.com/fib>. Image : http://www.orsayphysics.com/upload/big_1527510474.jpg.

- [16] K. Takamasu, Y. Iwaki, S. Takahashi, H. Kawada, and M. Ikota, “3d-profile measurement of advanced semiconductor features by using FIB as reference metrology,” in *Metrology, Inspection, and Process Control for Microlithography XXXI* (M. I. Sanchez and V. A. Ukraintsev, eds.), SPIE, apr 2017.
- [17] C. Fang, M. D. Smith, A. V. Pret, J. J. Biafore, S. A. Robertson, and J. Bekaert, “A compact physical CD-SEM simulator for IC photolithography modeling applications,” in *Scanning Microscopies 2014* (M. T. Postek, D. E. Newbury, S. F. Platek, and T. K. Maugel, eds.), SPIE, sep 2014.
- [18] J.-L. Palomares, *Une nouvelle méthode d’appariement de points d’intérêt pour la mise en correspondance d’images*. Theses, Université Montpellier II - Sciences et Techniques du Languedoc, Oct. 2012.
- [19] N. Cornille, *Accurate 3D Shape and Displacement Measurement using a Scanning Electron Microscope*. Theses, INSA de Toulouse, June 2005.
- [20] A. Kudryavtsev, *3D Reconstruction in Scanning Electron Microscope : from image acquisition to dense point cloud*. Theses, Université Bourgogne Franche-Comté, Oct. 2017.
- [21] R. I. Hartley and P. Sturm, “Triangulation,” in *Computer Analysis of Images and Patterns*, pp. 190–197, Springer Berlin Heidelberg, 1995.
- [22] E. Mouragnon, *Realtime Structure from Motion by local bundle adjustment*. Theses, Université Blaise Pascal - Clermont-Ferrand II, Dec. 2007.
- [23] V. N. Tondare, J. S. Villarrubia, and A. E. Vladar, “Three-dimensional (3d) nanometrology based on scanning electron microscope (SEM) stereophotogrammetry,” *Microscopy and Microanalysis*, vol. 23, pp. 967–977, sep 2017.
- [24] J. S. Villarrubia, V. N. Tondare, and A. E. Vladár, “Virtual rough samples to test 3d nanometer-scale scanning electron microscopy stereo photogrammetry,” in *Metrology, Inspection, and Process Control for Microlithography XXX* (M. I. Sanchez and V. A. Ukraintsev, eds.), SPIE, mar 2016.
- [25] M. Eulitz and G. Reiss, “3d reconstruction of SEM images by use of optical photogrammetry software,” *Journal of Structural Biology*, vol. 191, pp. 190–196, aug 2015.
- [26] M. Tanaka, C. Shishido, Y. Takagi, H. Morokuma, O. Komuro, and H. Mori, “Cross-sectional gate feature identification method using top-down SEM images,” in *Metrology, Inspection, and Process Control for Microlithography XVII* (D. J. Herr, ed.), SPIE, may 2003.
- [27] X. Zhang, H. Zhou, Z. Ge, A. Vaid, D. Konduparthi, C. Osorio, S. Ventola, R. Meir, O. Shoval, R. Kris, O. Adan, and M. Bar-Zvi, “Addressing FinFET metrology challenges in 1x node using tilt-beam CD-SEM,” in *Metrology, Inspection, and Process Control for Microlithography XXVIII* (J. P. Cain and M. I. Sanchez, eds.), SPIE, apr 2014.
- [28] T. Marschner and C. Stief, “Characterization of 193-nm resist layers by CD-SEM sidewall imaging,” in *Metrology, Inspection, and Process Control for Microlithography XVII* (D. J. Herr, ed.), SPIE, may 2003.
- [29] B. Su, R. Oshana, M. Menaker, Y. Barak, and X. Shi, “Shape control using sidewall imaging,” in *Metrology, Inspection, and Process Control for Microlithography XIV* (N. T. Sullivan, ed.), SPIE, jun 2000.
- [30] H. Morokuma, A. Miyamoto, M. Tanaka, M. Kazui, and A. Takane, “New technique to reconstruct effective 3d profile from tilt images of CD-SEM,” in *Metrology, Inspection, and Process Control for Microlithography XVIII* (R. M. Silver, ed.), SPIE, may 2004.
- [31] E. Prados and O. Faugeras, “Shape from shading : A well-posed problem?,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05)*, IEEE, 2005.

- [32] A. Meyer, H. M. Briceño, and S. Bouakaz, “User-guided shape from shading to reconstruct fine details from a single photograph,” in *Computer Vision – ACCV 2007* (Y. Yagi, S. B. Kang, I. S. Kweon, and H. Zha, eds.), (Berlin, Heidelberg), pp. 738–747, Springer Berlin Heidelberg, 2007.
- [33] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, “The bas-relief ambiguity,” *International Journal of Computer Vision*, vol. 35, no. 1, pp. 33–44, 1999.
- [34] W. Drzazga, J. Paluszynski, and W. Slowko, “Three-dimensional characterization of microstructures in a SEM,” *Measurement Science and Technology*, vol. 17, pp. 28–31, nov 2005.
- [35] W. Slówko, “Directional detection of secondary electrons for electron beam profilography,” *Vacuum*, vol. 52, pp. 441–449, apr 1999.
- [36] H. Fukaya, T. Murakawa, S. Shida, M. Kuribara, T. Iwai, J. Matsumoto, T. Nakamura, H. Hakii, I. Yonekura, M. Kawashita, Y. Nishiyama, K. Tanaka, and Y. Kikuchi, “New CD-SEM metrology method for the side wall angle measurement using multiple detectors,” in *Photomask and Next-Generation Lithography Mask Technology XVIII* (T. Konishi, ed.), SPIE, apr 2011.
- [37] W. Ito, B. Bunday, S. Harada, A. Cordes, T. Murakawa, A. Arceo, M. Yoshikawa, T. Hara, T. Arai, S. Shida, M. Yamagata, J. Matsumoto, and T. Nakamura, “Novel three dimensional (3d) CD-SEM profile measurements,” in *Metrology, Inspection, and Process Control for Microlithography XXVIII* (J. P. Cain and M. I. Sanchez, eds.), SPIE, apr 2014.
- [38] A. Blake, A. Zisserman, and G. Knowles, “Surface descriptions from stereo and shading,” *Image and Vision Computing*, vol. 3, pp. 183–191, nov 1985.
- [39] Y. Leclerc and A. Bobick, “The direct computation of height from shading,” in *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Comput. Soc. Press, 1991.
- [40] C. Wu, B. Wilburn, Y. Matsushita, and C. Theobalt, “High-quality shape from multi-view stereo and shading under general illumination,” in *CVPR 2011*, IEEE, jun 2011.
- [41] J. S. Villarrubia, N. W. M. Ritchie, and J. R. Lowney, “Monte carlo modeling of secondary electron imaging in three dimensions,” in *Metrology, Inspection, and Process Control for Microlithography XXI* (C. N. Archie, ed.), SPIE, mar 2007.
- [42] J. S. Villarrubia and Z. J. Ding, “Sensitivity of SEM width measurements to model assumptions,” in *Metrology, Inspection, and Process Control for Microlithography XXIII* (J. A. Allgair and C. J. Raymond, eds.), SPIE, mar 2009.
- [43] D. C. Joy and S. Luo, “An empirical stopping power relationship for low-energy electrons,” *Scanning*, vol. 11, no. 4, pp. 176–180, 1989.
- [44] T. Verduin, *Quantum Noise Effects in e-Beam Lithography and Metrology*. PhD thesis, 2017.
- [45] D. Nyyssonen, “A new approach to image modeling and edge detection in the SEM,” in *Integrated Circuit Metrology, Inspection, and Process Control II* (K. M. Monahan, ed.), SPIE, jan 1988.
- [46] S. Labbé, T. Faouzi, J. Belissard, and J. HAZART, “Limits of model-based CD-SEM metrology,” in *34th European Mask and Lithography Conference* (U. F. Behringer and J. Finders, eds.), SPIE, sep 2018.
- [47] A. Seeger, *Surface Reconstruction from Afm and Sem Images*. PhD thesis, Chapel Hill, NC, USA, 2004. AAI3156203.
- [48] P. C. Sabatier, “Past and future of inverse problems,” *Journal of Mathematical Physics*, vol. 41, pp. 4082–4124, jun 2000.

- [49] C. G. Frase, E. Buhr, and K. Dirscherl, “CD characterization of nanostructures in SEM metrology,” *Measurement Science and Technology*, vol. 18, pp. 510–519, jan 2007.
- [50] C. A. Mack and B. D. Bunday, “Improvements to the analytical linescan model for SEM metrology,” in *Metrology, Inspection, and Process Control for Microlithography XXX* (M. I. Sanchez and V. A. Ukraintsev, eds.), SPIE, mar 2016.
- [51] A. Seeger and H. Haussecker, “Shape-from-shading and simulation of SEM images using surface slope and curvature,” *Surface and Interface Analysis*, vol. 37, no. 11, pp. 927–938, 2005.
- [52] M. S. Chung and T. E. Everhart, “Simple calculation of energy distribution of low-energy secondary electrons emitted from metals under electron bombardment,” *Journal of Applied Physics*, vol. 45, pp. 707–709, feb 1974.
- [53] T. Koshikawa and R. Shimizu, “A monte carlo calculation of low-energy secondary electron emission from metals,” *Journal of Physics D : Applied Physics*, vol. 7, pp. 1303–1315, jun 1974.
- [54] R. Penrose and J. A. Todd, “A generalized inverse for matrices,” *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, p. 406, jul 1955.
- [55] M. A. . Efroymson, “Multiple regression analysis,” *Mathematical Methods for Digital Computers, Ralston A. and Wilf, H. S., (eds.), Wiley, New York, 1960.*
- [56] R. R. Hocking and R. R. Hocking, “A biometrics invited paper. the analysis and selection of variables in linear regression,” *Biometrics*, pp. 1–49, 1976.
- [57] S. Derksen and H. J. Keselman, “Backward, forward and stepwise automated subset selection algorithms : Frequency of obtaining authentic and noise variables,” *British Journal of Mathematical and Statistical Psychology*, vol. 45, pp. 265–282, nov 1992.
- [58] B. Efron, R. Tibshirani, I. Johnstone, and T. Hastie, “Least angle regression,” *The Annals of Statistics*, vol. 32, pp. 407–499, apr 2004.
- [59] M. Hebiri, *Some variable selection procedures based on the LASSO estimator*. Theses, Université Paris-Diderot - Paris VII, June 2009.
- [60] H. Zou, T. Hastie, and R. Tibshirani, “On the “degrees of freedom” of the lasso,” *The Annals of Statistics*, vol. 35, pp. 2173–2192, oct 2007.
- [61] J. Shao, “An asymptotic theory for linear model selection,” *Statistica Sinica*, vol. 7, no. 2, pp. 221–242, 1997.
- [62] Scikit learn. Site internet : https://scikit-learn.org/stable/modules/linear_model.html.
- [63] Mayavi : 3D scientific data visualization and plotting in Python. Site internet : <https://docs.enthought.com/mayavi/mayavi/>.
- [64] Scipy interpolate. Site internet : <https://docs.scipy.org/doc/scipy-0.18.1/reference/interpolate.html>.
- [65] P. Dierckx, “An algorithm for smoothing, differentiation and integration of experimental data using spline functions,” *Journal of Computational and Applied Mathematics*, vol. 1, pp. 165–184, sep 1975.
- [66] E. P. Solecky, C. N. Archie, J. Mayer, R. S. Cornell, and O. Adan, “Characterizing and understanding stray tilt : the next major contributor to CD SEM tool matching,” in *Metrology, Inspection, and Process Control for Microlithography XVII* (D. J. Herr, ed.), SPIE, may 2003.
- [67] K. T. Arat, J. Bolten, A. C. Zonneville, P. Kruit, and C. W. Hagen, “Estimating step heights from top-down SEM images,” *Microscopy and Microanalysis*, vol. 25, pp. 903–911, may 2019.

Résumé

Avec l'avancée des technologies de la microélectronique, l'architecture des composants électroniques devient de plus en plus compliquée. Or, la connaissance des caractéristiques dimensionnelles des structures réalisées est importante pour pouvoir comprendre et optimiser le comportement de ces composants. C'est pourquoi il existe un besoin de développer des méthodes de mesure tridimensionnelles rapides et non destructives.

Le microscope électronique à balayage (SEM) est largement utilisé pour réaliser des mesures dimensionnelles car il répond aux problématiques de rapidité et de non-destructivité. Cependant l'obtention d'informations tridimensionnelles quantitatives et précises est un challenge.

Grâce à un microscope électronique dont le faisceau électronique peut être incliné, il est possible d'obtenir des images à différents angles de vue. A partir de l'analyse de ces images, la hauteur et l'angle des flancs du motif observé peuvent être déterminés géométriquement.

Cependant, l'imagerie électronique étant le résultat des interactions électrons-matière, il est important de comprendre l'origine de la formation des images SEM, pour pouvoir les analyser correctement. C'est pourquoi une étude a été menée grâce à un logiciel de simulation physique pour observer et comprendre l'impact de la topographie d'un motif sur l'image SEM résultante.

A partir de ces observations, des métriques ont été créées sur les images SEM pour les analyser quantitativement. Un modèle linéaire a ensuite été créé grâce aux simulations physiques pour estimer les grandeurs topographiques à partir de ces métriques. Il a ensuite été calibré sur des mesures SEM réelles, en les comparant à des mesures tridimensionnelles de référence par microscopie à force atomique (AFM). Ce modèle a été créé pour la reconstruction de motifs de type "ligne" en silicium gravé. Grâce à ce modèle, des reconstructions de motifs réelles ont été réalisées. Enfin un travail sur la création d'un modèle pour les motifs de type "tranchée" et "dense" en silicium gravé a été initié.

Abstract

With the advancement of microelectronics technologies, the architecture of electronic components is becoming increasingly complicated. However, knowledge of the dimensional characteristics of the structures is important in order to be able to understand and optimize the behavior of these components. This is why there is a need to develop rapid, non-destructive three-dimensional measurement methods.

The scanning electron microscope (SEM) is widely used to carry out dimensional measurements because it responds to the problems of speed and non-destructivity. However, obtaining quantitative and precise three-dimensional information is a challenge.

Thanks to an electron microscope whose electron beam can be tilted, it is possible to obtain images at different viewing angles. From the analysis of these images, the height and the sidewall angles of the observed pattern can be determined geometrically.

However, since electronic imaging is the result of electron-matter interactions, it is important to understand the origin of the formation of SEM images, in order to be able to analyze them correctly. This is why a study was carried out using physical simulation software to observe and understand the impact of the topography of a pattern on the resulting SEM image.

From these observations, metrics were created on the SEM images to analyze them quantitatively. A linear model was then created using physical simulations to estimate the topographic quantities from these metrics. It was then calibrated on real SEM measurements, by comparing them to three-dimensional reference measurements by atomic force microscopy (AFM). This model was created for the reconstruction of "line" type patterns in etched silicon. Thanks to this model, reconstructions of real patterns were made. Finally, work was started on the creation of a model for "trench" and "dense" type patterns in etched silicon.